IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 21 July 2022, accepted 1 August 2022, date of publication 5 August 2022, date of current version 10 August 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3196677

RESEARCH ARTICLE

Practical Side-Channel Attack on Free-Space QKD Systems With Misaligned Sources and Countermeasures

PABLO ARTEAGA-DÍAZ, DANIEL CANO, AND VERONICA FERNANDEZ
Institute of Physical and Information Technologies, Spanish National Research Council(CSIC), 28006 Madrid, Spain

Corresponding author: Pablo Arteaga-Díaz (pablo.arteaga@csic.es)

This work was supported in part under Grant PID2020-118178RB-C22 and Grant AEI/10.13039/501100011033; in part by the Community of Madrid, Spain, under the Cybersecurity, Network Analysis And Monitoring for the next generation Internet (CYNAMON) Project, under Grant P2018/TCS-4566; in part by the European Social Fund and European Union (EU) Fondo Europeo de Desarrollo Regional (FEDER) funds; in part by the Spanish National Research Council (CSIC), under Project 202050E232; in part by the CSIC's Interdisciplinary Thematic Platform (PTI+) on Quantum Technologies (PTI-QTEP+); in part by the CSIC's Program for the Spanish Recovery, Transformation and Resilience Plan through the Recovery and Resilience Facility of the EU, established by the Regulation (EU) 2020/2094; and in part by the Ministerio de Ciencia e Innovación (MCIN) from EU NextGenerationEU under Grant PRTR-C17.I1.

ABSTRACT Practical implementations of quantum key distribution (QKD) protocols can introduce additional degrees of freedom in the quantum states that may render them distinguishable to an eavesdropper. This is the case of QKD systems using a different laser source to generate each quantum state, which can lead to temporal, spectral and/or spatial differences among them that can be exploited by a malicious party to extract information of the key. In this work we characterize, and experimentally verify, a side-channel attack on spatially distinguishable states against free-space QKD systems with misaligned laser sources. Specifically, for those emitting Gaussian beams, which is the most common case in free-space QKD. The attack makes theoretically unsafe any QKD system with any angular misalignment between the laser sources. Finally, we propose two countermeasures to eliminate the spatial distinguishability and secure the key exchange.

INDEX TERMS Countermeasures, free-space QKD, quantum key distribution, side-channel attack, spatial distinguishability.

I. INTRODUCTION

Quantum key distribution (QKD) is an application of quantum information theory that allows two authenticated distant parties to exchange a cryptographic key with unconditional security [1], [2]. The security of most QKD protocols is based on the fact that non-orthogonal quantum states are not distinguishable [3], [4], or at least not without a loss of information [5]. Thus, encoding the information of the key using non-orthogonal quantum states, it is possible to guarantee the secrecy of the key transmission in different QKD protocols [6], [7]. Nevertheless, in the physical implementation of these protocols, additional degrees of freedom not considered in theory may appear, making the states distinguishable, and jeopardizing the security of the key exchange [8]. Attacks that take advantage of these additional degrees of freedom are called side-channel attacks. There is a wide variety of possible side-channel attacks to QKD systems at device level, such as attacks to the sources and detectors [9]–[13], and attacks to the physical implementation of some post-processing steps [14]. Further, there are side-channel attacks to the signals traveling between the QKD terminals, which will be addressed in this work. There are two different approaches to protect QKD against side-channel attacks. The first approach is to use protocols in which the security does not depend on the implementation, such as device independent (DI) QKD [15]. The second approach is to detect and close all the possible security loopholes caused by side-channels. In order to do that, we must take into account all possible side-channel attacks and either characterize the information leakage they may cause and consider it in the privacy amplification step, or design countermeasures by hardware or software modifications to remove the side-channel. The best approach depends on the application, the users'

The associate editor coordinating the review of this manuscript and approving it for publication was Luca Barletta.

security needs, and budget restrictions. On the receiving station, both approaches have been used: detecting and closing security loopholes due to side-channels [10], [16]–[20], and shielding the receiver from any side-channel attack with measurement-device independent (MDI) QKD [21], [22]. Different QKD protocols based on the principles of quantum interference, such as the original MDI [21], [22], twin-field (TF) [23], and sending-or-not-sending (SNS) [24] protocols have been proposed and experimentally demonstrated [25]–[32], achieving side-channel-free receiving terminals. However, protecting the transmitting station is a more challenging task, since the implementation of DI-QKD has shown some practical limitations. Thus, to protect QKD against side-channel attacks to the transmitter terminal and the transmitted signals, the most extended approach is to detect and close all the possible security loopholes caused by side-channels [9], [16], [20], [33]–[36]. Therefore, we chose the latter approach to characterize a possible side-channel attack that can be targeted to many free-space QKD transmitters implemented to date, and propose countermeasures to protect them.

The polarization of light is one of the most extended physical observables used to encode the information of the keys in prepare-and-measure QKD with free-space or atmospheric transmission channels. To generate the states of polarization in protocols such as B92 [37] and BB84 [38], it is widespread to use a different laser source for each quantum state [39]–[46]. This design with multiple lasers in the transmitter, Alice, has the advantage of being able to obtain very stable states of polarization over long periods of time. Specifically, those using free-space polarizers and not any other fiber-optic component after the states are polarized, since optical fibers may induce polarization variations in the states dependent on temperature and pressure, making them unstable over time. Furthermore, the use of passive elements for the codification of the information of the key avoids some Trojan-horse attacks [10] and simplify the electronics to control the sources. However, the use of multiple lasers introduces different degrees of freedom that could be used to distinguish the sources. For these cases, side-channel attacks that take advantage of the spectral, temporal and/or spatial distinguishability of the states generated by the different lasers have been proposed [9], [33], [35], [36], [47]. Since each quantum state is generated with a different laser source, it is possible to distinguish the states by their wavelength spectrum, by disparity in emission times, or by the possibility of spatially separating the states to identify them. For the aforementioned attacks, some countermeasures have been proposed too [9], and secure QKD systems against them have been demonstrated [33], [43].

In this work, we will focus on a practical attack that takes advantage of the spatial distinguishability of quantum states due to angular misalignment between the laser sources, as this topic is not fully covered in the references mentioned above. More specifically, we analyze an attack that uses an optical system to discriminate beams with different angles of arrival, exploiting the fact that each one of them is focused in a different area of the focal plane. With this strategy, the attacker can distinguish the states depending on the area of the focal plane they are focused. The description of the attack is extended in the methods section (II). In [33] they propose a similar attack considering point sources, spherical waves, and a transmitter that truncates the beams introducing diffraction effects. However, this is not the most general or optimal design, since spherical waves are unusual in QKD systems, and the transmitter usually does not truncate the beams since it increases diffraction losses, which is a critical parameter that should be minimized in long distance QKD links. Further, truncating the beam is not an effective countermeasure against an attacker with infinite capabilities, which is generally considered in QKD security proofs. We thus consider that the transmitter emits Gaussian beams, which is the most common case in free-space QKD systems, and that its exit aperture does not introduce truncation effects on the Gaussian beams to reduce diffraction losses. We first assume free-space as the transmission channel without atmospheric turbulence and no pointing error. Nevertheless, the results could be applied to the case of atmospheric transmission with wavefront distortion and pointing errors considering that the attacker corrects those effects with adaptive optics and beam stabilization technologies. With these assumptions, the attack makes theoretically unsafe any QKD system that generates the different states with angularly misaligned laser sources, which could be the case of many experimental QKD transmitters like the Micius satellite [42]. We analytically characterize the maximum information that an attacker can extract depending on the angular misalignment between the optical beams, the wavelength of the quantum signal, and the beams' radii (sections II-A and III). Further, we experimentally verify the possibility of performing the attack (sections II-C and III). Finally, we propose two countermeasures to protect the system and discuss their advantages and disadvantages (section III-A).

II. METHODS

A. ATTACK MODELLING

Beams with different angles of arrival upon reception on an optical system are focused on different points at its focal plane. From this idea, we propose a practical side-channel attack in which the eavesdropper, Eve, uses an optical system to discriminate the source generating each quantum state, and thus, determine the codification of the key. For the attack characterization, we assume two beams with the same wavelength and beam radius, and an angular divergence between them of $\Delta\theta$. To spatially distinguish the beams at the focal plane, Eve can use an array of single-photon detectors. We will assume a simpler design (see figure 1) in which she divides her focal plane into two areas, and that she measures all the photons that fall in each area with a different measurement base. In the case of the BB84 protocol, these bases would be the Z (rectilinear) and the X (diagonal) one. This can be done by placing a wedge-shaped mirror that splits the beam in two portions, deflecting each towards a measurement base. We will consider the most beneficial case for the attacker in which the states of each base are aligned, and those of different bases have an angular divergence $\Delta\theta$. In other cases, the same procedure could be used, taking into account the divergence between each state and the two states of other base. Under certain circumstances that we will now analyze, the spatial separation of the beams at the focal plane, Δx, could be greater than twice their radius at the focal plane, w_f, such that most of the photons from each beam are measured by the attacker with the basis on which they were prepared. Measuring the states with the base in which they were prepared allows an unambiguous discrimination of the states. After measuring, the attacker generates the measured states and resends them to the original receiver of the QKD link.

To separate two optical beams with a certain angular divergence, $\Delta\theta = \theta_2 - \theta_1$, we can use an optical system with an effective focal length, f. As each beam reaches the system with a different angle of incidence, θ_1 and θ_2, each one will focus on a different point at the focal plane. The position of the centroid of each beam at the focal plane will be $x_i = f \cdot tan(\theta_i)$, being i = 1,2 [48]. The centroids of the beams will be separated by a distance $\Delta x = x_2 - x_1$ at the focal plane, which is $\Delta x = f(tan\theta_2 - tan\theta_1)$. For small angles of incidence: $tan(\theta_i) \approx \theta_i$, obtaining that

$$\Delta x \approx f(\theta_2 - \theta_1) = f\Delta\theta. \tag{1}$$

We will consider that collimated Gaussian beams are emitted with wavelength λ at the output of the QKD transmitter. Furthermore, at a certain distance z from the transmitter, the beams will have a radius w_z. We use the beam radius definition as the distance to the center of the beam at which the optical intensity drops by $exp(-2)$. If an attacker focuses these beams using an optical system with an effective focal length f, the radius of the beams at the focal plane will be at least

$$w_f = \lambda f/\pi w_z. \tag{2}$$

Equation (2) defines the radius of a focused Gaussian beam at the focal plane, assuming that the aperture of the attacker's optical system is infinitely large [49], that is to say, neglecting the effects of truncation. In practice, this does not vary significantly if the radius of the aperture is at least twice the radius of the beam [50]. Using equations (1) and (2), and dividing one by the other, we obtain

$$\Delta x/w_f = \pi w_z \Delta\theta/\lambda. \tag{3}$$

The quotient between the radius of the beams, w_f, and the distance between them, Δx, at the focal plane does not depend directly on the focal length of the optical system, but on w_z, $\Delta\theta$ and λ. As we can see in figure 1, and we will show more rigorously below, this quotient is what determines the amount of information that an attacker can obtain. We can calculate the information that the attacker obtains by considering the irradiance pattern of a Gaussian beam, and calculating the total power that is measured in the correct base. For a Gaussian beam, the irradiance distribution at the focal plane is Gaussian too [51]:

$$I(x, y) = I_0 exp(-2[x^2 + y^2]/w_f^2), \tag{4}$$

being I_0 the maximum irradiance. In the considered case in which we try to split the beams, we would have that the pattern is $I_1 = I(x + \Delta x/2, y)$ for the beam displaced to the negative x values, and $I_2 = I(x - \Delta x/2, y)$ for the beam displaced to the positive x values. The total power is

$$P_{tot} = 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I(x, y)dxdy = \pi w_f^2 I_0. \tag{5}$$

whereas the power that falls on the correct half of the focal plane, in which its measurement base is located, is

$$P_{corr} = \int_{-\infty}^{\infty}\int_{-\infty}^{0} I(x + \Delta x/2, y)dxdy + \int_{-\infty}^{\infty}\int_{0}^{\infty} I(x - \Delta x/2, y)dxdy, \tag{6}$$

which, by symmetry is

$$P_{corr} = 2\int_{-\infty}^{\infty}\int_{-\infty}^{0} I(x + \Delta x/2, y)dxdy \tag{7}$$

With these powers, the probability that the attacker measures the photons with the correct base is

$$p_c = P_{corr}/P_{tot}. \tag{8}$$

Using equations (4), (5), (7) and (8), and making the variable changes: $x + \Delta x/2 = x'$, $x' = n_x w_f$ and $y = n_y w_f$ we obtain

$$p_c = (2/\pi)\int_{-\infty}^{\infty}\int_{-\infty}^{\Delta x/2w_f} exp(-2(n_x^2 + n_y^2))dn_x dn_y. \tag{9}$$

Solving the integral (9) we get that

$$p_c(\Delta x/w_f) = (1/2)[erf((1/\sqrt{2})\Delta x/w_f) + 1]. \tag{10}$$

We can see how the probability that the attacker measures with the correct base depends on the quotient of $\Delta x/w_f$. Using equation (3), we can calculate $\Delta x/w_f$ as a function of w_z, $\Delta\theta$ and λ, and substitute it into (10) to calculate p_c, so

$$p_c(w_z, \Delta\theta, \lambda) = (1/2)[erf((\pi/\sqrt{2})w_z\Delta\theta/\lambda) + 1]. \tag{11}$$

In addition to measurements with the correct base, half of the measurements with the wrong base give correct measurement results. Thus, the probability of a correct measurement result is

$$p_{cm} = p_c + 0.5(1 - p_c) = 0.5(1 + p_c). \tag{12}$$

Finally, considering that the attacker measures all the photons, the probability that the attacker uses the wrong base and that the resent state generates an error in Bob is

$$p_{err} = 0.5(1 - p_c). \tag{13}$$

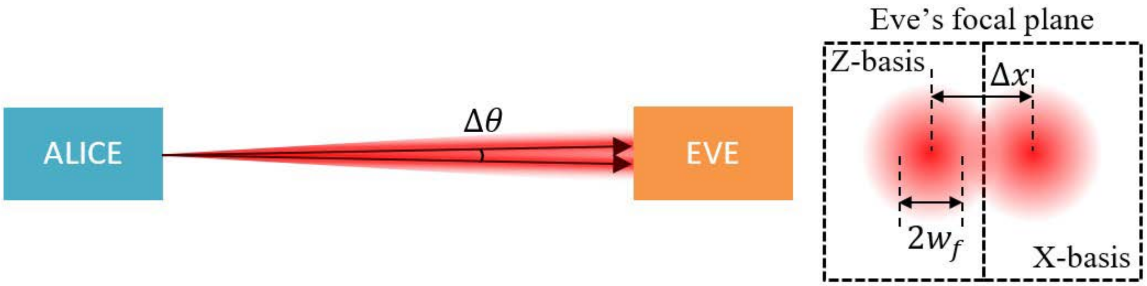

FIGURE 1. Alice transmits two beams with some divergence $\Delta\theta$, Eve intercepts the signal and uses an optical system to separate the beams at her focal plane. The demarcated areas within Eve's focal plane mark the division for each measurement base (Z and X). The parameters w_f and Δx are the beam radius and the distance between the beam centroids at the focal plane, respectively.

B. REALISTIC CHANNEL AND POINTING ERROR MODELLING

We have modelled the attack described in this work for the case of Gaussian beams, which is the case for a transmission channel without the presence of atmospheric turbulence and without considering pointing errors generated by the transmitter of the QKD system. Considering both effects, they will produce a broadening and a random movement of the beams at the focal plane, thus affecting the probability of obtaining correct and incorrect measurements. However, in QKD security proofs, an attacker with unlimited resources (compatible with the laws of Physics) is commonly assumed to guarantee unconditional security. This means that we must assume that the attacker has access to ideally perfect correction systems. The attacker can thus use both adaptive optics and beam stabilization systems to correct the wavefront distortion and stabilize the positions of the beams at the focal plane, respectively. Therefore, with these assumptions, the developed model here is still valid for a realistic channel. Additionally, we must consider that, in the case of a transmission channel with atmospheric turbulence, the attacker must use a larger aperture for the optical system to distinguish the states, since the long-term beam irradiance distribution at the aperture is broadened by atmospheric turbulence and pointing errors. We can model the long-term irradiance profile of a gaussian beam that has been propagated through atmospheric turbulence according to [52]. Assuming small pointing errors, we can approximate the results to those of [53]. This approximation assumes a Gaussian profile with radius:

$$w_S = \sqrt{w_{LT}^2 + (8\ln 2)\sigma_\theta^2 z^2}, \tag{14}$$

being w_{LT} the long-term beam radius due to atmospheric turbulence, σ_θ the standard deviation of the angular pointing error, and z the distance from the transmitter. Finally, we can also model w_{LT} according to [54] depending on the atmospheric turbulence conditions. In brief, if we consider a realistic channel with atmospheric turbulence and pointing errors, the attacker must use an optical system with an aperture of radius at least twice of w_S, which is larger than the radius of the beam in the case of ideal free-space propagation. However, if we assume the attacker is capable of correcting the wavefront and compensating the pointing errors, this does not change the information obtained by the attack, since it is determined by the radius of the corrected beam, w_z, and not w_S according to our model.

C. EXPERIMENTAL VERIFICATION OF THE ATTACK

To demonstrate a proof-of-principle of the attack, we have used the experimental setup shown in figure 2. We have measured the power of a focused beam by a lens, and we have varied the angle of incidence of the beam to the lens with a voice-coil-driven steering mirror. We used different angular deviations of the beam to simulate the values of angular divergence in the attack. We have used a photodetector with a round surface and covered half of it with a piece of metal, so the effective surface exposed to the light beam has a straight edge. The measured power with the photodetector corresponds to the measurements in the correct base, and the blocked power by the metal piece corresponds to the wrong base. We only measure the power of one beam and assume that the other one is symmetric. The diameter of the surface of the photodetector is more than ten times greater than that of the focused beam, which makes the power that falls outside the detector negligible, and therefore, the integration to infinity in equation (7) is a good approximation. The diameter of the lens was 70 mm, the focal length of the lens was 300 mm, the radius of the collimated beam was 3 mm, and the wavelength of the laser was 850 nm. The reason why we chose 850 nm as the wavelength is because it is widely used in free-space QKD systems. To assess the experimental verification, we measured the power on the exposed surface of the photodetector for each angle of deviation set by the steering mirror. Dividing each measured power by the total power of the beam as in equation (8), we obtained the probability of measuring with the correct base, p_c, as a function of $\Delta\theta$. The steering mirror has an internal optical sensor that we have used to obtain the angle of deviation of the beam. The deviation of the beam generated by the steering mirror is equivalent to half the angular divergence, $\Delta\theta$, in the proposed attack (see figure 2).

III. RESULTS AND DISCUSSION

In order to show the probability that Eve measures with the correct bases using the proposed attack, in figure 3 we

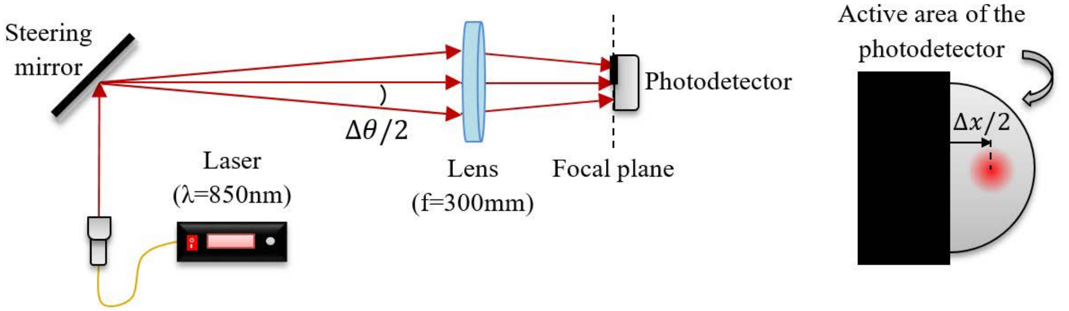

FIGURE 2. Schematic diagram of the experimental setup. A laser generates a continuous 850 nm-wavelength beam which is collimated and transmitted to the steering mirror, the steering mirror deviates the beam with different angles, and the lens focuses the beam on a photodetector with half of its surface covered.

show this probability calculated as a function of the quotient $\Delta x/w_f$ with equation (10). It can be seen that if the separation between the beams at the focal plane is $\Delta x = 0$, the probability is 50%, which agrees with that obtained in a conventional intercept-and-resend attack in which a measurement base is randomly chosen each time. As Δx increases with respect to the beams' radii at the focal plane, w_f, the beams get further and further apart and the probability of measuring with the correct base increases. For $\Delta x/w_f > 3$ the probability p_c is practically 100%.

Assuming a wavelength $\lambda = 850nm$, figure 4 represents p_c calculated with equation (11) as a function of $\Delta\theta$ for the different cases of w_z indicated in the legend. Note that the horizontal axis is in logarithmic scale so all the traces can be shown clearly, but the behavior in arithmetic scale is the same as the results in figure 3. With these results, we can get an idea of the information that Eve obtains with this attack, and the diameter of the telescope she needs, $4w_z$, as a function of $\Delta\theta$. We can see how the attack obtains more information for the case of larger beam sizes and larger angular divergences.

Finally, we e have experimentally verified the attack. In figure 5, the probability that Eve measures in the correct base, p_c, is represented versus the angular divergence between the beams, $\Delta\theta$. The dots are the experimental results, and the dashed line is the simulation calculated with equation (11).

The behavior of the experimental results of p_c represented in figure 5 agrees with the theoretical predictions, although some values are slightly lower than those of the simulation. This could be due to some imperfections in the experimental setup. Placing the photodetector at the exact focal plane with high precision is not always easy. For instance, if the photodetector is placed slightly out of focus, we obtain results with less probability of a correct measurement. This is due to $\Delta x/w_f$ being maximum at the focal plane. In addition, if the displacement of the beam is not perpendicular to the vertical axis defined by the edge between the metal piece and the active area of the photodetector, the irradiance obtained with each displacement is less than the expected one. Furthermore, the lens can introduce optical aberrations that increase the size of the beam at the focal plane, also reducing p_c. Thus, we find different practical aspects that could reduce the theoretical probability of distinguishing the states in a physical implementation of the attack. However, despite the fact that the implementation can be improved, the results still show high values of p_c that compromise the security of QKD systems.

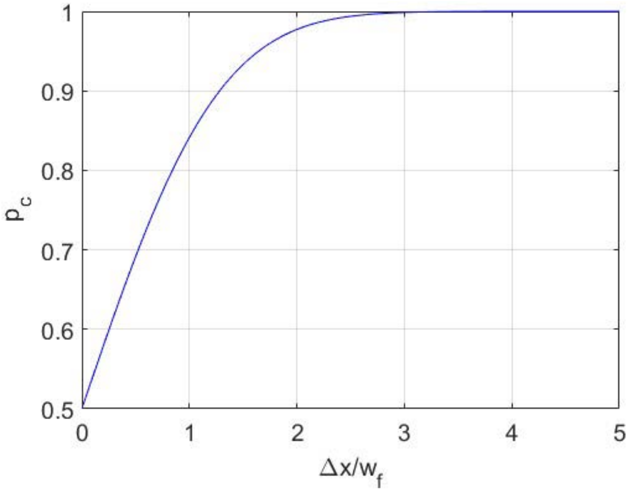

FIGURE 3. Probability, p_c, that the attacker measures with the correct base as a function of the quotient, $\Delta x/w_f$, being Δx the distance between focused beam centroids and w_f the radius of the beams at the focal plane of the attacker.

According to equation (11), if there were some methods of increasing the product $w_z \cdot \Delta\theta$, this would allow an attacker to obtain more information of the key. This is the case of Gaussian beam propagation, in which diffraction increases the beam size without varying $\Delta\theta$, thus increasing $w_z \cdot \Delta\theta$. That is to say, the attack is more effective in long distance links, and thus, we need to be specially careful in applications, such as satellite QKD. In fact, an attacker could modify the optical path of the beams between Alice and Bob and transmit them longer distances to arbitrarily increase their sizes before focusing them, thus extracting all the information of the key. Therefore, in theory, the information that Eve can obtain is not limited. Although we have considered for the analysis a simple case in which states of the same base are aligned and states of different bases are misaligned, in a general case with any other set of misalignments between states, the protocol is no longer secure since the attacker can arbitrarily modify

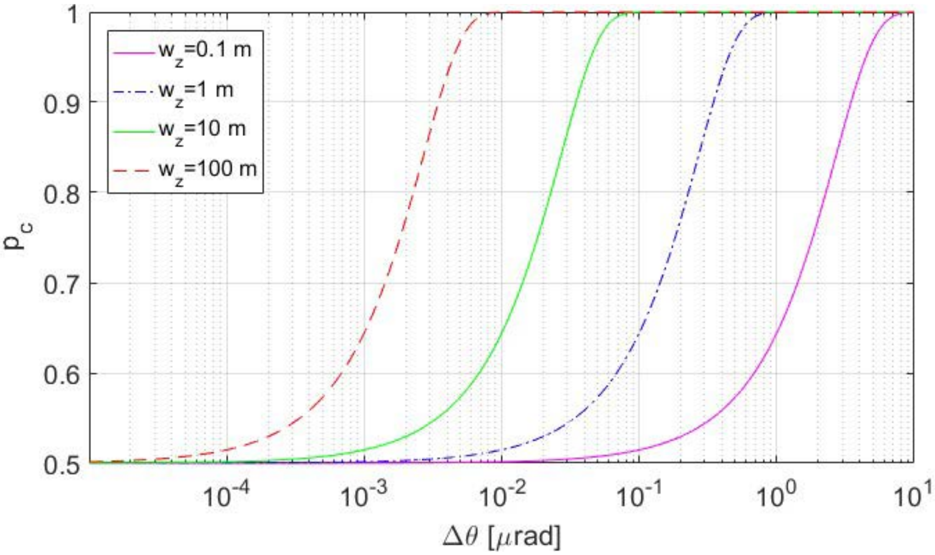

FIGURE 4. **Probability of using the correct measurement base, p_c, as a function of the angular divergence between sources, $\Delta\theta$, for the different cases of beam radius, w_z, indicated in the legend.**

the beam size and obtain the required angular resolution. However, in practice, this attack might not be so easy to implement since Eve must propagate the beams over long distances. For example, for the case of a collimated beam with a $10cm$ beam radius at the output of the transmitter, a transmission of more than $360km$ is needed to obtain $w_z \geq 1\ m$. In addition, Eve must use an optical system to detect the sates with a large aperture, even tens or hundreds of meters in the case of $\Delta\theta \leq 1\ nrad$ (see figure 4). On the other hand, for $\Delta\theta \geq 0.1\mu rad$, the beam size magnitudes obtained in figure 4 are close to the sizes of telescope apertures that we can obtain in practice, less than $10m$. For these cases, the attack could be carried out successfully with current technology. To get an idea of the possible impact of the attack with current technology, let us consider a QKD transmission between a LEO satellite and a ground station, specifically we assume the Micius satellite [42]. The wavelength of the quantum signal is $850nm$. Assuming the near-diffraction-limited far-field divergence of the beam obtained in [42], $10\mu rad$, we estimate that the collimated beam radius at the aperture of the QKD transmitter should be around $0.025\ m$, which we will use in our calculations. We consider that the attacker is capable of deviating the beam from the QKD transmitter to its own receiver, which could be another space telescope in a nearby orbit, measure, generate and resend the states to the original receiver. The attack could be performed before the propagation of the beam trough the lower layers of the atmosphere, thus requiring a smaller telescope aperture. We can consider a representative case for the attacker's space telescope aperture diameter using that of the Hubble telescope ($2.4\ m$). Considering a distance between the QKD transmitter and the space telescope of the attacker of $55km$, the beam radius of the quantum signal at the aperture of the attacker's telescope is slightly smaller than $0.6\ m$. With the considered aperture of $2.4\ m$, the truncation effects on the beam are low. We do not know the angular misalignment $\Delta\theta$ between the different beams in the case of the Micius satellite, therefore, we will assume different values of $\Delta\theta$. Assuming an angular misalignment $\Delta\theta = 1\mu rad$, we obtain a probability of correct measurement $p_c = 0.986$ and therefore a probability of generating an error in Bob $p_{err} = 0.007$. With an analogous procedure, in the case of an angular misalignment $\Delta\theta = 0.5\mu rad$, the attacker obtains a probability of correct measurement $p_c = 0.865$ and causes an error with probability $p_{err} = 0.068$. In the case of $\Delta\theta = 0.25\mu rad, p_c = 0.709$ and $p_{err} = 0.146$. Finally, considering an angular misalignment $\Delta\theta = 0.1\mu rad$, the attacker obtains a probability of correct measurement $p_c = 0.587$ and causes an error with probability $p_{err} = 0.206$. Thus, with current space telescope sizes, QKD systems using different sources with an angular misalignment greater than $1\mu rad$ are totally insecure, and lower values of angular misalignment they are still highly insecure but obtaining less information leakage. We can consider conditional security if we assume a limited technological power on the part of the attackers, and design safe systems under these assumptions. For example, we can consider a limit in the size of the optical telescopes of the attackers, which is reasonable due to the difficulty of constructing large mirrors. With this limit in the size of the apertures of the optical systems, we can align the beams with sufficient precision to obtain secure QKD. In the case of atmospheric transmission, our model assumes attackers with ideally perfect adaptive optics and beam stabilization systems. For attackers with limited technological power our result serves as an upper bound of the information that could be gained by an eavesdropper, but a more complex model to calculate more accurate probabilities of correct and incorrect measurements is required. However, QKD was potentially able to guarantee the security versus

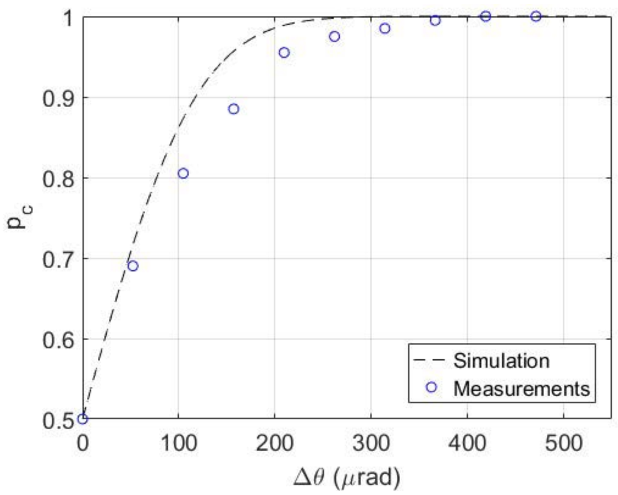

FIGURE 5. Probability, p_c, that the attacker measures with the correct base as a function of the angular divergence, $\Delta\theta$. The dashed line represents the simulated results, while the dots represent the measurements.

an attacker with infinite capabilities. If we do not want to assume technological limitations, increasing the precision of the alignment between the laser sources is not enough. Thus, we must design countermeasures that prevent the attack altogether.

A. COUNTERMEASURES

We propose two countermeasures to make the states spatially indistinguishable. The first one involves coupling all states into the same single-mode optical fiber before launching them into the free-space channel. In this way, by confining them in the fiber, a single spatial mode of propagation is achieved that makes them almost spatially indistinguishable [35], obtaining low information leakage. Considering the corresponding information leakage, the privacy amplification could be sufficient to protect the key exchange. This is the case of QKD systems such as those of [40] and [41], whereby the states are propagated through the same fiber before transmitting them into free-space. However, this design is still susceptible to attacks that take advantage of spectral and temporal distinguishability. It is possible though, to protect the system against those side-channel attacks as they do in [43], by making the temporal and spectral profiles of the different sources as indistinguishable as possible, but it may limit the secure key rate. Alternatively, other systems such as those of [55] and [56] generate the optical pulses using a single laser and select the polarization state with a polarization modulator, thus eliminating the possibility of side-channel attacks that take advantage of the spectral, spatial, and temporal distinguishability of the different sources. Since only one source is used, there are no differences in the wavelength spectrum, spatial mode, propagation direction, or temporal profile among the different states. Even in the case of a polarization modulator that introduces information leakage, it could be quantified and considered in the privacy amplification step. The disadvantage of a single laser design is that the polarization modulator can expose the system to Trojan horse attacks [10]. However, this can be solved by the techniques proposed in [10], [19] and [20], which guarantee the security of the key exchange against these attacks.

The existing QKD systems mentioned in the introduction [39], [42]–[46] are no longer unconditionally secure against the described attack. Thus, the proposed countermeasures could be applied to these QKD systems in order to secure them without drastically changing the QKD protocol.

IV. CONCLUSION

We have proposed and modelled a side-channel attack that takes advantage of the spatial distinguishability of the generated states by a QKD system with different misaligned laser sources, obtaining an analytical expression for the probability of discriminating the quantum states. We have also carried out an experimental proof-of-principle of the attack, validating the analytical result of the probability that the attacker discriminates the states. The described attack makes theoretically unsafe any system with any angular misalignment between the sources considering an attacker with infinite technological power. There are two possible ways of facing this problem: assuming limitations in the technological power of the attacker guaranteeing ‘‘only’’ conditional security for the QKD systems, or changing the design of the QKD systems to completely secure them against the described attack. For the second option we have proposed two countermeasures that protect QKD systems against the described attack: the use of single mode optical fiber to generate a unique spatial mode of propagation, and the use of a single laser to avoid spatial distinguishability of the different sources. Finally, we have pointed out the security considerations that must be taken into account with each countermeasure to guarantee secure key exchange. As possible future work, the attack could be performed in a more realistic situation. In addition, following the path of conditional security with limited technological power on the part of attackers, developing new beam alignment techniques and more realistic models with atmospheric transmission channels would be useful. On the other hand, following the path of unconditional security, work on the development and practical implementation of countermeasures that completely eliminate spatial distinguishability should be addressed.

REFERENCES

[1] D. Mayers, ‘‘Unconditional security in quantum cryptography,’’ *J. ACM*, vol. 48, no. 3, pp. 351–406, May 2001, doi: 10.1145/382780.382781.

[2] H.-K. Lo and H. F. Chau, ‘‘Unconditional security of quantum key distribution over arbitrarily long distances,’’ *Science*, vol. 283, no. 5410, pp. 2050–2056, Mar. 1999, doi: 10.1126/science.283.5410.2050.

[3] M. A. Nielsen and I. L. Chuang, ‘‘Quantum computation and quantum information,’’ *Amer. J. Phys.*, vol. 70, no. 5, pp. 558–559, 2002, doi: 10.1119/1.1463744.

[4] I. B. Djordjevic, *Physical-Layer Security and Quantum Key Distribution*. Cham, Switzerland: Springer, 2019.

[5] B. Huttner, A. Müller, J. D. Gautier, H. Zbinden, and N. Gisin, ‘‘Unambiguous quantum measurement of nonorthogonal states,’’ *Phys. Rev. A, Gen. Phys.*, vol. 54, no. 5, pp. 3783–3789, Nov. 1996, doi: 10.1103/PhysRevA.54.3783.

[6] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, "The security of practical quantum key distribution," *Rev. Mod. Phys.*, vol. 81, no. 3, pp. 1301–1350, Sep. 2009, doi: 10.1103/RevModPhys.81.1301.

[7] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, "Secure quantum key distribution with realistic devices," *Rev. Modern Phys.*, vol. 92, no. 2, May 2020, Art. no. 025002, doi: 10.1103/RevModPhys.92.025002.

[8] V. Scarani and C. Kurtsiefer, "The black paper of quantum cryptography: Real implementation problems," *Theor. Comput. Sci.*, vol. 560, pp. 27–32, Dec. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304397514006938

[9] N. Jain, B. Stiller, I. Khan, D. Elser, C. Marquardt, and G. Leuchs, "Attacks on practical quantum key distribution systems (and how to prevent them)," *Contemp. Phys.*, vol. 57, no. 3, pp. 366–387, Jul. 2016, doi: 10.1080/00107514.2016.1148333.

[10] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, and G. Ribordy, "Trojan-horse attacks on quantum-key-distribution systems," *Phys. Rev. A, Gen. Phys.*, vol. 73, no. 2, Feb. 2006, Art. no. 022320, doi: 10.1103/PhysRevA.73.022320.

[11] A. N. Bugge, S. Sauge, A. M. M. Ghazali, J. Skaar, L. Lydersen, and V. Makarov, "Laser damage helps the eavesdropper in quantum cryptography," *Phys. Rev. Lett.*, vol. 112, no. 7, Feb. 2014, Art. no. 070503, doi: 10.1103/PhysRevLett.112.070503.

[12] V. Makarov and D. R. Hjelme, "Faked states attack on quantum cryptosystems," *J. Modern Opt.*, vol. 52, no. 5, pp. 691–705, Mar. 2005, doi: 10.1080/09500340410001730986.

[13] M. Bozzio, A. Cavaillès, E. Diamanti, A. Kent, and D. Pitalúa-García, "Multiphoton and side-channel attacks in mistrustful quantum cryptography," *PRX Quantum*, vol. 2, no. 3, Sep. 2021, Art. no. 030338, doi: 10.1103/PRXQuantum.2.030338.

[14] D. Park, G. Kim, D. Heo, S. Kim, H. Kim, and S. Hong, "Single trace side-channel attack on key reconciliation in quantum key distribution system and its efficient countermeasures," *ICT Exp.*, vol. 7, no. 1, pp. 36–40, Mar. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405959521000138

[15] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, "Device-independent security of quantum cryptography against collective attacks," *Phys. Rev. Lett.*, vol. 98, no. 23, Jun. 2007, Art. no. 230501, doi: 10.1103/PhysRevLett.98.230501.

[16] T. F. da Silva, G. B. Xavier, G. P. T. ao, and J. P. von der Weid, "Real-time monitoring of single-photon detectors against eavesdropping in quantum key distribution systems," *Opt. Exp.*, vol. 20, no. 17, pp. 18911–18924, Aug. 2012. [Online]. Available: http://opg.optica.org/oe/abstract.cfm?URI=oe-20-17-18911

[17] T. F. da Silva, G. C. do Amaral, G. B. Xavier, G. P. Temporão, and J. P. von der Weid, "Safeguarding quantum key distribution through detection randomization," *IEEE J. Sel. Topics Quantum Electron.*, vol. 21, no. 3, pp. 159–167, Oct. 2015.

[18] P. González, L. Rebón, T. F. da Silva, M. Figueroa, C. Saavedra, M. Curty, G. Lima, G. B. Xavier, and W. A. T. Nogueira, "Quantum key distribution with untrusted detectors," *Phys. Rev. A, Gen. Phys.*, vol. 92, no. 2, Aug. 2015, Art. no. 022337, doi: 10.1103/PhysRevA.92.022337.

[19] N. Jain, B. Stiller, I. Khan, V. Makarov, C. Marquardt, and G. Leuchs, "Risk analysis of trojan-horse attacks on practical quantum key distribution systems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 21, no. 3, pp. 168–177, May 2015.

[20] A. V. Borisova, B. D. Garmaev, I. B. Bobrov, S. S. Negodyaev, and I. V. Sinil'shchikov, "Risk analysis of countermeasures against the trojan-horse attacks on quantum key distribution systems in 1260–1650 nm spectral range," *Opt. Spectrosc.*, vol. 128, no. 11, pp. 1892–1900, Nov. 2020.

[21] H.-K. Lo, M. Curty, and B. Qi, "Measurement-device-independent quantum key distribution," *Phys. Rev. Lett.*, vol. 108, p. 130503, Mar. 2012, doi: 10.1103/PhysRevLett.108.130503.

[22] S. L. Braunstein and S. Pirandola, "Side-channel-free quantum key distribution," *Phys. Rev. Lett.*, vol. 108, no. 13, Mar. 2012, Art. no. 130502, doi: 10.1103/PhysRevLett.108.130502.

[23] M. Lucamarini, Z. L. Yuan, J. F. Dynes, and A. J. Shields, "Overcoming the rate–distance limit of quantum key distribution without quantum repeaters," *Nature*, vol. 557, pp. 400–403, May 2018, doi: 10.1038/s41586-018-0066-6.

[24] X.-B. Wang, Z.-W. Yu, and X.-L. Hu, "Twin-field quantum key distribution with large misalignment error," *Phys. Rev. A, Gen. Phys.*, vol. 98, no. 6, Dec. 2018, Art. no. 062323, doi: 10.1103/PhysRevA.98.062323.

[25] Y. Liu, T.-Y. Chen, L.-J. Wang, H. Liang, G.-L. Shentu, J. Wang, K. Cui, H.-L. Yin, N.-L. Liu, L. Li, X. Ma, J. S. Pelc, M. M. Fejer, C.-Z. Peng, Q. Zhang, and J.-W. Pan, "Experimental measurement-device-independent quantum key distribution," *Phys. Rev. Lett.*, vol. 111, Sep. 2013, Art. no. 130502, doi: 10.1103/PhysRevLett.111.130502.

[26] Z. Tang, K. Wei, O. Bedroya, L. Qian, and H.-K. Lo, "Experimental measurement-device-independent quantum key distribution with imperfect sources," *Phys. Rev. A, Gen. Phys.*, vol. 93, no. 4, Apr. 2016, Art. no. 042308, doi: 10.1103/PhysRevA.93.042308.

[27] X. Zhong, J. Hu, M. Curty, L. Qian, and H.-K. Lo, "Proof-of-Principle experimental demonstration of twin-field type quantum key distribution," *Phys. Rev. Lett.*, vol. 123, no. 10, Sep. 2019, Art. no. 100506, doi: 10.1103/PhysRevLett.123.100506.

[28] J.-P. Chen, C. Zhang, Y. Liu, C. Jiang, W.-J. Zhang, Z.-Y. Han, S.-Z. Ma, X.-L. Hu, Y.-H. Li, H. Liu, F. Zhou, H.-F. Jiang, T.-Y. Chen, H. Li, L.-X. You, Z. Wang, X.-B. Wang, Q. Zhang, and J.-W. Pan, "Twin-field quantum key distribution over a 511 km optical fibre linking two distant metropolitan areas," *Nature Photon.*, vol. 15, no. 8, pp. 570–575, Aug. 2021, doi: 10.1038/s41566-021-00828-5.

[29] H. Liu, C. Jiang, H. T. Zhu, M. Zou, Z. W. Yu, X. L. Hu, H. Xu, S. Ma, Z. Han, J. P. Chen, and Y. Dai, "Field test of twin-field quantum key distribution through sending-or-not-sending over 428 km," *Phys. Rev. Lett.*, vol. 126, no. 25, Jun. 2021, Art. no. 250502, doi: 10.1103/PhysRevLett.126.250502.

[30] Y. Liu, Z.-W. Yu, W. Zhang, J.-Y. Guan, J.-P. Chen, C. Zhang, X.-L. Hu, H. Li, C. Jiang, J. Lin, T.-Y. Chen, L. You, Z. Wang, X.-B. Wang, Q. Zhang, and J.-W. Pan, "Experimental twin-field quantum key distribution through sending or not sending," *Phys. Rev. Lett.*, vol. 123, no. 10, Sep. 2019, Art. no. 100505, doi: 10.1103/PhysRevLett.123.100505.

[31] V. Zapatero and M. Curty, "Long-distance device-independent quantum key distribution," *Sci. Rep.*, vol. 9, no. 1, Dec. 2019, Art. no. 260503, doi: 10.1103/PhysRevLett.125.260503.

[32] C. Zhang, X.-L. Hu, C. Jiang, J.-P. Chen, Y. Liu, W. Zhang, Z.-W. Yu, H. Li, L. You, Z. Wang, X.-B. Wang, Q. Zhang, and J.-W. Pan, "Experimental side-channel-secure quantum key distribution," *Phys. Rev. Lett.*, vol. 128, no. 19, May 2022, Art. no. 190503, doi: 10.1103/PhysRevLett.128.190503.

[33] W.-S. Huang, W. Zhang, and Y.-D. Huang, "Elimination of spatial side-channel information for compact quantum key distribution senders," *J. Electron. Sci. Technol.*, vol. 17, no. 3, pp. 195–203, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1674862X19300321

[34] N. Jain, C. Wittmann, L. Lydersen, C. Wiechers, D. Elser, C. Marquardt, V. Makarov, and G. Leuchs, "Device calibration impacts security of quantum key distribution," *Phys. Rev. Lett.*, vol. 107, no. 11, Sep. 2011, Art. no. 110501, doi: 10.1103/PhysRevLett.107.110501.

[35] S. Nauerth, M. Fürst, T. Schmitt-Manderbach, H. Weier, and H. Weinfurter, "Information leakage via side channels in freespace BB84 quantum cryptography," *New J. Phys.*, vol. 11, no. 6, Jun. 2009, Art. no. 065001, doi: 10.1088/1367-2630/11/6/065001.

[36] A. Biswas, A. Banerji, P. Chandravanshi, R. Kumar, and R. P. Singh, "Experimental side channel analysis of BB84 QKD source," *IEEE J. Quantum Electron.*, vol. 57, no. 6, pp. 1–7, Dec. 2021.

[37] C. H. Bennett, "Quantum cryptography using any two nonorthogonal states," *Phys. Rev. Lett.*, vol. 68, no. 21, pp. 3121–3124, May 1992, doi: 10.1103/PhysRevLett.68.3121.

[38] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," *Theor. Comput. Sci.*, vol. 560, pp. 7–11, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304397514004241

[39] R. J. Hughes, J. E. Nordholt, D. Derkacs, and C. G. Peterson, "Practical free-space quantum key distribution over 10 km in daylight and at night," *New J. Phys.*, vol. 4, p. 43, Jul. 2002, doi: 10.1088/1367-2630/4/1/343.

[40] T. Schmitt-Manderbach, H. Weier, M. Fürst, R. Ursin, F. Tiefenbacher, T. Scheidl, J. Perdigues, Z. Sodnik, C. Kurtsiefer, J. G. Rarity, A. Zeilinger, and H. Weinfurter, "Experimental demonstration of free-space decoy-state quantum key distribution over 144 km," *Phys. Rev. Lett.*, vol. 98, no. 1, Jan. 2007, Art. no. 010504.

[41] S.-K. Liao, H.-L. Yong, C. Liu, G.-L. Shentu, D.-D. Li, J. Lin, H. Dai, S.-Q. Zhao, B. Li, and J.-Y. E. A. Guan, "Long-distance free-space quantum key distribution in daylight towards inter-satellite communication," *Nature Photon.*, vol. 11, no. 8, pp. 509–513, 2017.

[42] S.-K. Liao, W.-Q. Cai, W.-Y. Liu, L. Zhang, Y. Li, J.-G. Ren, J. Yin, Q. Shen, Y. Cao, and Z.-P. E. A. Li, ''Satellite-to-ground quantum key distribution,'' *Nature*, vol. 549, no. 7670, pp. 43–47, 2017.

[43] H. Ko, B.-S. Choi, J.-S. Choe, K.-J. Kim, J.-H. Kim, and C. J. Youn, ''High-speed and high-performance polarization-based quantum key distribution system without side channel effects caused by multiple lasers,'' *Photon. Res.*, vol. 6, no. 3, pp. 214–219, Mar. 2018. [Online]. Available: http://opg.optica.org/prj/abstract.cfm?URI=prj-6-3-214

[44] J.-Y. Wang, B. Yang, S.-K. Liao, L. Zhang, Q. Shen, X.-F. Hu, J.-C. Wu, S.-J. Yang, H. Jiang, and Y.-L. E. A. Tang, ''Direct and full-scale experimental verifications towards ground-satellite quantum key distribution,'' *Nature Photon.*, vol. 7, no. 5, pp. 387–393, 2013.

[45] W.-Y. Liu, X.-F. Zhong, T. Wu, F.-Z. Li, B. Jin, Y. Tang, H.-M. Hu, Z.-P. Li, L. Zhang, W.-Q. Cai, S.-K. Liao, Y. Cao, and C.-Z. Peng, ''Experimental free-space quantum key distribution with efficient error correction,'' *Opt. Exp.*, vol. 25, no. 10, pp. 10716–10723, May 2017. [Online]. Available: http://opg.optica.org/oe/abstract.cfm?URI=oe-25-10-10716

[46] X. Tang, L. Ma, A. Mink, A. Nakassis, H. Xu, B. Hershman, J. Bienfang, D. Su, R. F. Boisvert, C. Clark, and C. Williams, ''Quantum key distribution system operating at sifted-key rate over 4 Mbit/s,'' *Proc. SPIE*, vol. 6244, pp. 182–189, May 2006, doi: 10.1117/12.664455.

[47] H. Ko, B.-S. Choi, J.-S. Choe, K.-J. Kim, J.-H. Kim, and C. J. Youn, ''Critical side channel effects in random bit generation with multiple semiconductor lasers in a polarization-based quantum key distribution system,'' *Opt. Exp.*, vol. 25, no. 17, pp. 20045–20055, 2017. [Online]. Available: http://opg.optica.org/oe/abstract.cfm?URI=oe-25-17-20045

[48] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[49] S. A. Self, ''Focusing of spherical Gaussian beams,'' *Appl. Opt.*, vol. 22, no. 5, pp. 658–661, Mar. 1983. [Online]. Available: http://opg.optica.org/ao/abstract.cfm?URI=ao-22-5-658

[50] Z. L. Horváth and Z. Bor, ''Focusing of truncated Gaussian beams,'' *Opt. Commun.*, vol. 222, nos. 1–6, pp. 51–68, Jul. 2003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0030401803015621

[51] E. Hecht, *Optics*. Reading, MA, USA: Addison-Wesley, 2017.

[52] D. Jiang, Y. Yuan, Y. Lin, T. Lyu, B. Zhu, and Y. Yao, ''Mean irradiance profile of a Gaussian beam under random jitter,'' *Opt. Exp.*, vol. 26, no. 21, pp. 27472–27481, Oct. 2018. [Online]. Available: http://opg.optica.org/oe/abstract.cfm?URI=oe-26-21-27472

[53] M. Toyoshima, S. Yamakawa, T. Yamawaki, K. Arai, M. R. Garcia-Talavera, A. Alonso, Z. Sodnik, and B. Demelenne, ''Long-term statistics of laser beam propagation in an optical ground-to-geostationary satellite communications link,'' *IEEE Trans. Antennas Propag.*, vol. 53, no. 2, pp. 842–850, Feb. 2005.

[54] L. C. Andrews and R. L. Phillips, *Laser Beam Propagation Through Random Media*. Bellingham, WA, USA: SPIE, 2005.

[55] M. Avesani, L. Calderaro, M. Schiavon, A. Stanco, C. Agnesi, A. Santamato, M. Zahidy, A. Scriminich, G. Foletto, G. Contestabile, M. Chiesa, D. Rotta, M. Artiglia, A. Montanaro, M. Romagnoli, V. Sorianello, F. Vedovato, G. Vallone, and P. Villoresi, ''Full daylight quantum-key-distribution at 1550 nm enabled by integrated silicon photonics,'' *NPJ Quantum Inf.*, vol. 7, no. 1, pp. 1–8, Dec. 2021.

[56] C. Ma, W. D. Sacher, Z. Tang, J. C. Mikkelsen, Y. Yang, F. Xu, T. Thiessen, H.-K. Lo, and J. K. S. Poon, ''Silicon photonic transmitter for polarization-encoded quantum key distribution,'' *Optica*, vol. 3, no. 11, pp. 1274–1278, Nov. 2016. [Online]. Available: http://opg.optica.org/optica/abstract.cfm?URI=optica-3-11-1274

PABLO ARTEAGA-DÍAZ received the B.Sc. degree in physics from the University of La Laguna, in 2018, and the M.S. degree in new technologies for electronics and photonics from the Complutense University of Madrid, in 2019. He is currently pursuing the Ph.D. degree with the Department of Information and Communication Technologies, Institute of Physical and Information Technologies, Spanish National Research Council, Madrid. His research interests include free-space optics and quantum key distribution.

DANIEL CANO received the B.Sc. degree in physics from Universidad Complutense, Madrid, Spain, in 2002, and the Ph.D. degree in physics from the University of Tübingen, Germany, in 2009. From 2010 to 2013, he was a Postdoctoral Researcher at the Center for Collective Quantum Phenomena and Their Applications, Tübingen. From 2014 to 2016, he was a Postdoctoral Researcher at IMDEA Nanoscience, Madrid. From 2017 to 2019, he was a Postdoctoral Researcher at the Institute of Photonic Sciences in Barcelona. Since 2021, he has been with the Institute of Physical and Information Technologies, Spanish National Research Council, Madrid. He has authored 19 peer-reviewed articles and one patent. His research interests include the development of novel quantum instruments with applications in Nanotechnology, photonics, quantum information, and secure optical communications. He was awarded the Marie Sklodowska-Curie Fellowship for a project on fast Quantum Optics in graphene.

VERONICA FERNANDEZ received the B.Sc. degree (Hons.) in physics with electronics from the University of Seville, in 2002, and the Ph.D. degree in physics from Heriot-Watt University, Edinburgh, U.K., in 2006. She joined the Institute for Physical and Information Technologies of the Spanish National Research Council (CSIC), Madrid, in November 2007 with the Postdoctoral ''Juan de la Cierva'' Contract. In 2009, she received a permanent position at the Group of Cryptography and Information Security (GiCSI), CSIC. She leads the Quantum Communication research line, focused in experimental quantum key distribution systems in free-space transmission channels for high-speed metropolitan applications, and related technologies, such as beam and polarization tracking systems for mobile aerial QKD. She has published more than 50 scientific contributions in the field, is the co-coordinator of the Quantum Technologies Platform from CSIC, and advisory committee of optical satellite communication programs for the European Space Agency (ESA). Her research interests include QKD for terrestrial and satellite communications, for both discrete and continuous-variables protocols.