Bounding passive light-source side channels in quantum key distribution via Hong-Ou-Mandel interference

Alexander Duplinskiy[1,2,3,*] and Denis Sych[3,4,5]
[1]*QRate, Novaya Avenue 100, Moscow, Russia*
[2]*Russian Quantum Center, Bolshoy Boulevard 30, Building 1, Skolkovo, Moscow 121205, Russia*
[3]*NTI Center for Quantum Communications, National University of Science and Technology MISiS, Leninsky Prospekt 4, Moscow 119049, Russia*
[4]*P. N. Lebedev Physical Institute, Russian Academy of Sciences, Leninskiy Prospekt 53, 119991 Moscow, Russia*
[5]*Sirius University of Science and Technology, 1 Olympic Avenue, Sochi 354340, Russia*

(Received 21 April 2021; accepted 5 May 2021; published 6 July 2021; corrected 29 November 2021)

Passive side-channel attacks in quantum key distribution (QKD) aim at obtaining information about the quantum signals without disturbing them and hence compromise real-world QKD security. Currently, there are no reliable tools for the assessment of QKD signal generation imperfections. In this work we propose a generic experimental method which allows to upper-bound QKD light-source imperfections and directly integrate them into the modern security proofs. The method relies on Hong-Ou-Mandel interference between different emitted signals: the maximum interference visibility reveals overall signal distinguishability that could lead to passive side-channel information leakage. We apply it for the standard decoy-state BB84 protocol and calculate a lower bound on the secret key rate for realistic values of interference visibility. The method can be readily implemented in practical QKD setups and is especially relevant for multiple-laser QKD systems such as the one installed on the Micius satellite.

DOI: 10.1103/PhysRevA.104.012601

I. INTRODUCTION

Quantum key distribution (QKD) is a secure method for distributing keys between distant users—the transmitter (Alice) and the receiver (Bob). Although the security of QKD is based on fundamental laws of quantum physics, the real-world QKD implementations can significantly deviate from their theoretical models and compromise the security statement. The real devices which are used in QKD allow for side-channel information leakage, which can be exploited by an eavesdropper (Eve) to hack a QKD system [1,2].

Development of measurement-device-independent (MDI)-QKD allowed to secure the receiving part; however, the sender of the quantum states is still at risk [3]. The side-channel attacks on Alice's systems can be divided into two major types: active and passive ones. In an active attack, or Trojan horse attack (THA), Eve sends optical signals to Alice, and tries to obtain information about the settings from the reflected light. The natural technical countermeasures against THA include optical isolation from the quantum channel, thus the influence of the THA on the QKD security can be upper bounded according to the level of isolation and maximum optical power which can be injected into the optical fiber (for recent comprehensive analysis of the THA see Refs. [4,5]).

The basic idea of passive side-channel attacks is to exploit imperfections within the signal generation stage and measure auxiliary degrees of freedom, such as signal timing or spectral differences, which may leak information about the secret key. A general model of possible passive leaks and flaws of Alice's setup includes modulation deviation and nonoperational degrees of freedom (the quantum state outside of the single mode space) [6]. Modulation deviation, i.e., deviations of the actual quantum signal from the desired one, can be directly measured in the experiment, as it deals only with the known (operational) degree of freedom of the signals.

Even when the models include device imperfections, the theoretical figures of merit are related to the fidelity of different density matrices, which has no direct operational meaning [6]. As a result, modern QKD implementations do not include any part that allows to estimate such figures of merit; hence theoretical security proofs are not fully taken up in practice. As a simple example, the first-ever QKD setup used high-voltage Pockels cells, which reveal their settings by acoustical noise [7]. Talking about modern QKD, a common approach in high-speed free-space systems is to use different laser sources to encode different quantum states, without applying any external modulation [8–11]. Also, there are many fiber-optic QKD setups that are based on multiple lasers; some of them implement quantum cryptography beyond QKD [12,13]. Such setups are immune to conventional THAs as they do not use any modulators, but can be insecure because of information leakage beyond the single-mode space. Even if two lasers seem identical at first sight, the actual parameters of the emitted light, such as spectrum, temporal, and spatial shape, always vary [14].

The problem of potential distinguishability estimation of quantum signals due to the passive side channels remains open. Previous work proposed to perform individual measurements of different degrees of freedom [2,14]. This approach,

*duplinskii@phystech.edu

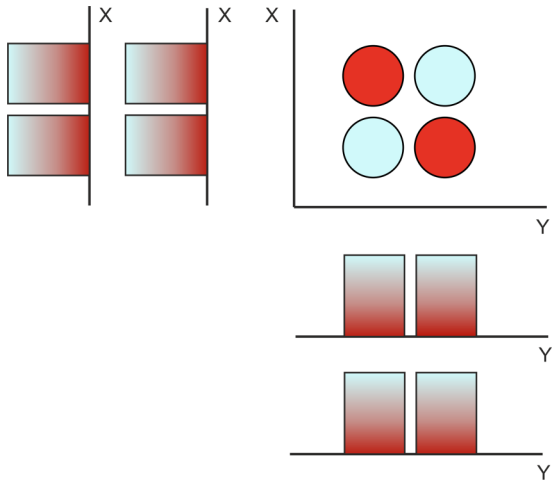

FIG. 1. An example of two nonoverlapping joint probability distributions, colored red and blue, corresponding to a pair of completely distinguishable quantum states. The partial distributions (shown by red-blue gradient) perfectly coincide; hence any individual measurement fails to distinguish the quantum states.

however, fails to reveal the actual distinguishability of the signals. For example, the same temporal and spatial distributions do not guarantee that the joint distributions of intensity in time and space for two states also match. In addition, spectral distribution may depend on time due to the frequency chirp of a laser source [15]. A simple illustration of two completely distinguishable quantum states that have identical partial distributions (and hence are indistinguishable within the existing approaches) is shown in Fig. 1.

In contrast to individual measurements of separate degrees of freedom, the effect of interference allows to study the overall distinguishability of the pulses, since distinguishable states do not interfere [16]. For example, fourth-order, or Hong-Ou-Mandel (HOM), interference is a standard method for characterization of single-photon sources [17,18]. The possibility of utilizing HOM interference to characterize unmeasured degrees of freedom for the QKD sources with multiple lasers has been mentioned in Ref. [19].

In this work, we show how to use the HOM interference of optical pulses for estimation of their distinguishability in the context of QKD, and upper bound the influence of passive side-channel effects caused by the optical mode mismatch between different pulses. We calculate the secret key generation rate of the BB84 protocol with phase-randomized weak coherent pulses (PRWCPs), taking into account the realistic values of visibility between different quantum states, and show the practical applicability of the proposed method. The proposed method allows to close the gap between security proofs and the real-world QKD.

The paper is organized as follows. In Sec. II we describe Hong-Ou-Mandel effect for PRWCPs from the quantum point of view, and link visibility of HOM interference with distinguishability of interfering states. Next, in Sec. III the results of the HOM interference are used to calculate the value of bases imbalance. Then, in Sec. IV the value of bases imbalance are used to correct estimation of the single-photon error rate, and, finally, obtain the key generation rate for different values of HOM visibility. Based on currently known experimental results, the applicability of the proposed method is discussed in Sec. V.

II. HONG-OU-MANDEL INTERFERENCE AND NONORTHOGONALITY OF QUANTUM STATES

The fourth-order interference of single photons was introduced by Hong, Ou, and Mandel in 1987 in order to measure the time delay between single photons precisely [17]. The key idea is that two indistinguishable single photons matched on a 50:50 beam splitter always exit it pairwise. This is a purely quantum effect, which is a result of two-photon wave function interference [20]. Later experiments allowed to demonstrate Hong-Ou-Mandel interference even for the photons emitted by different atoms [21]. Schematics of the optical setup to observe the effect is illustrated in Fig. 2.

Single photon detectors in both arms allow to count coincidence clicks, when two photons leave the beam splitter in different arms. While the time difference Δt between single-photon wave packets is large enough so that the photons do not overlap, their behavior is independent from each other. As a result, coincidence clicks are observed in half of the detection events.

As soon as overlap becomes nonzero, the number of coincidence clicks decreases. In the limit of perfectly mode-matched pure single-photon states, coincidence clicks completely vanish. In order to characterize the indistinguishability of photons, the visibility of HOM interference is introduced as the difference between the maximum and minimum numbers of coincidence clicks, divided by the maximum number of coincidence clicks: $V = (N_{\max} - N_{\min})/N_{\max}$ (Fig. 2). It is easy to see that visibility equals zero for completely orthogonal (distinguishable) states and equals 1 in the case of perfect indistinguishability. In the case one knows the density matrices of the single-photon states ($\hat{\rho}_1^{\mathrm{sp}}$ and $\hat{\rho}_2^{\mathrm{sp}}$), the visibility can be calculated as

$$V_{\hat{\rho}_1^{\mathrm{sp}}\hat{\rho}_2^{\mathrm{sp}}} = \mathrm{Tr}\left(\hat{\rho}_1^{\mathrm{sp}}\hat{\rho}_2^{\mathrm{sp}}\right). \tag{1}$$

Most of the practical QKD systems utilize PRWCPs instead of the single-photon states. Nevertheless, a HOM-like effect also takes place for two PRWCPs as well, though the coincidence clicks never totally disappear. The maximum possible visibility obtained in this type of experiment is 0.5. In contrast to the single-photon case, the HOM effect for PRWCPs can be explained both from classical and quantum points of view. It is quite natural as coherent states could be treated as the "most classical" quantum states. Anyway, in this paper we focus on the quantum approach and study how the visibility depends on the characteristics of interfering states.

The PRWCPs can be expressed as a Poissonian combination of independent Fock states [22]. Let us consider two such states $\hat{\rho}_1$ and $\hat{\rho}_2$ with equal intensity μ at a 50:50 beam splitter. Their density matrices can be written in Fock basis as follows:

$$\hat{\rho} = \sum_{n=0}^{\infty} \rho_{nn} |n\rangle\langle n| = \sum_{n=0}^{\infty} e^{-\mu} \frac{\mu^n}{n!} (\hat{a}^\dagger)^n |0\rangle\langle 0| (\hat{a})^n. \tag{2}$$

To simulate a real system one has to take into account detection efficiency, including both optical losses and single-

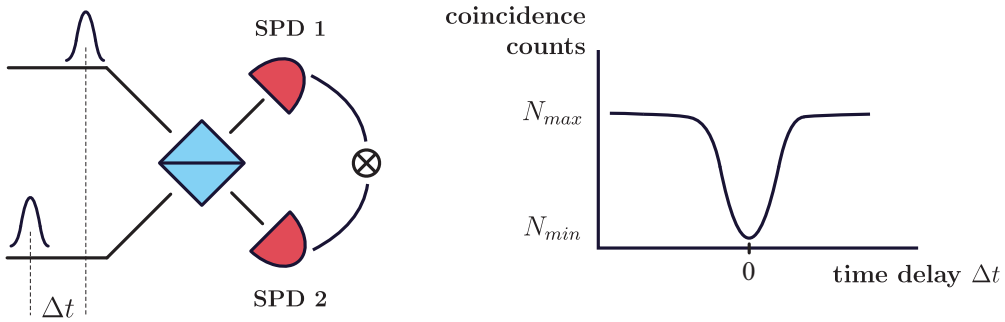

FIG. 2. Typical setup for the HOM experiment (left). Two single-photon pulses enter a beam splitter, while coincidence counts are measured with two single-photon detectors (SPDs). Dependence of the number of coincidence counts on the time delay Δt has a dip that corresponds to the maximum overlap of photons on the beam splitter (right).

photon detector efficiency. Here we consider μ as the effective intensity of states, corrected for detection imperfections. Thereby the input state on a beam splitter is

$$\begin{aligned}\hat{\rho}_{\text{in}} &= \hat{\rho}_1 \otimes \hat{\rho}_2 \\ &= e^{-\mu}\Big(|0_1 0_2\rangle\langle 0_1 0_2| + \mu(|0_1 1_2\rangle\langle 1_1 0_2| + |1_1 0_2\rangle\langle 0_1 1_2|) \\ &\quad + \frac{\mu^2}{2}(|0_1 2_2\rangle\langle 0_1 2_2| + 2|1_1 1_2\rangle\langle 1_1 1_2| + |2_1 0_2\rangle\langle 2_1 0_2|) \\ &\quad + O(\mu^3)\Big). \end{aligned} \tag{3}$$

We omit the terms higher than the second degree, as we consider μ to be small. As soon as the terms with the total photon number less than 2 do not contribute to coincidence clicks, the only states that we need to study are

$$\hat{\rho}'_{\text{in}} = \frac{e^{-\mu}\mu^2}{2}(|0_1 2_2\rangle\langle 0_1 2_2| + 2|1_1 1_2\rangle\langle 1_1 1_2| + |2_1 0_2\rangle\langle 2_1 0_2|). \tag{4}$$

In the case of orthogonal modes, all three terms $|0_1 2_2\rangle\langle 0_1 2_2|$, $|1_1 1_2\rangle\langle 1_1 1_2|$, and $|2_1 0_2\rangle\langle 2_1 0_2|$ contribute equally to coincidence clicks. The middle term $|1_1 1_2\rangle\langle 1_1 1_2|$ leads to the standard single-photon HOM effect; i.e., for perfect node matching, coincidences vanish. At the same time, coincidence clicks due to the other terms, $|0_1 2_2\rangle\langle 0_1 2_2|$ and $|2_1 0_2\rangle\langle 2_1 0_2|$, never change, since there is no interference with vacuum. Thus the maximum visibility value is 0.5. In the intermediate case of partial distinguishability, the number of coincidence clicks from the state $|1_1 1_2\rangle\langle 1_1 1_2|$ reduces by the factor of $(1 - V_{\hat{\rho}_1^{\text{sp}}\hat{\rho}_2^{\text{sp}}})$, according to (1), resulting in visibility for PRWCPs:

$$V_{\hat{\rho}_1\hat{\rho}_2} = \tfrac{1}{2}V_{\hat{\rho}_1^{\text{sp}}\hat{\rho}_2^{\text{sp}}} = \tfrac{1}{2}\text{Tr}\big(\hat{\rho}_1^{\text{sp}}\hat{\rho}_2^{\text{sp}}\big). \tag{5}$$

In these calculations we neglect high-order terms. To check the applicability limit of the obtained result, we carry out numerical simulations, including terms up to the 20th degree. Simulation results (Fig. 3) show that for μ below 0.025 photons per pulse Eq. (5) is very close to the actual behavior. Even for higher values of μ (e.g., 0.25) the dependence remains very close to linear; however, the visibility limit should be slightly corrected. For μ value more than 1 photon per pulse, Eq. (5) is no longer applicable. Note that the mean photon number used for the visibility measurement does not have to be the same as the one used for quantum communication signals in the QKD channel (see Fig. 6).

III. SIDE-CHANNEL INFORMATION AND HOM VISIBILITY

Side channel is a collective name for vulnerabilities that can cause information leakage to Eve, bypassing the communication protocol. Here we address only the passive type of side channels for Alice's device, meaning that information can be partially revealed via the distinguishability of pulses in nonoperational degrees of freedom. In other words, we consider mode mismatch between different optical signals, emitted by Alice.

We address conventional BB84 protocol with decoy states and study distinguishability between the two bases [4,23], caused by the nonoperational degrees of freedom. This approach is sufficient to quantify all BB84 vulnerabilities caused

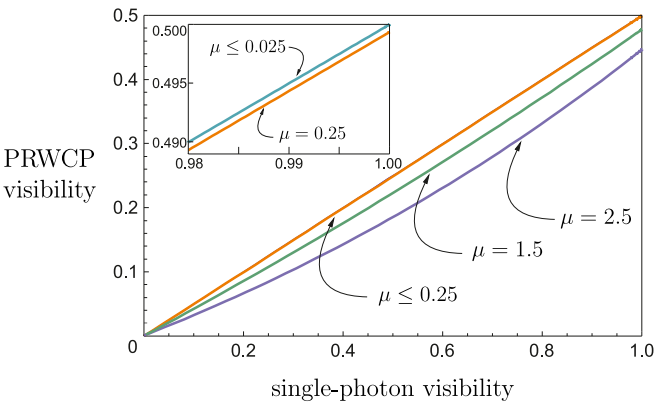

FIG. 3. HOM visibility of PRWCPs as a function of HOM visibility for single-photon states within the PRWCPs. PRWCP visibility can be measured directly in the experiment, while the single-photon visibility is used to integrate the measurement results into the security proof. The plot shows the dependence for different mean photon numbers per pulse μ. Calculations are performed including the terms (3) up to the 20th degree. Note that the use of large values of μ ($\mu > 1$) leads to the significant underestimation of the single-photon visibility, and hence to the overestimation of the potential side channels and to the reduction of the secret key rate. On the other hand, the choice of small values of μ leads to large statistical fluctuations and long data acquisition time.

by distinguishability, as it has been proven that for identical bases density matrices all bit flaws cause an increase of the respective error rate [24].

Ideally, the protocol implies that Eve is not able to discriminate one basis from the other, as their density matrices $\hat{\rho}_x = \frac{1}{2}(|H\rangle\langle+||V\rangle\langle|)$ and $\hat{\rho}_z = \frac{1}{2}(|D\rangle\langle+||A\rangle\langle|)$ are supposed to be the same. Nonperfect mode matching between the different bits causes the differences in the bases' density matrices, resulting in a vulnerability that could be exploited by Eve. The value that allows to quantify the impact of this effect is the so-called bases imbalance [4,23]:

$$\Delta = \frac{1 - \sqrt{F(\hat{\rho}_x, \hat{\rho}_z)}}{2}, \tag{6}$$

where $F(\hat{\rho}_x, \hat{\rho}_z)$ is the fidelity between the density matrices [25,26],

$$F(\hat{\rho}_x, \hat{\rho}_z) = (\mathrm{Tr}\sqrt{\sqrt{\hat{\rho}_x}\hat{\rho}_z\sqrt{\hat{\rho}_x}})^2 = (\mathrm{Tr}|\sqrt{\hat{\rho}_x}\sqrt{\hat{\rho}_z}|)^2. \tag{7}$$

It is easy to see that in the perfect case, when $\hat{\rho}_x$ and $\hat{\rho}_z$ are indistinguishable, Δ equals zero, whereas in the worst-case scenario, it reaches its maximum value of $1/2$.

We can express density matrices of X and Z bases as sums of bits, each of which is a tensor product of density matrices in operational and nonoperational degrees of freedom:

$$\hat{\rho}_x = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \hat{\rho}^{\lambda}_{x,0} + \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \hat{\rho}^{\lambda}_{x,1}, \tag{8}$$

$$\hat{\rho}_z = \frac{1}{2}\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \otimes \hat{\rho}^{\lambda}_{z,0} + \frac{1}{2}\begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix} \otimes \hat{\rho}^{\lambda}_{z,1}, \tag{9}$$

where $\hat{\rho}^{\lambda}_{b,i}$ accounts for nonoperational degrees of freedom for the bit i in the basis b.

In Appendix B we construct a lower bound on fidelity between X and Z bases depending on the fidelity between density matrices of nonoperational degrees of freedom. Then, using the triangle inequality for the Bures angle [27], we obtain the result

$$\arccos\sqrt{\mathrm{F}(\hat{\rho}_x, \hat{\rho}_z)} \leqslant \arccos \max_{i,j\in 0,1} \sqrt{F\left(\hat{\rho}^{\lambda}_{x,i}, \hat{\rho}^{\lambda}_{z,j}\right)} + \sum_{b\in\{x,z\}} \arccos \frac{1 + \sqrt{F\left(\hat{\rho}^{\lambda}_{b,0}, \hat{\rho}^{\lambda}_{b,1}\right)}}{2}. \tag{10}$$

Here the first term indicates the difference between the bases, whereas the second term represents differences between the bits within a basis. This result allows us to estimate bases imbalance with only pairwise interference experiments with separate pulses. To conduct an interference experiment on nonoperational degrees of freedom, we only need to match the bits in operational space. Figure 4 illustrates an example of a possible optical setup that could be used to carry out such an experiment for a QKD system with polarization encoding. A delayed Mach-Zehnder interferometer allows two consequent pulses to overlap on a beam splitter with two SPDs at the output. A half-wave plate is used to match the polarization of different bits. It can also be used instead of the time delay to vary the degree of mode overlap.

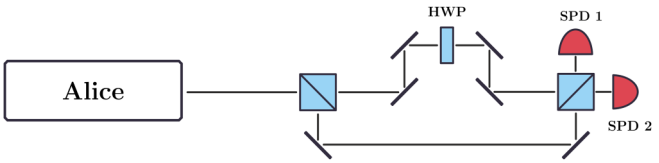

FIG. 4. Concept of a possible optical scheme that can be used for a HOM-based bases test of a polarization-encoding QKD system. One of the interferometer arms is delayed, so that two consequent pulses are matched on a beam splitter. A half-wave plate is used for polarization matching of two states. This testing routine can be done on the fly during a QKD session, and the photodetection events are postselected to collect statistics on the desired pair of states within a random pulse sequence.

Finally, we establish a connection between visibility and fidelity. In Appendix A we derive an equation to calculate fidelity for two arbitrary PRWCPs $(\hat{\rho}_1, \hat{\rho}_2)$ with equal mean photon number μ depending on their HOM visibility:

$$\sqrt{F(\hat{\rho}_1, \hat{\rho}_2)} = \exp(\mu(\sqrt{2\,\mathrm{V}_{\hat{\rho}_1\hat{\rho}_2}} - 1)). \tag{11}$$

We substitute this result into Eq. (10). In principle, every two pulses have their own value of visibility, but to simplify the visualization of the final result, we demonstrate the special case when every pair of bits has the same visibility value V:

$$1 - 2\Delta \geqslant \cos\left(2\arccos\frac{1 + e^{\mu(\sqrt{2V}-1)}}{2} + \arccos e^{\mu(\sqrt{2V}-1)}\right). \tag{12}$$

IV. KEY GENERATION RATE

In the previous section we have shown how HOM experiments' results could be used for upper-bounding the degree of distinguishability for both BB84 and decoy-state techniques. The final goal of this procedure is to find the secret key rate, depending on the measurement results. In this section we simulate the key generation rate for a system with BB84 protocol and two decoy states (i.e., three intensities). The asymptotic key generation rate can be lower bounded as follows [28]:

$$K \geqslant \max_{I_s, I_d}\left[\tfrac{1}{2}\left(p^s_1 Y^s_{1L}\left[1 - h\left(e^s_{1U}\right)\right] - f(E^s)Q^s h(E^s)\right)\right]. \tag{13}$$

Here maximization is done among the intensities of signal (I_s) and decoy (I_d) states. The second decoy-state intensity is considered to be zero. The probability of generating a single photon for a signal state is $p^s_1 = I_s e^{-I_s}$, according to the Poisson distribution. The lower bound for the yield of a single photon for signal states is Y^s_{1L}, and e^s_{1U} is the upper bound for the single-photon error rate within a signal state. Q^s is the signal state gain, $f(E^s)$ is the efficiency of error correction, and $h(E^s)$ is the binary Shannon entropy with the signal error rate E^s as an argument.

The parameter affected by the bases imperfections is e^s_{1U}, as Eve has an ability to get more information with the same overall error rate.

For simulation we use the error correction coefficient $f(E^s) = 1.2$ and fiber attenuation 0.2 dB/km, Bob's device losses 3 dB; and we use detector efficiency 25%, optical error

rate 1%, and dark count probability per gate 10^{-5}. These values match with the ones used in previous simulations [4,5].

According to the state-of-the-art security models, Bob has to perform active and continuous calibration on his optical transmittance and single-photon detector efficiency to prevent Eve from manipulating it [1,2,29]. This important requirement allows to treat Bob's losses as calibrated. If this assumption is omitted, this means that Bob has no control on his efficiency and in the limit can be under a blinding attack that totally compromises the key [30].

Finally, to minimize the number of displayed parameters we substitute the same visibility values for all pairs of sources within the simulation.

To simulate the effect of bases distinguishability, we use a method developed in Ref. [4]. The calculated imbalance is corrected taking into account the ability of Eve to use a lossless channel:

$$\Delta' = \frac{\Delta}{\tilde{Y}^s_{1L}}, \tag{14}$$

where $\tilde{Y}^s_{1L}$ is the minimum single-photon yield among two bases, obtained using the decoy-state method. Here we assume that the decoy-state method does not have any additional vulnerabilities. Once corrected, the bases imbalance is included in the upper bound of single-photon error of signal state:

$$\left(e^s_{1U}\right)' = 4(1-\Delta')\Delta'\left(1-2e^s_{1U}\right) + 4(1-2\Delta')\sqrt{\Delta'(1-\Delta')e^s_{1U}\left(1-e^s_{1U}\right)}. \tag{15}$$

With the help of the above estimations, we perform a simulation of key generation rate. Visibility values used for simulation are chosen so that they can be compared with current experimental results that we discuss in the next section.

Simulation results shown in Fig. 5 indicate that the key generation rate is highly sensitive to the distinguishability of the bases of the BB84 protocol (Fig. 5). Even for pretty high values of visibility like 0.495, the key rate drops significantly. On the other hand, for visibility as low as 0.47, key generation is still possible.

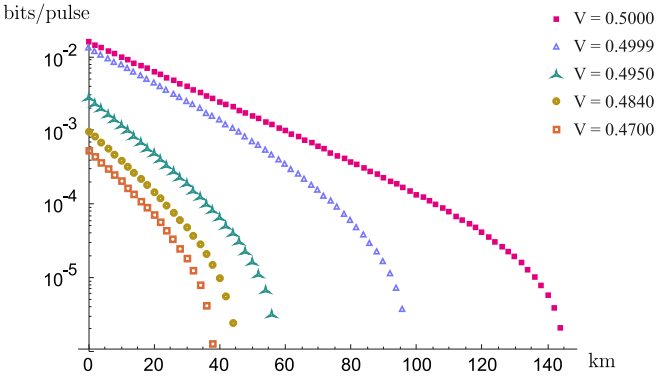

FIG. 5. Key generation rate depending on the communication distance for different values of HOM visibility for mode-matched states.

V. APPLICATIONS AND DISCUSSION

To check applicability of the proposed method for practical setups, we take the best-to-date fourth-order PRWCP interference visibilities from various experiments carried out for the MDI-QKD [3,31–34] (Table I).

According to the results of the previous section, it is clear that all listed results provide a positive key rate. It should be also noted that the latest results saturate the theoretical limit [34]. This fact allows to expect more precise experiments in the future.

As it has been mentioned [34], one of the main reasons for nonperfect interference between the pulses from independent gain-switched semiconductor laser diodes is the combination of frequency chirp and emission time jitter. While the former is a vulnerability that could be used by Eve to extract information about the states, the latter can be truly quantum as the time uncertainty includes spontaneous emission jitter. This leads to possible underestimation of the key generation rate, as quantum jitter dramatically decreases the interference visibility but does not compromise the setup.

To date, the best visibility has been achieved with the help of optical seeding, which leads to reduction of jitter from approximately 30 to 10 ps [34]. Moreover, authors performed postselection on the results, removing double clicks which are separated by a time interval longer than the full width at half maximum (FWHM) of the double click peak, and improved the visibility up to 0.499 ± 0.004 saturating the theoretical limit. Since we try to upper-bound all possible differences between the pulses, it is not clear if this kind of postselection can be used for distinguishability evaluation purposes. If we are sure that jitter is mainly quantum and out of Eve's control, such a kind of postselection could be treated as removing the events with the worst time overlap due to the quantum randomness of the jitter. In this case we make a tighter bound on Eve's abilities. We have to be sure, however, that the jitter is not caused by the classical effects, which Eve may possibly account for.

The results by Comandar *et al.* [34] show that pulsed laser seeding is a promising technology to design a modulator-free polarization encoding device with multiple sources. Indeed, the method proposed in this paper together with such a design allows to construct a provably secure source, as an alternative to current ones, used in satellite QKD.

It should also be mentioned that the proposed method is limited by the single-photon detector wavelength sensitivity. If the source contains side channels in the wavelength range that are out of the detector's sensitivity, or even not electromagnetic, this kind of side channel will not be detected by the proposed approach. One such example is the first ever

TABLE I. Best-to-date HOM visibility results with PRWCPs.

	HOM visibility
Ferreira da Silva *et al.* [31]	0.478
Rubenok *et al.* [32]	0.47 ± 0.01
Tang *et al.* [33]	0.475 ± 0.010
Comandar *et al.* [34] (without postselection)	0.487 ± 0.003
Comandar *et al.* [34] (with postselection)	0.499 ± 0.004

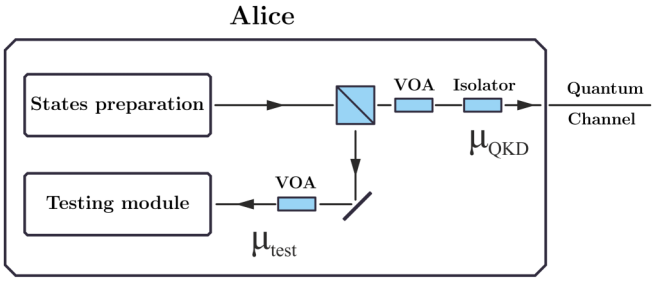

FIG. 6. The proposed testing module for the visibility measurement can be integrated into Alice's protected area, so that verification of indistinguishability can be carried out continuously. The light intensities μ_{test} and μ_{QKD} for the visibility measurement and for the communication signal, respectively, can be optimized independently. Optical isolation at the output of the device protects both the state preparation part and the testing module from Eve's potential light injections.

QKD prototype, where one could hear which state has been generated [7,35] due to the high-voltage power supply of the Pockels cell.

Even if the laser sources have been perfectly matched at some point, there is no guarantee that their behavior does not change over a long time interval. To prevent this kind of a loophole, one has to repeat the procedure described above continuously. An automatized testing module can be integrated into Alice's device, which is shown in Fig. 6. After the state preparation, optical signals enter a beam splitter that takes a part of the signal to the testing module (shown separately in Fig. 4). Variable optical attenuators can be used to set two independent amplitudes μ_{test} and μ_{QKD} for the visibility measurement and for the communication signal, respectively.

The mean photon number per pulse, μ_{QKD}, varies depending on the quantum channel parameters (e.g., length of the communication channel, bit error rate, etc.). The mean photon number per pulse, μ_{test}, used for testing can also be optimized. If it is too small, then the time to collect statistics can be too long. On the other hand, as it is shown in Sec. II, very high mean photon number leads to a nonlinear dependence of the single-photon visibility on the measurable PRWCP visibility (see Fig. 3), and results in underestimation of the secret key rate since PRWCP visibility is significantly lower than the maximum value of 0.5.

This conceptual procedure allows Alice to control the mismatch of her lasers on the fly and continuously monitor potential passive side channels.

As far as the beam splitter is a part of Alice's protected area, the losses introduced by it should not be treated as quantum channel losses. In other words the mean photon number included in the model is the value measured outside the whole system. To prevent Eve from having any influence on the testing procedure, optical isolators that are used to bound the THA [4,5] should be placed at the output of Alice, after the beam splitter.

The procedure described above is analogous to the method for calibration of Bob's detection efficiency, mentioned in the previous section [1]. Bob has his own light source to check both his optical losses and single-photon detectors' characteristics. The common idea between these methods is that the legitimate parties use a verification module inside their protected area to check the performance of the QKD components.

The same approach can be straightforwardly implemented for any other quantum cryptography tasks, which suffer from the same kind of vulnerability because their optical setups are usually similar. For example, experimentally demonstrated quantum digital signatures use the same four laser schematics together with the assumption of the states' indistinguishability to prevent Eve from forgery [13,36].

VI. CONCLUSION

We introduce an explicit method for an integral evaluation of passive side-channel information leakage due to the optical mode mismatch in the emitted pulses. We show that for relatively small signal intensities, fourth-order interference visibility linearly depends on the degree of nonorthogonality between two PRWCPs. This allows us to relate the results of Hong-Ou-Mandel experiments with the existing bases imbalance security proofs. We include the typical experimentally obtained visibility values in the security model, and calculate the key generation rate.

Our results indicate that the secret key is highly sensitive to the HOM interference visibility value. In order to maintain a reasonably long communication distance, the visibility must be very close to the theoretical limit of 0.5. We note that the experimentally obtainable value of interference visibility depends not only on the actual distinguishability of optical signals, but also on various imperfections of the measurement setup, i.e., the jitter of laser pulses. Thus our approach gives an upper bound on signal distinguishability and hence a lower bound on the secret key rate.

In this work we address the conventional BB84 protocol with decoy states; however, the concept could be applied to other protocols as well. Current experimental data have been used to check the applicability of our method to real-world systems. Future studies on this topic may include estimation of distinguishability between signal and decoy states, derivation of tighter bounds on the secret key rate, as well as security analysis of other protocols. We anticipate the developed method is highly prospective in terms of real-world QKD certification applications.

ACKNOWLEDGMENTS

A.D. acknowledges support from Russian Science Foundation Grant No. 17-71-20146 for work reported in the sections "Side-channel information and HOM visibility" (Sec. III) and "Key generation rate" (Sec. IV). D.S. acknowledges funding by RFBR, Sirius University of Science and Technology, JSC Russian Railways and Educational Fund "Talent and Success" Project No 20-32-51004 for work reported in the section "Hong-Ou-Mandel interference and nonorthogonality of quantum states" (Sec. II).

APPENDIX A: FIDELITY ESTIMATION

In this section we calculate the fidelity for two PRWCPs with density matrices $\hat{\rho}_1$ and $\hat{\rho}_2$, from Eq. (2) as a function of HOM visibility:

$$\hat{\rho}_1 = \sum_{n=0}^{\infty} \rho_{nn} |n\rangle_1\langle n|_1 = \sum_{n=0}^{\infty} e^{-\mu^2} \frac{\mu^n}{n!} (\hat{a}_1^\dagger)^n |0\rangle\langle 0| (\hat{a}_1)^n, \tag{A1}$$

$$\hat{\rho}_2 = \sum_{n=0}^{\infty} \frac{\rho_{nn}}{n!} (\sqrt{\gamma}\hat{a}_1^\dagger + \sqrt{1-\gamma}\hat{a}_{\perp 1}^\dagger)^n |0\rangle\langle 0| (\sqrt{\gamma}\hat{a}_1 + \sqrt{1-\gamma}\hat{a}_{\perp 1})^n. \tag{A2}$$

Here we decomposed the $a_2^\dagger$ operator into a sum of the $a_1^\dagger$ and $a_{\perp 1}^\dagger$. Therefore, $a_{\perp 1}^\dagger$ represents all the modes that are orthogonal to $a_1^\dagger$. The parameter γ indicates the similarity of these states. We rewrite $\hat{\rho}_2$ as

$$\hat{\rho}_2 = \sum_{n=0}^{\infty} \rho_{nn} \left(\sum_{k=0}^{n} \sqrt{C_n^k} (\sqrt{\gamma})^k (\sqrt{1-\gamma})^{n-k} |k\rangle_1 |n-k\rangle_{\perp 1} \right) \left(\sum_{q=0}^{n} \langle q|_1 \langle n-q|_{\perp 1} \sqrt{C_n^q} (\sqrt{\gamma})^q (\sqrt{1-\gamma})^{n-q} \right). \tag{A3}$$

For simplicity we replace $(1-\gamma)$ with a new letter, δ. Visualizing the density matrices $\hat{\rho}_1$ and $\hat{\rho}_2$, we have

$$\hat{\rho}_1 = \begin{pmatrix} \rho_{00} & & & \\ & \rho_{11}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & & \\ & & \rho_{22}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \\ & & & \cdots \end{pmatrix}, \tag{A4}$$

$$\hat{\rho}_2 = \begin{pmatrix} \rho_{00} & & & \\ & \rho_{11}\begin{pmatrix} \gamma & \sqrt{\gamma\delta} \\ \sqrt{\gamma\delta} & \delta \end{pmatrix} & & \\ & & \rho_{22}\begin{pmatrix} \gamma^2 & \sqrt{2\gamma^3\delta} & \gamma\delta \\ \sqrt{2\gamma^3\delta} & 2\gamma\delta & \sqrt{2\gamma\delta^3} \\ \gamma\delta & \sqrt{2\gamma\delta^3} & \delta^2 \end{pmatrix} & \\ & & & \cdots \end{pmatrix}. \tag{A5}$$

According to the fidelity definition (7), we straightforwardly obtain the result

$$\sqrt{F(\hat{\rho}_1, \hat{\rho}_2)} = \rho_{00} + \sum_{n=1}^{\infty} \rho_{nn} \gamma^{\frac{n}{2}} = \exp(\mu(\sqrt{\gamma} - 1)). \tag{A6}$$

Now we address the visibility equation (5) and find that $V_{\hat{\rho}_1\hat{\rho}_2} = \frac{\gamma}{2}$. So the final result is

$$\sqrt{F(\hat{\rho}_1, \hat{\rho}_2)} = \exp(\mu(\sqrt{2V_{\hat{\rho}_1\hat{\rho}_2}} - 1)). \tag{A7}$$

APPENDIX B: BASES IMBALANCE ESTIMATION

To include bases flaw into the security model one needs to calculate the fidelity between two bases ($F(\hat{\rho}_x, \hat{\rho}_z)$). Density matrices of the X and Z bases could be expressed as follows:

$$\hat{\rho}_x = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \hat{\rho}_{x,0}^\lambda + \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \hat{\rho}_{x,1}^\lambda, \tag{B1}$$

$$\hat{\rho}_z = \frac{1}{2}\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \otimes \hat{\rho}_{z,0}^\lambda + \frac{1}{2}\begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix} \otimes \hat{\rho}_{z,1}^\lambda. \tag{B2}$$

Here, $\hat{\rho}_{i,j}^\lambda$ represents density matrices of nonoperational degrees of freedom.

We calculate the fidelity between the bases in two steps. First, we introduce a set of auxiliary density matrices; we then calculate the fidelity between the initial and auxiliary states and then, between auxiliary matrices for different bases. After that, we use the distance and apply triangle inequalities, that allow us to bound the fidelity between the initial states. The idea of the estimation is illustrated in Fig. 7.

We define auxiliary density matrices as bases for which nonoperational degrees of freedom are the same for both bits:

$$\hat{\rho}_i^j = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \hat{\rho}_{i,j}^\lambda + \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \hat{\rho}_{i,j}^\lambda. \tag{B3}$$

We calculate the fidelity between $\hat{\rho}_x$ and an auxiliary matrix, which for both bits are the same in nonoperational degrees of freedom. Without loss of generality, we start with a case when both bits have a density matrix $\hat{\rho}_{x,0}^\lambda$ for nonoperational parameters: $\hat{\rho}_x^0 = \frac{1}{2}\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \otimes \hat{\rho}_{x,0}^\lambda$,

$$F(\hat{\rho}_x, \hat{\rho}_x^0) = F\left(\frac{1}{2}\begin{pmatrix} \hat{\rho}_{x,0}^\lambda & 0 \\ 0 & \hat{\rho}_{x,1}^\lambda \end{pmatrix}, \frac{1}{2}\begin{pmatrix} \hat{\rho}_{x,0}^\lambda & 0 \\ 0 & \hat{\rho}_{x,0}^\lambda \end{pmatrix} \right). \tag{B4}$$

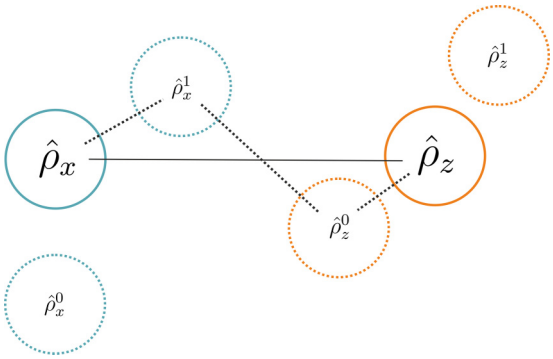

FIG. 7. Scheme of bases fidelity estimation. A pair of auxiliary matrices for each basis allow to bound the distance using two triangle inequalities. The distance between two closest auxiliary matrices is used.

As a result, using the fidelity definition, we obtain

$$\sqrt{F\left(\hat{\rho}_x, \hat{\rho}_x^0\right)} = \frac{1+\sqrt{F\left(\hat{\rho}_{x,0}^\lambda, \hat{\rho}_{x,1}^\lambda\right)}}{2} = \sqrt{F\left(\hat{\rho}_x, \hat{\rho}_x^1\right)}. \quad \text{(B5)}$$

The fidelities between the initial $\hat{\rho}_x$ and both auxiliary states $\hat{\rho}_x^0$ and $\hat{\rho}_x^1$ are the same. A similar result can be obtained for the Z basis as a Hadamard transformation can be applied to the operational degree of freedom for $\hat{\rho}_z$, so the resulting fidelity is the same:

$$\sqrt{F\left(\hat{\rho}_z, \hat{\rho}_z^0\right)} = \frac{1+\sqrt{F\left(\hat{\rho}_{z,0}^\lambda, \hat{\rho}_{z,1}^\lambda\right)}}{2} = \sqrt{F\left(\hat{\rho}_z, \hat{\rho}_z^1\right)}. \quad \text{(B6)}$$

Now, to apply the triangle inequality, we use the Bures angle [27], which constitutes a metric in Hilbert space:

$$\mathrm{A}(\hat{\rho}_1, \hat{\rho}_2) = \arccos\sqrt{F(\hat{\rho}_1, \hat{\rho}_2)}. \quad \text{(B7)}$$

As soon as we find the fidelity values between the initial states and the auxiliary ones, we need to look for a pair of auxiliary matrices from both bases that have maximum fidelity (minimum distance).

As a result of two triangle inequalities we obtain a result:

$$\arccos\sqrt{F(\hat{\rho}_x, \hat{\rho}_z)} \leqslant \arccos\max_{i,j\in 0,1}\sqrt{F\left(\hat{\rho}_{x,i}^\lambda, \hat{\rho}_{z,j}^\lambda\right)} + \sum_{b\in\{x,z\}}\arccos\frac{1+\sqrt{F\left(\hat{\rho}_{b,0}^\lambda, \hat{\rho}_{b,1}^\lambda\right)}}{2}. \quad \text{(B8)}$$

The final result is expressed as follows:

$$\Delta \leqslant \frac{1}{2} - \frac{1}{2}\cos\left(\arccos\max_{i,j\in 0,1}\sqrt{F\left(\hat{\rho}_{x,i}^\lambda, \hat{\rho}_{z,j}^\lambda\right)} + \sum_{b\in\{x,z\}}\arccos\frac{1+\sqrt{F\left(\hat{\rho}_{b,0}^\lambda, \hat{\rho}_{b,1}^\lambda\right)}}{2}\right). \quad \text{(B9)}$$

In the special case when all fidelities equal the same value F, Eq. (B9) simplifies to

$$\Delta \leqslant \frac{1}{2} - \frac{1}{2}\cos\left(2\arccos\frac{1+\sqrt{F}}{2} + \arccos\sqrt{F}\right). \quad \text{(B10)}$$

[1] Ø. Marøy, V. Makarov, and J. Skaar, Secure detection in quantum key distribution by real-time calibration of receiver, Quantum Sci. Technol. **2**, 044013 (2017).

[2] A. Huang, S.-H. Sun, Z. Liu, and V. Makarov, Quantum key distribution with distinguishable decoy states, Phys. Rev. A **98**, 012330 (2018).

[3] L. Comandar, M. Lucamarini, B. Fröhlich, J. Dynes, A. Sharpe, S.-B. Tam, Z. Yuan, R. Penty, and A. Shields, Quantum key distribution without detector vulnerabilities using optically seeded lasers, Nat. Photon. **10**, 312 (2016).

[4] M. Lucamarini, I. Choi, M. B. Ward, J. F. Dynes, Z. L. Yuan, and A. J. Shields, Practical Security Bounds Against the Trojan-Horse Attack in Quantum key Distribution, Phys. Rev. X **5**, 031030 (2015).

[5] K. Tamaki, M. Curty, and M. Lucamarini, Decoy-state quantum key distribution with a leaky source, New J. Phys. **18**, 065008 (2016).

[6] M. Pereira, M. Curty, and K. Tamaki, Quantum key distribution with flawed and leaky sources, npj Quantum Inf. **5**, 62 (2019).

[7] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, Experimental quantum cryptography, J. Cryptol. **5**, 3 (1992).

[8] 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. **4**, 43 (2002).

[9] T. Schmitt-Manderbach, H. Weier, M. Fürst, R. Ursin, F. Tiefenbacher, T. Scheidl, J. Perdigues, Z. Sodnik, C. Kurtsiefer, J. G. Rarity *et al.*, Experimental Demonstration of Free-Space Decoy-State Quantum Key Distribution over 144 km, Phys. Rev. Lett. **98**, 010504 (2007).

[10] M. García-Martínez, N. Denisenko, D. Soto, D. Arroyo, A. Orue, and V. Fernandez, High-speed free-space quantum key distribution system for urban daylight applications, Appl. Opt. **52**, 3311 (2013).

[11] G. Vallone, V. D'Ambrosio, A. Sponselli, S. Slussarenko, L. Marrucci, F. Sciarrino, and P. Villoresi, Free-Space Quantum Key Distribution by Rotation-Invariant Twisted Photons, Phys. Rev. Lett. **113**, 060503 (2014).

[12] C.-Z. Peng, J. Zhang, D. Yang, W.-B. Gao, H.-X. Ma, H. Yin, H.-P. Zeng, T. Yang, X.-B. Wang, and J.-W. Pan, Experimental Long-Distance Decoy-State Quantum key Distribution Based on Polarization Encoding, Phys. Rev. Lett. **98**, 010505 (2007).

[13] H.-L. Yin, Y. Fu, H. Liu, Q.-J. Tang, J. Wang, L.-X. You, W.-J. Zhang, S.-J. Chen, Z. Wang, Q. Zhang *et al.*,

Experimental quantum digital signature over 102 km, Phys. Rev. A **95**, 032334 (2017).
[14] 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. **11**, 065001 (2009).
[15] T. L. Koch and J. E. Bowers, Nature of wavelength chirping in directly modulated semiconductor lasers, Electron. Lett. **20**, 1038 (1984).
[16] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
[17] C.-K. Hong, Z.-Y. Ou, and L. Mandel, Measurement of Subpicosecond Time Intervals between Two Photons by Interference, Phys. Rev. Lett. **59**, 2044 (1987).
[18] C. Osorio, N. Sangouard, and R. T. Thew, On the purity and indistinguishability of down-converted photons, J. Phys. B: At. Mol. Opt. Phys. **46**, 055501 (2013).
[19] Z. L. Yuan, M. Lucamarini, J. F. Dynes, B. Fröhlich, M. B. Ward, and A. J. Shields, Interference of Short Optical Pulses from Independent Gain-Switched Laser Diodes for Quantum Secure Communications, Phys. Rev. Appl. **2**, 064006 (2014).
[20] L. Gui-Lu, General quantum interference principle and duality computer, Commun. Theor. Phys. **45**, 825 (2006).
[21] J. Beugnon, M. P. Jones, J. Dingjan, B. Darquié, G. Messin, A. Browaeys, and P. Grangier, Quantum interference between two single photons emitted by independently trapped atoms, Nature (London) **440**, 779 (2006).
[22] A. Allevi, M. Bondani, P. Marian, T. A. Marian, and S. Olivares, Characterization of phase-averaged coherent states, J. Opt. Soc. Am. B **30**, 2621 (2013).
[23] M. Koashi, Simple security proof of quantum key distribution based on complementarity, New J. Phys. **11**, 045018 (2009).
[24] M. Koashi and J. Preskill, Secure Quantum Key Distribution with an Uncharacterized Source, Phys. Rev. Lett. **90**, 057902 (2003).
[25] A. Uhlmann, The transition probability in the state space of a *-algebra, Rep. Math. Phys. **9**, 273 (1976).
[26] R. Jozsa, Fidelity for mixed quantum states, J. Mod. Opt. **41**, 2315 (1994).
[27] Z. Ma, F.-L. Zhang, and J.-L. Chen, Fidelity induced distance measures for quantum states, Phys. Lett. A **373**, 3407 (2009).
[28] X. Ma, B. Qi, Y. Zhao, and H.-K. Lo, Practical decoy state for quantum key distribution, Phys. Rev. A **72**, 012326 (2005).
[29] Ø. Marøy, L. Lydersen, and J. Skaar, Security of quantum key distribution with arbitrary individual imperfections, Phys. Rev. A **82**, 032337 (2010).
[30] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, Hacking commercial quantum cryptography systems by tailored bright illumination, Nat. Photon. **4**, 686 (2010).
[31] T. Ferreira da Silva, D. Vitoreti, G. B. Xavier, G. C. do Amaral, G. P. Temporão, and J. P. von der Weid, Proof-of-principle demonstration of measurement-device-independent quantum key distribution using polarization qubits, Phys. Rev. A **88**, 052303 (2013).
[32] A. Rubenok, J. A. Slater, P. Chan, I. Lucio-Martinez, and W. Tittel, Real-World Two-Photon Interference and Proof-of-Principle Quantum Key Distribution Immune to Detector Attacks, Phys. Rev. Lett. **111**, 130501 (2013).
[33] Z. Tang, Z. Liao, F. Xu, B. Qi, L. Qian, and H.-K. Lo, Experimental Demonstration of Polarization Encoding Measurement-Device-Independent Quantum Key Distribution, Phys. Rev. Lett. **112**, 190503 (2014).
[34] L. Comandar, M. Lucamarini, B. Fröhlich, J. Dynes, Z. Yuan, and A. Shields, Near perfect mode overlap between independently seeded, gain-switched lasers, Opt. Express **24**, 17849 (2016).
[35] G. Brassard, Brief history of quantum cryptography: A personal perspective, in *IEEE Information Theory Workshop on Theory and Practice in Information-Theoretic Security, 2005* (IEEE, Piscataway, NJ, 2005), pp. 19–23.
[36] A. Huang, S. Barz, E. Andersson, and V. Makarov, Implementation vulnerabilities in general quantum cryptography, New J. Phys. **20**, 103016 (2018).

Correction: Funding information for the second author was incomplete and has been fixed.