Explicit attacks on the Bennett-Brassard 1984 protocol with partially distinguishable photons

D. Babukhin,[1,2] D. Kronberg,[2] and D. Sych[2,3]
[1]*QRate LLC, Novaya Avenue 100, Moscow 143026, Russia*
[2]*Department of Mathematical Methods for Quantum Technologies, Steklov Mathematical Institute, Russian Academy of Sciences, Gubkina Street 8, Moscow 119991, Russia*
[3]*P.N. Lebedev Physical Institute, Russian Academy of Sciences, 53 Leninskiy Prospekt, Moscow 119991, Russia*

(Received 8 June 2022; revised 15 September 2022; accepted 20 September 2022; published 3 October 2022)

The distinguishability of photons in nonoperational degrees of freedom compromises the unconditional security of quantum key distribution since an eavesdropper can improve attack strategies by exploiting this distinguishability. However, the optimal eavesdropping strategies in the presence of light-source side channels are not known. Here we provide several explicit attack strategies on the Bennett-Brassard 1984 (BB84) protocol with partially distinguishable photons. In particular, we consider the phase-covariant cloning attack, which is optimal in the absence of side channels, and show that there are even better strategies in the presence of side channels. The improved strategies exploit a measurement of the side-channel state followed by an attack on the signal photon. Our results explicitly demonstrate a reduction of the critical error rate and set an upper bound on the practical secret key rate.

DOI: 10.1103/PhysRevA.106.042403

I. INTRODUCTION

Quantum key distribution (QKD) is a method to share a secret key between two legitimate parties (Alice and Bob) in the presence of an eavesdropper (Eve) [1,2]. This possibility is based on the no-cloning theorem, which prohibits information gain from a carrier photon without disturbing its quantum state [3,4]. In theory, QKD is unconditionally secure [5,6]. In practice, however, there are numerous deviations of practical QKD from theory [7], which leads to loopholes and side channels [8] and, consequently, to quantum hacking [9,10]. The practical security of QKD is an active research topic [11,12].

Loopholes on Bob's (receiver) side can be closed with the use of a measurement-device-independent QKD protocol [13–15] and twin-field QKD [16,17], where measurement devices are excluded from the private space of the legitimate sides and transferred to a third (untrusted) party. This protocol allows getting rid of all hacking attacks on the measurement devices, which made up the majority of already known hacking strategies in practical QKD [8]. But this protocol does not close loopholes on Alice's (transmitting) side, which require specific characterization and countermeasures ([18–21]).

Light-source side channels can be described with a nonoperational degree of freedom [18,22] from which Eve gains information. For example, if Alice and Bob use polarization encoding in the QKD protocol, then all other photon degrees of freedom (e.g., spatial and temporal profile) are nonoperational for QKD. The efficiency of eavesdropping depends on Eve's choice of how to measure and process information from the side channel. In the Bennett-Brassard 1984 (BB84) protocol without side channels, the critical quantum bit error rate of 11% is an information-theoretic result [6], which is also achieved with an explicit eavesdropping strategy [23]. But if photons used in the BB84 protocol are partially distinguishable in other than operational degrees of freedom (e.g., photons' spatial profiles do not perfectly coincide in BB84 with polarization encoding), the security of the protocol is lower. In other words, the critical quantum bit error rate is less than 11%. In BB84 with partially distinguishable photons, the lower bound on the possible secret key rate was provided in [21]. This result shows a pessimistically low secret key generation rate. The explicit attack strategy, which gives this lower bound of the secret key rate, is unknown. Previous research on this issue [24–26] has not reached the secret key rate of [21], thus opening a gap between explicit attacks and theoretical secrecy. Therefore, investigating explicit attack strategies can help close this gap.

In this paper, we provide explicit eavesdropping strategies on the BB84 protocol in the presence of a binary side channel of the light source. The strategies consist of a measurement of the side-channel state and an attack on the signal photon. We investigate two main types of measurements: unambiguous state discrimination (USD) and minimum-error (ME) measurement. For unambiguous state discrimination of the side channel, there are two options: in the case of a conclusive result, Eve obtains full information about the quantum signal, while in the case of an inconclusive result, Eve performs the standard optimal cloner attack. Therefore, we expect a reduction of the critical error rate. For the minimum-error side-channel measurement, the measurement result introduces a bias in the signal states' ensemble probabilities. Eve can use this bias as an additional binary information source, or alternatively, she can use it to adjust the eavesdropping strategy on the signal photon with the use of postselection. We use a Hong-Ou-Mandel (HOM) interference visibility [21] to estimate information leakage through the side channel and

calculate how the critical error rate depends on the HOM visibility.

This paper is organized as follows. In Sec. II we formulate the BB84 protocol in the presence of a binary-source side channel and discuss possible strategies for information extraction. In Sec. III we formulate three eavesdropping strategies on the protocol, and in Sec. IV we discuss and conclude our results.

II. BB84 PROTOCOL WITH PARTIALLY DISTINGUISHABLE PHOTONS

A. Side-channel model

In the balanced BB84 protocol, Alice and Bob use two equiprobable bases of quantum states to distribute bits of a secret key: the X basis $\{|0\rangle_x, |1\rangle_x\}$ and the Y basis $\{|0\rangle_y, |1\rangle_y\}$. Alice randomly chooses a secret bit value (0 or 1) and randomly chooses a basis (X or Y) and sends a quantum state into the communication channel. This leads to the following ensemble of states:

$$\left\{\tfrac{1}{4} : |0_x\rangle, \tfrac{1}{4} : |1_x\rangle, \tfrac{1}{4} : |0_y\rangle, \tfrac{1}{4} : |1_y\rangle\right\}. \tag{1}$$

Eavesdropping introduces mixedness in this ensemble, which occurs due to the entanglement of the state of the carrier photon with an auxiliary quantum system. By measuring her quantum system, Eve obtains outcomes which correlate with secret bits. At the same time, eavesdropping introduces errors in Bob's measurement results. These errors allow estimating the leakage of information to the eavesdropper. This eavesdropping-error connection is guaranteed with the no-cloning theorem and is a cornerstone of quantum key distribution.

The eavesdropper has an informational side channel when the photon source produces photons with physical distinguishability besides the signal degree of freedom. For example, if the protocol uses polarization encoding, other photon physical properties such as spatial and temporal degrees of freedom can be correlated with polarization and hence leak information about the polarization state. The eavesdropper can use these correlations to enhance the distinguishability of photons and thus potentially get more information about the encoded secret bit than assumed in a QKD protocol. To estimate this information leakage, a model of nonoperational degrees of freedom should be incorporated into the security analysis.

In this work, we consider a specific binary model of the side channel of the photon source. This model accounts for the photon distinguishability of Alice's photon source while providing Eve no information either about the secret bit or about the basis choice. Thus, Alice's choice of basis remains equiprobable to Eve even in the presence of a side channel; hence, the protocol is still basis balanced. We note that in the current practical QKD systems the basis choice can be intentionally unbalanced; that is, one basis is used by Alice more often than the other in order to generate the secret key more efficiently. In our consideration we do not take this imbalance into account. When Eve measures a state of a side channel, she finds that one pair of quantum states is more probable in a quantum channel. At the same time, these states belong to different bases and encode different bits; thus, this is not a side channel of the form "0 is more probable than 1." Our model is interesting from the point of eavesdropping possibilities: even such an uninformative side channel allows Eve to increase the eavesdropping efficiency.

The model of the side channel is

$$\{|0_x\rangle \otimes |0_\Delta\rangle, |1_x\rangle \otimes |1_\Delta\rangle\} \tag{2}$$

for the X basis and

$$\{|0_y\rangle \otimes |1_\Delta\rangle, |1_y\rangle \otimes |0_\Delta\rangle\} \tag{3}$$

for the Y basis, where $|0_\Delta\rangle$ and $|1_\Delta\rangle$ are two nonorthogonal states $\langle 0_\Delta||1_\Delta\rangle = \Delta$ of a side-channel degree of freedom. This form of the side channel allows for special eavesdropping strategies, which we describe in the following sections.

B. Hong-Ou-Mandel visibility as a measure of side-channel leakage

Legitimate sides can use Hong-Ou-Mandel interference to keep track of information that leaks to Eve through the photon-distinguishability side channel. The Hong-Ou-Mandel interference is a two-photon interference that prohibits two indistinguishable photons from exiting at different ends of a balanced beam splitter [27]. If there is any photon distinguishability of the two photons, incident on a beam splitter, there will be a nonzero probability to measure the photon counts at both exits of a beam splitter in photodetectors placed at each exit. These coincidence counts indicate the physical distinguishability of photons for any physical difference and can be a sign that photons contain additional information available to the eavesdropper.

Let us consider a two-photon state which is incident on a balanced beam splitter. The initial state of the two photons is

$$|\Psi_{\text{in}}\rangle_{ab} = a^\dagger_{p1} b^\dagger_{p2} |0_z\rangle_{ab} = |1; p_1\rangle_a |1; p_2\rangle_b, \tag{4}$$

where $a^\dagger$ and $b^\dagger$ are creation operators for modes a and b, which correspond to the incoming sides of a beam splitter and $p_{1,2}$ are arbitrary discrete degrees of freedom of the two photons. The unitary transform of the beam splitter is

$$U_{\text{BS}} a^\dagger = \frac{1}{\sqrt{2}} a^\dagger + \frac{1}{\sqrt{2}} b^\dagger, \tag{5}$$

$$U_{\text{BS}} b^\dagger = \frac{1}{\sqrt{2}} a^\dagger - \frac{1}{\sqrt{2}} b^\dagger. \tag{6}$$

Applying this operator to the initial two-photon state, one obtains an output state of the form

$$|\Psi_{\text{out}}\rangle_{ab} = \tfrac{1}{2}(a^\dagger_{p1} a^\dagger_{p2} + a^\dagger_{p2} b^\dagger_{p1} - a^\dagger_{p1} b^\dagger_{p2} - b^\dagger_{p1} b^\dagger_{p2})|0_z\rangle_{ab}. \tag{7}$$

When all degrees of freedom are the same for a pair of incident photons, then this state reduces to

$$|\Psi_{\text{out}}\rangle_{ab} = \frac{1}{2}(a^\dagger_p a^\dagger_p - b^\dagger_p b^\dagger_p)|0_z\rangle_{ab} = \frac{1}{\sqrt{2}}(|2; p\rangle_a - |2; p\rangle_b). \tag{8}$$

This state allows two photons to exit one of the exits of the beam splitter in pairs but not separately. If one places photodetectors at each end of the beam splitter, there will be no coincidence counts (counts of both photodetectors in the same time bin). In contrast, if there is a mismatch in any degrees of freedom of two incident photons, a state (7) will be generated

after the beam splitter, and there will be coincidence counts. These coincidence events will affect the visibility of the HOM interference visibility, defined as

$$V(\rho_1, \rho_2) = \mathrm{Tr}[\rho_1 \rho_2] = \frac{N_{\max} - N_{\min}}{N_{\max}}, \tag{9}$$

where $N_{\min}$ and $N_{\max}$ are minimum and maximum values of coincidence counts. If Alice makes the signal degrees of freedom of her photons equal (e.g., she encodes equal bits in the polarization of photons), she can test her photons for additional distinguishability with the HOM visibility [21]. If there is no photon-distinguishability side channel, the visibility is unity. If this is not the case and the visibility is less than unity, then there is additional distinguishability among the photons, which leaks information to Eve.

Let us apply the HOM interference to our consideration of a single-photon QKD. Alice's ensemble in the presence of the binary side channel has the following form:

$$\mathcal{E} = \left\{\tfrac{1}{4} : |0_x\rangle \otimes |0_\Delta\rangle, \tfrac{1}{4} : |1_x\rangle \otimes |1_\Delta\rangle, \tfrac{1}{4} : |0_y\rangle \otimes |1_\Delta\rangle, \tfrac{1}{4} : |1_y\rangle \otimes |0_\Delta\rangle\right\}. \tag{10}$$

To check the distinguishability of states with HOM interference, Alice produces two different photons, e.g., encoding states 0_x and 1_x. She then brings a signal degree of freedom of these photons to the equal quantum state—she transforms one of the photons in such a way that its quantum state transforms to the state of another photon [21]. For example, she transforms the state $|1_\Delta\rangle\langle 1_\Delta|$ to state $|0_x\rangle\langle 0_x|$. In the end, she has two photons with states $|0_x\rangle\langle 0_x| \otimes |0_\Delta\rangle\langle 0_\Delta|$ and $|0_x\rangle\langle 0_x| \otimes |1_\Delta\rangle\langle 1_\Delta|$, which she causes to interfere on a balanced beam splitter. The visibility value of this interference then reads

$$\begin{aligned}&\mathrm{Tr}[(|0_x\rangle\langle 0_x| \otimes |0_\Delta\rangle\langle 0_\Delta|)(|0_x\rangle\langle 0_x| \otimes |1_\Delta\rangle\langle 1_\Delta|)]\\ &\quad = \mathrm{Tr}[|0_x\rangle\langle 0_x||0_x\rangle\langle 0_x|]\mathrm{Tr}[|0_\Delta\rangle\langle 0_\Delta||1_\Delta\rangle\langle 1_\Delta|]\\ &\quad = |\langle 0_\Delta||1_\Delta\rangle|^2.\end{aligned} \tag{11}$$

So in our consideration, HOM visibility is parametrized with a scalar product of side-channel states.

III. EAVESDROPPING ELEMENTS IN THE BB84 PROTOCOL WITH PARTIALLY DISTINGUISHABLE PHOTONS

A. Side-channel-state measurement strategies

Knowledge of the photon-distinguishability side channel allows Eve a variety of strategies to attack the protocol. The choice here is how to combine measurements of the signal and side-channel degrees of freedom to gain information about secret bits. In the following, we consider two strategies of interest.

1. Unambiguous-state-discrimination measurement

With this strategy, Eve performs unambiguous state discrimination of the side-channel degree of freedom. This measurement provides full knowledge of the side-channel state or yields an inconclusive result with no information. Physically, this measurement describes a device which has $N + 1$ outputs and which is a target for a quantum state belonging to an ensemble of arbitrary (in general, nonorthogonal) N states $\{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_N\rangle\}$. The ith port of the device fires if a corresponding quantum state is measured, while the last $(N + 1)$th port fires when the device fails to recognize the quantum state. In general, there is a nonzero probability to fail the measurement because of the theorem that unknown nonorthogonal quantum states cannot be distinguished reliably [28].

In terms of positive operator-valued measure (POVM) operators, this measurement is formulated as follows: for a set of linearly independent quantum states $\{|\psi_1\rangle, \ldots, |\psi_n\rangle\}$ there exist POVM operators $M_i, i = 0, \ldots, n$, $I = \sum_{i=0}^{n} M_i$, such that

$$p_i = \mathrm{Tr}[|\psi_i\rangle\langle\psi_i|M_i], \quad \mathrm{Tr}[|\psi_j\rangle\langle\psi_j|M_i] = 0, \text{ if } j \neq i, \tag{12}$$

is the probability of conclusive results of measuring the ith state (if Eve has a conclusive measurement, she obtains full information about the measured state), and

$$p^{\mathrm{inc}} = 1 - \sum_{i=1}^{N} p_i \tag{13}$$

is the probability of an inconclusive measurement with no information about the state.

2. Minimum-error measurement

With this strategy, Eve makes a minimum-error measurement of the side-channel degree of freedom, which gives her information about the side-channel state with a minimum (but nonzero for nonorthogonal states) error probability. Physically, this measurement describes a device which has N outputs and which is a target for a quantum state belonging to an ensemble of arbitrary (in general, nonorthogonal) N states $\{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_N\rangle\}$. If the ith port of the device fires, then it is likely that a $|\psi_i\rangle$ state was sent towards the device. At the same time, there is a nonzero probability that the measured state was different from the $|\psi_i\rangle$ state. This is again a consequence of the theorem of distinguishability of nonorthogonal quantum states.

In terms of POVM operators, this measurement is formulated as follows: for a set of quantum states $\{q_1 : |\psi_1\rangle, \ldots, q_n : |\psi_n\rangle\}$, where q_i is the probability of sending the ith state towards the measurement device, there exist POVM operators $M_i, i = 1, \ldots, n, I = \sum_{i=1}^{n} M_i$, such that

$$p_i = \mathrm{Tr}[|\psi_i\rangle\langle\psi_i|M_i] \tag{14}$$

is the correct outcome probability when measuring the ith state and

$$p_i^{\mathrm{error}} = \sum_{\substack{j\neq i \\ j=1}}^{n} \mathrm{Tr}[|\psi_j\rangle\langle\psi_j|M_i] \neq 0 \tag{15}$$

is the probability of an error for the ith state measurement. The condition of minimal error is formulated as a constraint of the maximal average probability of the correct measurement outcome,

$$P_{\mathrm{opt}} = \max_{\{M_i\}} \sum_{i=1}^{n} q_i p_i. \tag{16}$$

B. Soft filtering of an ensemble of quantum states

One of the main assumptions at the base of BB84 protocol security is a random choice of the basis and the secret bit, which Alice uses to encode the secret bit into the state of the photon and to send it to Bob. Formally, this lack of information means equal probabilities of ensemble states (1). With source side channels, an eavesdropper can obtain information about states from nonoperational degrees of freedom. This leads to the equiprobability violation of ensemble states. If states of the ensemble have different probabilities, Eve can use soft filtering [29,30] to enhance her attack efficiency on the protocol. The soft filtering is an extension of the USD measurement, which in the case of success maps nonorthogonal quantum states of the ensemble $\{q_i : |\psi_i\rangle\langle\psi_i|, i = 1, \ldots, n\}$ into another ensemble $\{q_i\langle\psi_i|Q^{-1}|\psi_i\rangle : \frac{Q^{-\frac{1}{2}}|\psi_i\rangle\langle\psi_i|Q^{\frac{1}{2}}}{\langle\psi_i|Q^{-1}|\psi_i\rangle}, i = 1, \ldots, n\}$, where $Q = \sum_j q_j|\psi_j\rangle\langle\psi_j|$. This quantum channel is defined as follows:

$$\Phi(|\psi_i\rangle\langle\psi_i|) = F_{\text{succ}}|\psi_i\rangle\langle\psi_i|F_{\text{succ}}^{\dagger} + F_{\text{fail}}|\psi_i\rangle\langle\psi_i|F_{\text{fail}}^{\dagger}, \quad (17)$$

where

$$F_{\text{succ}} = Q^{-\frac{1}{2}}, \quad (18)$$

$$F_{\text{fail}} = (I - Q^{-1})^{\frac{1}{2}}. \quad (19)$$

In the case of successful filtering, a new ensemble consists of states which are more distinguishable than the states of the initial ensemble. The inverse filtering map is given by

$$B_{\text{succ}} = Q^{\frac{1}{2}}, \quad (20)$$

$$B_{\text{fail}} = (I - Q)^{\frac{1}{2}}. \quad (21)$$

Using the described soft filtering, Eve can enhance her eavesdropping efficiency when the photon source has a binary side channel as introduced in (2) and (3). If Eve makes the minimum-error measurement of the side-channel degree of freedom, she efficiently introduces bias in quantum states of Alice's ensemble. The initial ensemble (1) transforms to

$$\left\{\frac{q}{2} : |0_x\rangle,\ \frac{1-q}{2} : |1_x\rangle,\ \frac{1-q}{2} : |0_y\rangle,\ \frac{q}{2} : |1_y\rangle\right\} \quad (22)$$

if Eve measures a side channel in the state $|0_\Delta\rangle$ or

$$\left\{\frac{1-q}{2} : |0_x\rangle,\ \frac{q}{2} : |1_x\rangle,\ \frac{q}{2} : |0_y\rangle,\ \frac{1-q}{2} : |1_y\rangle\right\} \quad (23)$$

if Eve measures a side channel in the state $|1_\Delta\rangle$. Here q is the probability to distinguish states $|0_\Delta\rangle$ and $|1_\Delta\rangle$ with a minimum-error measurement [31,32]. Details are given in the Appendix. Then, depending on the side-channel measurement, Eve applies soft filtering to make two states out of four possibly more distinguishable. This filtering allows her to apply a cloning unitary, tuned to the two most distinguishable states, to attack the protocol. We discuss quantum cloning in the next section.

C. Quantum cloning

1. The phase-covariant cloning

Among all attack strategies on the state of the signal photon (without taking side-channel information into account), the most efficient is a collective unitary attack. The essence of this attack is applying a particular unitary evolution to a system, which consists of a signal photon and an ancillary quantum system. The resulting entangled quantum state (after the unitary evolution) allows for maximal information for Eve at a given error on the Bob's side. Eve applies the same unitary operation to all signal states in a sequence of sent bits during Alice and Bob's communication and stores her ancillary system states in a quantum memory. After the basis-exchange step, Eve applies a collective measurement to her quantum memory, which gives her the possible maximum information about distributed secret bits.

A collective unitary attack strategy on the BB84 protocol results in a critical quantum bit error rate of $Q_c \approx 11\%$ and can be implemented with the use of a phase-covariant optimal cloning machine [33]. The optimal cloner is a special unitary transform which takes as input an arbitrary quantum state and produces two quantum states with a high overlap with the initial state. There are many cloner transforms [34] which provide optimal clones for different ensembles of quantum states and for different figures of merit. Although these cloners do not allow violating the no-cloning theorem [3], they can be used as an eavesdropping tool by an adversary. This phase-covariant machine is a unitary of the following form:

$$\begin{aligned} &U|\psi(\phi)\rangle_B|0\rangle_E|0\rangle_{\text{Anc}} \\ &\quad = \tfrac{1}{2}(|0\rangle_B|0\rangle_E|0\rangle_{\text{Anc}} + \cos\eta|0\rangle_B|1\rangle_E|1\rangle_{\text{Anc}} \\ &\qquad + \sin\eta|1\rangle_B|0\rangle_E|1\rangle_{\text{Anc}} \pm \cos\eta|1\rangle_B|0\rangle_E|0\rangle_{\text{Anc}} \\ &\qquad \pm \sin\eta|0\rangle_B|1\rangle_E|0\rangle_{\text{Anc}} \pm |1\rangle_B|1\rangle_E|1\rangle_{\text{Anc}}), \end{aligned} \quad (24)$$

where η is a cloning parameter. When $\eta = 0$, the cloner is the identity operator and has no effect on the compound quantum state, and when $\eta = \pi/2$, Eve has Bob's state in her space, and Bob's qubit becomes maximally mixed because of entanglement with Eve's ancillary qubit.

The described phase-covariant cloning attack allows for a critical bit error value $Q_c \approx 11\%$ for the standard BB84 protocol with two mutually unbiased bases X and Y (or Z and X). Photon distinguishability leads to lower critical error values. Here we consider the case when Eve can run a perfect optimal phase-covariant cloning transform, i.e., without errors from imperfect device construction. Imperfect cloning gives Eve less information than perfect cloning. Thus, allowing Eve to clone perfectly, we give her maximal opportunities, bounded only by the no-cloning theorem.

2. The two-state cloning

While an optimal phase-covariant cloner allows for the most efficient attack in the absence of passive side channels, it does not use photon distinguishability, which arises from side-channel-state measurements. If Eve measures the side-channel state with the minimum-error measurement, the probabilities of states in Alice's ensemble change, and Eve knows some

states are more probable than others. It is reasonable to use this knowledge to devise an adaptive eavesdropping strategy.

Such an eavesdropping strategy can be constructed with the use of a two-state optimal cloner machine [23]. This cloning machine is a unitary of the following form:

$$U|0\rangle_A|0\rangle_E = a|0\rangle_A|0\rangle_E + b(|0\rangle_A|1\rangle_E + |1\rangle_A|0\rangle)_E + c|1\rangle_A|1\rangle_E, \tag{25}$$

$$U|1\rangle_A|0\rangle_E = c|0\rangle_A|0\rangle_E + b(|0\rangle_A|1\rangle_E + |1\rangle_A|0\rangle_E) + a|1\rangle_A|1\rangle_E, \tag{26}$$

where

$$a = \frac{1}{\cos 2x}[(P + Q\cos 2x)\cos x - (P - Q\cos 2x)\sin x], \tag{27}$$

$$b = \frac{1}{\cos 2x}P(\cos x - \sin x)\sin 2x, \tag{28}$$

$$c = \frac{1}{\cos 2x}[(P - Q\cos 2x)\cos x - (P + Q\cos 2x)\sin x], \tag{29}$$

$$P = \frac{1}{2}\frac{\sqrt{1+\sin 2x}}{\sqrt{1+\sin^2 2x}}, \tag{30}$$

$$Q = \frac{1}{2}\frac{\sqrt{1-\sin 2x}}{\sqrt{1-\sin^2 2x}}, \tag{31}$$

and x is given by the scalar product $\langle\psi_1||\psi_2\rangle = \sin 2x$ of states $|\psi_1\rangle$ and $|\psi_2\rangle$, which the cloner is tuned on.

IV. ATTACK STRATEGIES ON THE BB84 PROTOCOL

Here we provide a description of eavesdropping strategies.

A. Minimal-error measurement of the side channel and phase-covariant cloning

The first strategy is minimal-error measurement of the side-channel state and phase-covariant cloning of the signal-photon state. If Eve makes a minimum-error measurement of the side channel, she obtains binary information about the quantum state of the signal photon. This information means that two states within a basis have different probabilities, and Eve now needs to distinguish between states from a quantum ensemble

$$\{p : \rho_0, \quad 1-p : \rho_1\}, \tag{32}$$

where p and $1-p$ represent the classical knowledge obtained through the side channel. Here Eve uses her classical information twofold: she obtains information from a classical channel and also uses this information in quantum ensemble discrimination. The attack can be divided into the following steps:

(1) Eve performs a minimal-error measurement of the side-channel state (Sec. III A 2). She uses information from the side channel in addition to information from a classical channel.

(2) After measuring the side-channel state, she executes a phase-covariant cloning attack (Sec. III C 1) on the transmitted photon and stores her clone in a quantum memory register.

(3) At the end of the communication, Eve obtains basis information for each position in the quantum memory register and makes a collective measurement of the memory register.

Her total information about secret bits is

$$I_{AE} = I_{\text{classical}}(p) + I_{\text{quantum}}(p, Q) = 1 - h_2(p) + h_2\left(\tfrac{1}{2}\{1 - \sqrt{1 - 4p(1-p)[1-(1-2Q)^2]}\}\right), \tag{33}$$

where Q is a bit error on Bob's side and we use the connection between Q and Eve's state distinguishability after the phase-covariant cloner [33,35]. The resulting secret key after this attack strategy is

$$R = 1 - h_2(Q) - I_{AE}. \tag{34}$$

B. Minimal-error measurement of the side-channel state, soft filtering, and two-state cloning

The second strategy is minimal-error measurement of the side channel with soft filtering and two-state cloning of the signal-photon state. Here again, Eve has classical binary information from the side channel, but now she uses this information to adjust the attack strategy on the signal photon. Measurement of the side channel leads to reweighting of the ensemble probabilities (see the Appendix): if Eve measures $|0_\Delta\rangle$ on the side channel, she obtains an ensemble of the form of (A8), and if she measures $|1_\Delta\rangle$, she obtains an ensemble of the form of (A9). Then Eve carries out the eavesdropping sequence soft filtering $\longrightarrow$ cloning $\longrightarrow$ backward soft filtering and obtains a quantum state correlated with a secret bit. At the end of the communication, Eve obtains basis information for every bit position and makes a collective measurement of the ensemble of states in her quantum memory. For an X-basis position, Eve discriminates states from the ensemble

$$\mathcal{E} = \{P_0^{0_\Delta} : \rho_{\text{Eve}}^{0,0_\Delta}, P_0^{1_\Delta} : \rho_{\text{Eve}}^{0,1_\Delta}, P_1^{0_\Delta} : \rho_{\text{Eve}}^{1,0_\Delta}, P_1^{1_\Delta} : \rho_{\text{Eve}}^{1,1_\Delta}, \}, \tag{35}$$

where $P_0^{0_\Delta}$ is a probability to receive a state $\rho_{\text{Eve}}^{0,0_\Delta}$. This state is a result of measuring $|0_\Delta\rangle$ in the side channel, tuning soft filtering on ensemble (22), postselecting successful filtering results, applying two-state cloning, and back soft filtering with postselection of successful results. The same logic applies to another pair of states in Alice's ensemble. To see that this ensemble correctly accounts for side-channel measurement, we here discuss two extreme cases. If side-channel states are orthogonal ($\langle 0_\Delta||1_\Delta\rangle = 0$), then this ensemble transforms to

$$\mathcal{E} = \{\tfrac{1}{2} : |0_x\rangle\langle 0_x|, 0 : \rho_{\text{Eve}}^{0,1_\Delta}, 0 : \rho_{\text{Eve}}^{1,0_\Delta}, \tfrac{1}{2} : |1_x\rangle\langle 1_x|, \}, \tag{36}$$

where states $\rho_{\text{Eve}}^{0,1_\Delta}$ and $\rho_{\text{Eve}}^{1,0_\Delta}$ are some nonorthogonal states (which are not important for calculation since their probability to reach Eve is zero). Here Eve has an ensemble of two equiprobable and orthogonal pure states, which give one bit of information. Hence, Eve can reliably distinguish a quantum state sent by Alice in a quantum channel. In contrast, if side-channel states coincide, then the final ensemble is

$$\mathcal{E} = \{\tfrac{1}{4} : \rho_{\text{Eve}}^{0,0_\Delta}, \tfrac{1}{4} : \rho_{\text{Eve}}^{0,1_\Delta}, \tfrac{1}{4} : \rho_{\text{Eve}}^{1,0_\Delta}, \tfrac{1}{4} : \rho_{\text{Eve}}^{1,1_\Delta}, \}, \tag{37}$$

where $\rho_{\text{Eve}}^{0,1}$ are the resulting states in Eve's quantum memory after the whole attack sequence. The lack of information from the side channel (due to complete indistinguishability of side-channel states) leads to no effect from the soft filtering [(18)

is a unity operator], and the only action Eve performs on the signal-photon state is two-state cloning, which now is chosen at random. The attack can be divided into the following steps:

(1) Eve applies a minimal-error measurement of the side-channel state (Sec. III A 2). This measurement gives her partial information about the transmitted state, which transforms states of the BB84 ensemble (1). If Eve measured the side channel in state $|0\rangle$, the ensemble (1) transforms into (22), and if she measured the side channel in state $|1\rangle$, the ensemble (1) transforms into (23).

(2) This change in probabilities in the ensemble allows Eve to apply a filtering transformation (17), which makes the two states with the highest probabilities more distinguishable while making the other two states less distinguishable.

(3) Then, Eve applies a two-state cloning transformation (Sec. III C 2), which is tuned on a more distinguishable pair of states, to a photon in the quantum channel and produces two clones of this photon.

(4) Then, she applies a backward-filtering transformation to both clones to make them closer to the state sent by Alice.

(5) Eve sends one of two identical clones to Bob and stores the other in her quantum memory.

(6) At the end of the communication, Eve obtains the basis information for each position in the quantum memory register and makes a collective measurement of the memory register.

In the general case, the secret key rate has the form

$$R = (1 - P_{\text{attack}})1 + P_{\text{attack}}[1 - h_2(Q) - \chi(\mathcal{E})], \tag{38}$$

where Q is a bit error on Bob's side. We here introduced the variable P_{attack} to have a controlled parameter for error on Bob's side. Here χ is the Holevo bound value. The Holevo value is a maximal number of bits per state one can extract from the ensemble of quantum states with the best collective measurement of an infinite number of states, defined as follows:

$$\chi(\mathcal{E}) = S\left(\sum_j p_j \rho_j\right) - \sum_j p_j S(\rho_j), \tag{39}$$

where $S(\rho)$ is the von Neumann entropy

$$S(\rho) = -\text{Tr}[\rho \log_2(\rho)]. \tag{40}$$

C. USD measurement of the side-channel state and phase-covariant cloning

The third strategy is a USD measurement of the side-channel state with phase-covariant cloning on the signal-photon state if the USD measurement failed.

(1) Eve uses a USD measurement of the side-channel state (see Sec. III A 1). If the USD measurement is successful, it reliably tells Eve what a side-channel state was. If the USD measurement fails, it provides no information about the side-channel state.

(2) In the case of successful measurement Eve does not attack the signal-photon state and waits until the basis exchange between Alice and Bob. The knowledge of the basis reveals to her what quantum state was sent among the two, corresponding to the measured side-channel state. In this case, she does not introduce any error in the communication act.

(3) In the case of measurement failure, Eve executes a phase-covariant cloning attack (Sec. III C 1) on the signal photon and stores her clone in a quantum memory register.

(4) At the end of the communication, Eve obtains information about used bases for each position in the quantum memory register and makes a collective measurement of the memory register.

Although the side-channel model we use here does not give Eve information about the communication basis, after basis exchange she knows both the basis and bit sent by Alice to Bob. In this way, the attack strategy is similar to the tagged-photon case of Gottesman *et al.* [36] applied to a USD attack of the side-channel degree of freedom (see also [18]):

$$R = (1 - P_{\text{USD}})\left(1 - h_2\left(\frac{Q}{1 - P_{\text{USD}}}\right)\right) - \chi(\mathcal{E}), \tag{41}$$

where P_{USD} is the probability of the USD measurement success, Q is a bit error on Bob's side, and $\mathcal{E} = \{\frac{1}{2} : \rho_{0,X}, \frac{1}{2} : \rho_{1,X}\}$ is the standard BB84 basis X after phase-covariant cloning eavesdropping.

D. Results

Here we provide the calculation results for the three eavesdropping strategies. In Fig. 1, we provide critical error rates (error rates for secret key rate $R = 0$) for different values of the HOM visibility. In Fig. 2, we provide secret key rates for three eavesdropping strategies for different values of single-photon HOM visibility.

In Fig. 1(a) we show how the critical error rate (i.e., maximum error rate that corresponds to the zero secret key rate) depends on the light-source HOM visibility (i.e., the "quality" of the light source). As we expect, the critical error rate becomes less than 11% in the presence of side channels. Among the three attacks considered here, one of the attacks (explained in Sec. IV B) provides a critical error rate greater than 11% in the absence of side channels (HOM visibility equals 1). This attack is based on a unitary two-state cloning. This cloning provides optimal clones of only two states, and thus, it is suboptimal for the symmetric alphabet of the BB84 protocol (four equiprobable states on the equator of the Bloch sphere). Although this attack provides a higher value of the critical error rate in the absence of side channels (soft filtering with two-state cloning), it becomes the most efficient attack when the source quality is very low. The reason is that information from side channels introduces asymmetry in the BB84 alphabet and makes two states more likely than the other two. This bias makes the two-state cloning preferable to use for eavesdropping. We see that at $V \lesssim 0.4$ soft filtering with two-state cloning becomes more efficient than phase-covariant cloning without soft filtering.

For high-quality light sources (i.e., high values of HOM visibility) the situation is different. As we can see in Fig. 1(a), two attacks (attacks in Secs. IV A and IV C) become the most efficient when $V \gtrsim 0.4$ and provide almost identical results in terms of the critical-error-rate reduction. Figure 1(b) indicates that for realistic single-photon sources that provide a HOM visibility value above 0.9, the critical error rate decreases to just 10%. Hence, a small deviation of the light source from the ideal model does not result in any significant reduction of

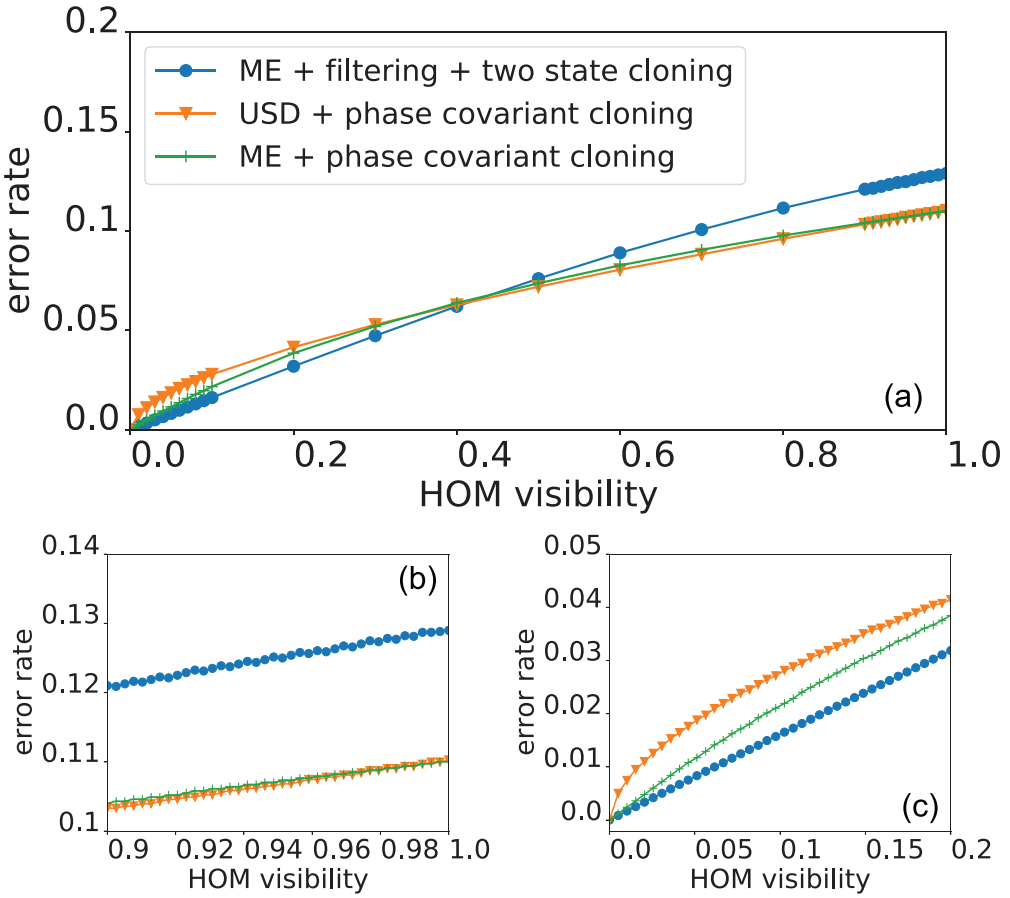

FIG. 1. (a) Critical error rates (secret key rate is zero) for the three eavesdropping strategies. Here Eve attacks the side-channel state with a minimum error measurement (ME) or an unambiguous state discrimination (USD). The attack on the side-channel state is followed by an attack on the signal photon. (b) Critical error rates for high visibility values. (c) Critical error rates for low visibility values.

security, at least for the considered eavesdropping strategies. In Fig. 1(c) we see that all three attacks provide zero key in the limit of zero HOM visibility, which indicates our calculations are self-consistent. We see that the USD-based attack gives strongly nonlinear dependence of the critical error rate on the visibility value (although single-photon sources of this quality are hardly applicable to QKD) and gives Eve less information than the other two attacks. As a result, the two-state cloning attack with soft filtering (see Sec. IV B) is the most efficient for low-quality light sources.

In Fig. 2 we see how the secret key rate of the BB84 protocol depends on the quantum bit error rate for different values of the HOM visibility of the light source. Apart from the trivial fact that the higher the quantum bit error rate is, the lower the secret key rate is, the dependence of the key rate on the visibility value is not that trivial. In the limit of zero error rate, eavesdropping strategies behave differently. In the attacks in Secs. IV A and IV C Eve uses the side channels separately from the operational degree of freedom. In particular, she can measure the side channel and does not attack the operational signal state. In this way she obtains some information about the secret key without introducing bit errors, and the secret key becomes less than 1 bit even when the error rate is zero [Figs. 2(a) and 2(b)]. The adaptive strategy [Fig. 2(c)], which uses filtering and two-state cloning, requires Eve to act with the operational degree of freedom in order to obtain information about secret bits. She does not use the side channel and operational degree of freedom separately and thus cannot obtain any information without introducing errors.

V. CONCLUSION AND OUTLOOK

We compared the efficiencies of three explicit eavesdropping strategies for the BB84 protocol with partially distinguishable photons. We demonstrated that among these attacks the USD measurement of the side channel along with a unitary attack on the signal-photon state is the most efficient eavesdropping for light sources with low information leakage (high values of the HOM visibility). In contrast, for light sources with high information leakage (low values of the HOM visibility), it is better to use adaptive eavesdropping with soft filtering of the signal-photon state.

We also found that eavesdropping with postselection (soft filtering in our studies) can enhance eavesdropping efficiency compared to eavesdropping without postselection. This enhancement is remarkable since this effect does not take place in the case of BB84 without side channels.

Our results open the following questions for future research. We demonstrated that it is possible to overcome the performance of the optimal attack without side channels, but we do not claim the global optimality of the proposed attacks.

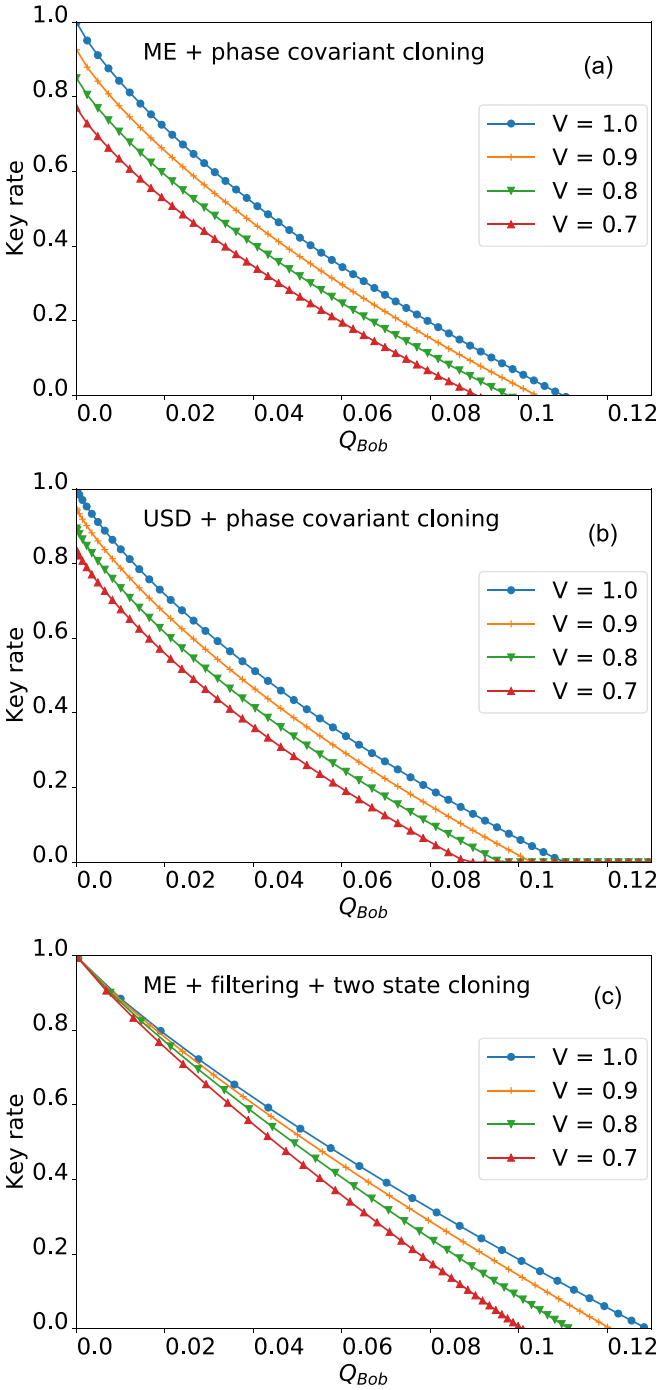

FIG. 2. Secret key rate for the three eavesdropping strategies: (a) corresponds to Sec. IV A, (b) corresponds to Sec. IV C, and (c) corresponds to Sec. IV B. Here we provide the key rate as a function of the quantum bit error rate for several values of the light-source HOM visibility values.

The optimal eavesdropping attack in the presence of side channels and the optimal model of the side channels remain an open problem. In particular, the dimensionality of the side-channel-state space potentially influences the efficiency of the eavesdropping; using different photon parameters separately (e.g., spectral and spatial profiles), Eve can construct more complicated sequences of measurements and filtering over both signal and side-channel degrees of freedom, which leads to the lower bound on the secret key rate.

Next, the critical error rate is not the only possible framework since there are attacks which compromise the protocol with a zero bit error rate. For example, USD of the side-channel state with blocking-inconclusive results leads to zero error on Bob's side. To take such strategies into account, we need to include the channel transmittance to the security analysis and connect it to the probability of conclusive results or other postselection-based strategies.

ACKNOWLEDGMENT

This work was funded by the Ministry of Science and Higher Education of the Russian Federation (Grant No. 075-15-2020-788).

APPENDIX: ENSEMBLE REWEIGHTING FROM MINIMUM-ERROR MEASUREMENT OF THE SIDE-CHANNEL STATE

Let us consider the process of minimal-error measurement of the side-channel degree of freedom in detail. We denote two possible states of the side-channel degree of freedom as $\{|0_\Delta\rangle, |1_\Delta\rangle\}$. We model the process of measurement as the interaction of a side-channel degree of freedom with an ancillary degree of freedom, whose states can be distinguished with certainty ($\{|r_0\rangle, |r_1\rangle\} : \langle r_0||r_1\rangle = 0$). We specify the Stinespring representation of this process as follows:

$$|0_\Delta\rangle \longrightarrow \sqrt{P_{\text{succ}}}|0_\Delta\rangle|r\rangle_0 + \sqrt{1-P_{\text{succ}}}|1_\Delta\rangle|r\rangle_1, \quad \text{(A1)}$$

$$|1_\Delta\rangle \longrightarrow \sqrt{1-P_{\text{succ}}}|0_\Delta\rangle|r\rangle_0 + \sqrt{P_{\text{succ}}}|1_\Delta\rangle|r\rangle_1, \quad \text{(A2)}$$

where P_{succ} is the probability to distinguish between two states $|0_\Delta\rangle$ and $|1_\Delta\rangle$. Then, the density matrices of the two resulting states of the compound system are

$$\begin{aligned}\rho_0^{\text{minerr}} = {} & P_{\text{succ}}|0_\Delta\rangle\langle 0_\Delta| \otimes |r_0\rangle\langle r_0| \\ & + \sqrt{P_{\text{succ}}(1-P_{\text{succ}})}|0_\Delta\rangle\langle 1_\Delta| \otimes |r_0\rangle\langle r_1| \\ & + \sqrt{P_{\text{succ}}(1-P_{\text{succ}})}|1_\Delta\rangle\langle 0_\Delta| \otimes |r_1\rangle\langle r_0| \\ & + (1-P_{\text{succ}})|1_\Delta\rangle\langle 1_\Delta| \otimes |r_1\rangle\langle r_1|, \end{aligned} \quad \text{(A3)}$$

$$\begin{aligned}\rho_1^{\text{minerr}} = {} & (1-P_{\text{succ}})|0_\Delta\rangle\langle 0_\Delta| \otimes |r_0\rangle\langle r_0| \\ & + \sqrt{P_{\text{succ}}(1-P_{\text{succ}})}|0_\Delta\rangle\langle 1_\Delta| \otimes |r_0\rangle\langle r_1| \\ & + \sqrt{P_{\text{succ}}(1-P_{\text{succ}})}|1_\Delta\rangle\langle 0_\Delta| \otimes |r_1\rangle\langle r_0| \\ & + P_{\text{succ}}|1_\Delta\rangle\langle 1_\Delta| \otimes |r_1\rangle\langle r_1|. \end{aligned} \quad \text{(A4)}$$

Now let us look at changes in ensembles of Alice's states. Suppose Alice chose an X basis state to send a secret bit to Bob. Eve applies unitary evolution to a side-channel state and a measurement-device state:

$$\rho_{\text{Alice}}^X = \tfrac{1}{2}|0_x\rangle\langle 0_x| \otimes \rho_0^{\text{minerr}} + \tfrac{1}{2}|1_x\rangle\langle 1_x| \otimes \rho_1^{\text{minerr}}. \quad \text{(A5)}$$

We can rewrite this state in a more convenient form as

$$\rho^{X}_{\text{Alice}} = \left(\frac{P_{\text{succ}}}{2}|0_x\rangle\langle 0_x| + \frac{1-P_{\text{succ}}}{2}|1_x\rangle\langle 1_x|\right) \times \otimes |0_\Delta\rangle\langle 0_\Delta| \otimes |r_0\rangle\langle r_0| \tag{A6}$$

$$+\left(\frac{1-P_{\text{succ}}}{2}|0_x\rangle\langle 0_x| + \frac{P_{\text{succ}}}{2}|1_x\rangle\langle 1_x|\right) \times \otimes |1_\Delta\rangle\langle 1_\Delta| \otimes |r_1\rangle\langle r_1| + \text{off-diagonal terms}. \tag{A7}$$

After doing a measurement of the ancillary state the off-diagonal terms of a compound density matrix vanish. Depending on the outcome of the ancilla measurement, the resulting states have the form

$$\text{Measured } |r_0\rangle : \rho'_{\text{Alice}} = P_{\text{succ}}|0_x\rangle\langle 0_x| + (1-P_{\text{succ}})|1_x\rangle\langle 1_x|, \tag{A8}$$

$$\text{Measured } |r_1\rangle : \rho'_{\text{Alice}} = (1-P_{\text{succ}})|0_x\rangle\langle 0_x| + P_{\text{succ}}|1_x\rangle\langle 1_x|. \tag{A9}$$

The same logic applies to all bases of the BB84 protocol. This leads to reweighting of Alice's ensemble states, which is used for adaptive eavesdropping.

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

[2] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. S. Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden, Adv. Opt. Photonics **12**, 1012 (2020).

[3] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982).

[4] D. Dieks, Phys. Lett. A **92**, 271 (1982).

[5] H.-K. Lo and H. F. Chau, Science **283**, 2050 (1999).

[6] P. W. Shor and J. Preskill, Phys. Rev. Lett. **85**, 441 (2000).

[7] V. Scarani and C. Kurtsiefer, Theor. Comput. Sci. **560**, 27 (2014).

[8] N. Jain, B. Stiller, I. Khan, D. Elser, C. Marquardt, and G. Leuchs, Contemp. Phys. **57**, 366 (2016).

[9] Ø. Marøy, V. Makarov, and J. Skaar, Quantum Sci. Technol. **2**, 044013 (2017).

[10] A. Huang, S.-H. Sun, Z. Liu, and V. Makarov, Phys. Rev. A **98**, 012330 (2018).

[11] E. Diamanti, H.-K. Lo, B. Qi, and Z. Yuan, npj Quantum Inf. **2**, 16025 (2016).

[12] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, Rev. Mod. Phys. **92**, 025002 (2020).

[13] H.-K. Lo, M. Curty, and B. Qi, Phys. Rev. Lett. **108**, 130503 (2012).

[14] G.-J. Fan-Yuan, F.-Y. Lu, S. Wang, Z.-Q. Yin, D.-Y. He, Z. Zhou, J. Teng, W. Chen, G.-C. Guo, and Z.-F. Han, Photonics Res. **9**, 1881 (2021).

[15] G.-J. Fan-Yuan, F.-Y. Lu, S. Wang, Z.-Q. Yin, D.-Y. He, W. Chen, Z. Zhou, Z.-H. Wang, J. Teng, G.-C. Guo, and Z.-F. Han, Optica **9**, 812 (2022).

[16] M. Lucamarini, Z. L. Yuan, J. F. Dynes, and A. J. Shields, Nature (London) **557**, 400 (2018).

[17] S. Wang, Z.-Q. Yin, D.-Y. He, W. Chen, R.-Q. Wang, P. Ye, Y. Zhou, G.-J. Fan-Yuan, F.-X. Wang, W. Chen, Y.-G. Zhu, P. V. Morozov, A. V. Divochiy, Z. Zhou, G.-C. Guo, and Z.-F. Han, Nat. Photonics **16**, 154 (2022).

[18] M. Lucamarini, I. Choi, M. B. Ward, J. F. Dynes, Z. L. Yuan, and A. J. Shields, Phys. Rev. X **5**, 031030 (2015).

[19] K. Tamaki, M. Curty, and M. Lucamarini, New J. Phys. **18**, 065008 (2016).

[20] S. Nauerth, M. Fürst, T. Schmitt-Manderbach, H. Weier, and H. Weinfurter, New J. Phys. **11**, 065001 (2009).

[21] A. Duplinskiy and D. Sych, Phys. Rev. A **104**, 012601 (2021).

[22] M. Pereira, M. Curty, and K. Tamaki, npj Quantum Inf. **5**, 62 (2019).

[23] D. Bruß, M. Cinchetti, G. M. D'Ariano, and C. Macchiavello, Phys. Rev. A **62**, 012302 (2000).

[24] D. Babukhin and D. Sych, J. Phys.: Conf. Ser. **1695**, 012119 (2020).

[25] D. Babukhin and D. Sych, J. Phys.: Conf. Ser. **1984**, 012008 (2021).

[26] D. Sych, A. Duplinskiy, and D. Babukhin, J. Phys.: Conf. Ser. **1984**, 012001 (2021).

[27] A. M. Brańczyk, Hong-Ou-Mandel interference, arXiv1711.00080v1 [quant-ph] (2017).

[28] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[29] D. A. Kronberg, A. S. Nikolaeva, Y. V. Kurochkin, and A. K. Fedorov, Phys. Rev. A **101**, 032334 (2020).

[30] D. Kronberg, Proc. Steklov Inst. Math. **313**, 113 (2021).

[31] C. Helstrom, J. Stat. Phys. **1**, 231 (1969).

[32] A. S. Kholevo, Theory Probab. Appl. **23**, 411 (1979).

[33] D. Bruß, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, and J. A. Smolin, Phys. Rev. A **57**, 2368 (1998).

[34] N. Gisin and S. Massar, Phys. Rev. Lett. **79**, 2153 (1997).

[35] C. A. Fuchs, N. Gisin, R. B. Griffiths, C.-S. Niu, and A. Peres, Phys. Rev. A **56**, 1163 (1997).

[36] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, Quant. Inf. Comput. **5**, 325 (2004).