Quantum-key-distribution protocol with pseudorandom bases

A. S. Trushechkin,[1,2,3,4] P. A. Tregubov,[2] E. O. Kiktenko,[1,4] Y. V. Kurochkin,[4,3] and A. K. Fedorov[4,3]

[1] *Steklov Mathematical Institute of Russian Academy of Sciences, Moscow 119991, Russia*
[2] *National Research Nuclear University "MEPhI," Moscow 115409, Russia*
[3] *Department of Mathematics and Russian Quantum Center, National University of Science and Technology MISiS, Moscow 119049, Russia*
[4] *Russian Quantum Center, Skolkovo, Moscow 143025, Russia*

(Received 27 June 2017; published 16 January 2018)

Quantum key distribution (QKD) offers a way for establishing information-theoretical secure communications. An important part of QKD technology is a high-quality random number generator for the quantum-state preparation and for post-processing procedures. In this work, we consider a class of prepare-and-measure QKD protocols, utilizing additional pseudorandomness in the preparation of quantum states. We study one of such protocols and analyze its security against the intercept-resend attack. We demonstrate that, for single-photon sources, the considered protocol gives better secret key rates than the BB84 and the asymmetric BB84 protocols. However, the protocol strongly requires single-photon sources.

DOI: 10.1103/PhysRevA.97.012311

I. INTRODUCTION

Quantum algorithms exploit the laws of quantum mechanics to solve problems exponentially faster than their best classical counterparts [1]. Shor's quantum algorithm for fast number factoring attracted a great attention since this problem is in the heart of public-key cryptosystems [2]. In view of Shor's algorithm, the only way to ensure the absolute long-term security is to use information-theoretical secure primitives, such as the one-time pad scheme [3–5]. However, the need for establishing secret keys between communicating parties invites the challenge of how to securely distribute these keys [5].

Fortunately, together with the tool for breaking public-key cryptographic primitives, quantum physics allows one to establish secure communications [6]. By encoding information in quantum states of photons, transmitting them through fiber channels, and communication via authenticated classical channel, quantum-key-distribution (QKD) systems offer a practical tool for private key distribution. Unlike classical cryptography, QKD promises information-theoretical security based on the quantum physics laws. During the last decades, great progress in theory, experimental study, and technology of QKD has been performed. However, QKD technology faces a number of challenges such as distance, key generation rate, practical security, and many others [7].

The idea behind the seminal proposal for QKD protocol, known as BB84 protocol [8], is inspired by the conjugate coding method [9]. The BB84 protocol employs the idea of usage of two orthogonal polarizations states of photons. The BB84 protocol has been widely studied, and its security has been proven [6]. Development of novel QKD protocols, offering ways to push the performance of QKD technology, is on the forefront of quantum information technologies. During the last decades, several extensions of the BB84 protocol and alternative QKD protocols, such as E91 (proposed independently of BB84) [10], B92 [11], six-state BB84 protocol [12], asymmetric BB84 (we will abbreviate it as aBB84) [13], SARG04 [14], differential-phase shift [15–17], coherent one way [18], and also setups with continuous variables [19], have been actively discussed.

The point we want to stress here is the fact that for the seminal BB84 protocol and most of its variations, something should provide the ignorance of an eavesdropper (Eve) about the bases in which quantum states are encoded [8]. The BB84 protocol provides this condition by the random independent choice of the bases by legitimate parties (Alice and Bob). To this end, Alice and Bob use true random number generators (TRNG). However, the cost is the sifting procedure: Alice and Bob must discard the positions with incompatible basis choices. This leads to a loss of approximately a half of the raw key. In order to reduce the losses in the sifting procedure, the aBB84 protocol has been proposed [13]. In this protocol, Alice and Bob use one basis with a high probability and a conjugate basis with a small probability. The first basis is used mainly to establish a secret key, while the second one is used to verify the absence of eavesdropping. We will refer to these bases as "the signal basis" and "the test basis," respectively. In the asymptotic case of an infinitely large number of transmitted quantum states, the probability of the use of the test basis can be made arbitrarily small. Hence, the basis choices of Alice and Bob almost always coincide and there is almost no sifting. Nevertheless, for a finite number of transmitted states, this probability cannot be made arbitrarily small since a reliable statistics for the test basis should be collected for tight estimation of the amount of eavesdropping [13,20].

In this work, we consider a class of QKD protocols, which utilize the pseudorandomness in the preparation of quantum states. Namely, Alice and Bob can use not random but pseudorandom sequence of bases generated from a common short secret key. On the one hand, their bases always coincide, so, the suggested scheme allows one to avoid the sifting procedure. On the other hand, for Eve, who does not know this key, the sequence is similar to a random one and she cannot predict it. On the basis of this idea, we study a protocol with pseudorandom choice of bases (abbreviated as PRB) and analyze its security against the intercept-resend attack. The suggested protocol is a formalization of a protocol (the floating basis protocol) described by one of the authors [21–23]. In this work, we assume that the sequence of logical bits is truly random, but the sequence of bases is pseudorandom. We then demonstrate that the PRB protocol gives higher secret key rates than the BB84 protocol and approximately the same and even slightly better rates as the asymmetric BB84 protocol. However, the PRB strongly requires single-photon sources.

A general motivation for the development of new QKD protocols is exploration of different ways of how we can exploit the properties of quantum information to provide information security. The idea of the protocol proposed here is a method of combining of classical pseudorandomness with quantum encoding of information. We should note that the known Y00 protocol [24–26] also uses pseudorandom quantum states. It provides a randomized stream cipher with information-theoretic security by a randomization based on quantum noise and additional tools. Another important example of utilizing pseudorandomness is a recently suggested mechanism for quantum data locking [27,28].

We can treat QKD protocols utilizing pseudorandomness in such a way. Since generators of true random numbers are not sufficiently fast, pseudorandom number generators (PRNGs) are used in practical setups instead [29]. It is interesting to study how the use of pseudorandom numbers instead of true random numbers affects the security of QKD protocols (see Ref. [30]) and, moreover, can it be even advantageous. Here, we assume that the sequence of logical bits is truly random, but the sequence of bases is pseudorandom. As it was explained above, if we are able to prove the security of such scheme, we can make it more advantageous due to avoiding the sifting (with the cost of an additional secret key consumption for initial secret random seed for the PRNG in future sessions).

The paper is organized as follows. The new QKD protocol, which we will refer to as the PRB protocol, is described in Sec. II. Its security against the intercept-resend attack is proved in Sec. III. We summarize the main results of our work in Sec. V. In Sec. IV, we analyze the photon-number splitting (PNS) attack and show that, unlike the BB84 protocol and its modifications, the PRB protocol strongly requires a single-photon source of light.

II. QKD PROTOCOL WITH PSEUDORANDOM BASES

Let Alice and Bob have a common preshared key

$$k = (k_1, \dots, k_l) \in \{0,1\}^l, \tag{1}$$

which is the seed for the pseudorandom number generator (PRNG), l is the key size. We use the following notation:

$$|\varphi\rangle = \cos\varphi|0\rangle + \sin\varphi|1\rangle, \tag{2}$$

where $\{|0\rangle, |1\rangle\}$ is the standard basis. As usual in QKD, Alice and Bob have a quantum channel and an authenticated public classical channel: Eve freely read the communication over this channel, but cannot interfere in it.

The considered class of QKD protocol based on pseudorandomness operates as follows.

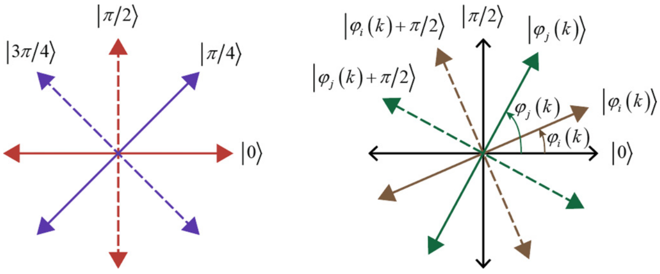

FIG. 1. Base patterns on the Poincaré sphere. The BB84 protocol (left) uses two maximally conjugated bases with the angle $\pi/4$ between each other. For each pulse, the bases are chosen by Alice and Bob randomly and independently, so, the sifting procedure (discarding of positions where Alice's and Bob's bases are different) is required. In the suggested protocol (right), the standard basis $\{|0\rangle, |1\rangle\}$ is rotated by an arbitrary angle (from a finite set) in not a random pseudorandom manner. Thus, the bases of Alice and Bob always coincide and there is no sifting.

(i) Using the common preshared key k and the PRNG, Alice and Bob generate a common pseudorandom sequence in the following form:

$$\varphi_1(k), \dots, \varphi_N(k), \quad \varphi_i(k) \in \left\{\frac{\pi j}{2M}\right\}_{j=0}^{M-1}, \tag{3}$$

where $M = 2^m$ for some $m \geqslant 1$. Schematically, such pseudorandom rotations of the standard basis $\{|0\rangle, |1\rangle\}$ are shown in Fig. 1. We assume that l is divisible by m and denote $l/m = l'$. We assume that $N = 2^{l'}$.

(ii) Using a TRNG, Alice generates the random bits as follows:

$$x_1, \dots, x_N. \tag{4}$$

(iii) Using sequence (3) and generated random bits (4), Alice prepares the following sequence of states:

$$|\varphi_1(k) + x_1\pi/2\rangle, \dots, |\varphi_N(k) + x_N\pi/2\rangle, \tag{5}$$

and sends them to Bob over the quantum channel.

(iv) Bob measures them using the following bases:

$$\{|\varphi_i(k)\rangle, |\varphi_i(k) + \pi/2\rangle\}, \quad i = 1, \dots, n. \tag{6}$$

(v) Bob then writes the results of these measurements in the binary variables y_i as follows: $|\varphi_i(k)\rangle$ corresponds to $y_i = 0$ and $|\varphi_i(k) + \pi/2\rangle$ corresponds to $y_i = 1$. In the case of ideal channel and no eavesdropping, $x_i = y_i$ for all i, hence, Alice and Bob can use their binary strings x and y as a common secret key. Due to noise in the channel and, probably, eavesdropping, there are some errors in these strings, and Eve potentially has some information about them. They are, thus, called the *raw keys*.

(vi) The following steps of post-processing of raw keys coincide with those of BB84, so, we only briefly mention them: Alice and Bob perform the error correction (using error-correcting codes or interactive error-correction protocols; for the last issues concerning the adaptation of error-correcting codes for QKD, see Ref. [31]) and calculate the number of detected errors [32]. If the number of errors exceeds a certain threshold, which make the secret key distribution impossible, Alice and Bob abort the protocol. Otherwise, they perform privacy amplification to reduce the potential Eve's information to a negligible level. The resulting key is called the *secret key* or the *final key*. It is the output of our QKD protocol.

The underlying idea of the protocol is as follows. If Eve does not know the initial secret key k, and the pseudorandom angles φ_i are similar to truly random, then she cannot guess all the bases correctly and her eavesdropping will cause disturbance in Alice's and Bob's raw keys. Of course, a rigorous analysis is required since it takes into account that the sequence $\{\varphi_i(k)\}_{i=1}^N$ is not truly random but pseudorandom.

Remark 1: We note that every quantum-key-distillation protocol that makes use of preshared key can be transformed into an equally efficient protocol which needs no preshared key [34]. However, this fact is irrelevant for our case since we use the preshared key not on the stage of secret key distillation from raw keys, but on the previous stage of transfer of quantum states.

Remark 2: If $M = 2$, then the protocol uses two BB84 bases. In this case, the protocol can be called "BB84 with pseudorandom sequence of bases," while the case $M > 2$ can be called the "multibasis protocol." We will see that the multibasis version gives higher secret key rates due to additional uncertainty for Eve.

The possibility of the use of arbitrary number of bases is a consequence of their correlated choice by Alice and Bob; otherwise, this would lead to an increase of the number of positions with inconsistent bases.

Remark 3: We may ask why do Alice and Bob need QKD if they can use a pseudorandom sequence as a "running key" for encryption, e.g., as a key in the one-time pad. This construction is known in classical cryptography as *stream cipher* [5]. However, the stream cipher cannot provide the information-theoretic security. Moreover, it is well known that the information-theoretic security is never possible in classical cryptography whenever the entropy of the initial secret key shared by Alice and Bob is smaller than the entropy of the message to be encrypted. The use of QKD makes the information-theoretic security possible. Using QKD, we obtain a long key whose entropy is close to maximal. This key can be used, e.g., for one-time pad encryption to provide information-theoretic security.

Also, it is worthwhile to stress that despite of the use of a PRNG, the resulting key is truly random since the bit values in our QKD protocol are still generated by a TRNG, while a PRNG is used only for bases choices.

Remark 4: Our protocol can be regarded as a generalization of BB84 without public announcement of bases proposed in Ref. [35]. In this protocol, Alice and Bob also share an initial short secret key $k = (k_1, \dots, k_l) \in \{0,1\}^l$, which determines first l choices of their bases. Then, Alice and Bob repeat this sequence of bases an arbitrary number of times. This corresponds to our protocol with $M = 2$ and $\varphi_i(k) = \pi k_{(i-1 \bmod l)+1}/4$. The advantage of our protocol is that the use of a good PRNG increases the security of the protocol. For example, we use a PRNG (see the next subsection) which generates sequences with the period of order 2^l instead of l. Also, the security analysis in Ref. [35] is valid only for the asymptotic case $l \to \infty$. For example, both protocols are certainly insecure for $l = 1$ (Eve correctly guesses the initial key and, hence, all the bases with the probability $\frac{1}{2}$). Thus, the security of these protocols depends on l, but an analysis of this dependence is lacking in Ref. [35]. Here, we provide such analysis for the simplest intercept-resend attack. This is the most complicated part of our analysis given in Sec. III and Appendix C.

PRNG based on the Legendre symbol

Our choice for the PRNG is the Legendre symbol PRNG since it provides an almost uniform distribution of patterns, which will be exploited in the security proof. The PRNG is defined as follows.

Let L be a prime number (public value) such that $L \equiv 3 \pmod 4$, and $k \in [0, L-1]$ be a secret key. Let us then define

$$\overline{a}_i = \begin{cases} 1, & \text{if } i \text{ is a quadratic residue modulo } L \\ & \text{and } i \not\equiv 0 \pmod{L}; \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

$$a_i(k) = \overline{a}_{k+i}. \tag{8}$$

Recall that x is called a quadratic residue modulo L if there exists an integer y such that $y^2 \equiv x \pmod{L}$. If $i \not\equiv 0 \pmod{L}$, the value $2\overline{a}_i - 1$ is called the Legendre symbol of i. We will refer to the sequence

$$\overline{a}_1, \overline{a}_2, \dots \tag{9}$$

as the Legendre sequence. It is periodic with the period L. For example, one period of the sequence $\{\overline{a}_i\}_{i=1}^{\infty}$ for $L = 7$ is 1101000.

Pseudorandom properties of Legendre sequences are known for a long time [36]. In particular, the distribution of patterns of Legendre sequences is known to be close to uniform [36–40] [for details, see property (A4) from Appendix A and, as its direct consequence, property (C14) from Appendix C]. This will be important for estimation of the number of bases correctly guessed by Eve.

Remark 5: It is worthwhile to stress that the usual sense of cryptographic security for PRNG is neither a necessary nor sufficient condition for us because we have a completely different context than in classical cryptography. In classical cryptography, PRNG can be secure only in the *computational* sense. The usual notion of cryptographically secure PRNG means that the eavesdropper has to do an unrealistic amount of computation to break the PRNG.

Here, as is common for QKD, we assume that Eve has unlimited computing power and consider the information-theoretic security. Under such an assumption, Eve can break any PRNG whenever she observes a pseudorandom sequence with the length greater than the length l of the initial key. But, fortunately, in our protocol Eve does not observe the pseudorandom sequence itself. Moreover, we will see that on the stage of quantum-state transmission she gets no knowledge on the pseudorandom sequence at all. Thus, we have a completely different context than in classical cryptography.

For this reason, we do not require our PRNG to be secure in the usual (computational) sense. Property (C14) from Appendix C with a suitable (not large) function $W(s)$ is the only property we need for the proof of the security of our protocol against the intercept-resend attack. This property means that the so-called pattern distribution for the PRNG is close to uniform.

Let us specify the use of this PRNG for our protocol. In the two-basis version of the protocol (BB84 with pseudorandom sequence of bases), the basis for each position i is specified by a single bit $a_i(k)$. The length of the key l is the number of bits required to specify k, i.e., $\lceil \log_2 L \rceil$. In the following, $\lceil x \rceil$ and $\lfloor x \rfloor$ denote the ceiling and the floor of x (the closest integer to x from above and from below), respectively.

In the multibasis version of the protocol, every basis is specified by m bits or m registers. Each register has its own PRNG based on the Legendre symbol, so that

$$\varphi_i(k) = \frac{\pi}{2} \sum_{j=1}^{m} a_i(k^{(j)}) 2^{-j}, \tag{10}$$

where $k^{(j)}$ is a subkey (of length $l' = \lceil \log_2 L \rceil$) for the jth register and the sequence $(a_1(k^{(j)}), a_2(k^{(j)}), \ldots)$ is specified by Eq. (8). The total key $(k^{(1)}, \ldots, k^{(m)})$ has the length $l = l'm$.

III. INTERCEPT-RESEND ATTACK

The simplest attack on BB84-like protocols is the intercept-resend attack. Here, we describe this attack for the considered class of QKD protocols:

(i) Eve chooses some positions $1 \leqslant i_1 < \cdots < i_n \leqslant N$ to intercept, where $0 < n \leqslant N$. Denote $\gamma = n/N$ the fraction of positions she intercepts. Then, for each $j = 1, \ldots, n$, Eve performs the next steps.

(ii) Eve chooses an angle β_{i_j}, measures the i_jth qubit in the basis

$$\left\{ |\beta_{i_j}\rangle, \left|\beta_{i_j} + \frac{\pi}{2}\right\rangle \right\}, \tag{11}$$

and writes the result in the variable z_{i_j} (0 or 1, respectively).

(iii) Eve sends a new qubit in the state $|\beta_{i_j} + z_{i_j}\pi/2\rangle$ to Bob.

The crucial point is that the results of Eve's measurements alone leak no information about the bases and, hence, about the initial secret key (the seed for the PRNG) k. This follows from the fact that the quantum state of a qubit for unknown x is independent on φ:

$$\frac{1}{2}|\varphi\rangle \left|\varphi + \frac{1}{2}\right\rangle \left|\varphi + \frac{\pi}{2}\right\rangle \left\langle \varphi + \frac{\pi}{2}\right| = \frac{1}{2} I, \tag{12}$$

where I is the identity operator.

Thus, on the stage of quantum-state transmission, Eve chooses the angle β_{i_j} with no information on the key and has to *guess* the bases or the initial key. Since it is unlikely that she correctly guesses all bases, we arrive at the keystone of the security of QKD: eavesdropping causes disturbance. Rigorous estimations of the number of bases that Eve can correctly guess is the main part of security proof.

From the other side, we assume that, after the accomplishment of all stages of the protocol and, moreover, after the transmission of the encrypted message, Eve is able to determine the initial key. Thus, *a posteriori*, she gets knowledge of the correct bases. In Appendix B we show that Eve needs of order l bits to intercept to guess the initial key if she knows the encrypted message ("known plain-text attack").

In our analysis, we assume that N and n are so large that we can neglect the statistical fluctuations since our aim is to give general analysis of the protocol, not the ultimate formulas for the practical applications.

A. BB84 with pseudorandom sequence of bases

For a transparent analysis, we first consider the protocol BB84 with pseudorandom bases. In this case, the basis choice is specified by a single bit a_i. Let the upper bound on the number of bases correctly guessed by Eve for a given γ be $n_{\text{correct}}(\gamma)$. Respectively, $n_{\text{incorrect}}(\gamma) = n - n_{\text{correct}}(\gamma)$ is the lower bound on the number of incorrect guesses. Recall that we assume that Eve eventually gets knowledge of the initial key k, hence, for each position, she knows whether she has correctly guessed the basis in this position or not. Consequently, the quantum bit error rate (QBER) (q) and the Eve's mean information on a raw key bit are as follows:

$$q(\gamma) = \frac{1}{2}\frac{n_{\text{incorrect}}(\gamma)}{N}, \tag{13}$$

$$I_{\text{E}}(\gamma) = \frac{n_{\text{correct}}(\gamma)}{N}. \tag{14}$$

The legitimate parties have the measured (or estimated) value of QBER. If we replace the left-hand side of Eq. (13) by this value, we can find the inverse function as follows:

$$\gamma = \gamma(q). \tag{15}$$

This is an estimation of the fraction of qubits intercepted by Eve for a given QBER. Then, one has

$$I_{\text{E}}(q) = \frac{n_{\text{correct}}(\gamma(q))}{N}. \tag{16}$$

From the other side, Bob's mean information on a bit of the Alice's raw key is

$$I_{\text{B}}(q) = 1 - h(q). \tag{17}$$

Here,

$$h(p) = -p \log_2 p - (1-p) \log_2 (1-p) \tag{18}$$

is the binary entropy function, $0 \leqslant h(p) \leqslant 1$. However, to fully exploit this information, Alice and Bob require error-correcting scheme that achieves the theoretical (Shannon) limit, in which $h(q)$ bits of information about raw keys are revealed over the public channel. Practically, $f(q)h(q)$ bits are revealed, where $f(q) \geqslant 1$ is the efficiency of the error-correction scheme. Thus, the "effective" Bob's mean information on a bit of the Alice's raw key is

$$I_{\text{B}}(q) = 1 - f(q)h(q). \tag{19}$$

Then, the secret key rate (per transmitted qubit, also called secret fraction) has the following form [6,41]:

$$R(q) = I_{\text{B}}(q) - I_{\text{E}}(q) = 1 - f(q)h(q) - I_{\text{E}}(q). \tag{20}$$

Eve can try to guess the elements of the sequence $\{a_i\}$ as it were a truly random sequence. In this case, she correctly guesses approximately

$$n_{\text{correct}} \approx \frac{n}{2} = \frac{\gamma N}{2} \tag{21}$$

of the bases. Thus,

$$q(\gamma)=\frac{\gamma}{4},\quad \gamma(q)=4q,\quad I_{\rm E}(q)=2q, \tag{22}$$

and the secret fraction is as follows:

$$R(q)=I_{\rm B}(q)-I_{\rm E}(q)=1-f(q)h(q)-2q. \tag{23}$$

But, Eve can exploit the fact that the sequence $\{a_i\}$ is not random, but pseudorandom and contains some regularities. The estimation of $n_{\rm correct}(\gamma)$ for this case is rather involved and is given in Appendix C. Here, we give a summary of the analysis and results of Appendix C.

The analysis of pseudorandom sequences is an important part of classical cryptography. But, in classical cryptography it is usually assumed that Eve has a limited computing power and cannot use the brutal force attack. Here, we assume that Eve has an unlimited computing power, which is common for quantum cryptography. Suppose that Eve succeeded to guess a subset $\mathcal{K}_1\subset\mathcal{K}$ which contains the actual key k. In other words, it is unlikely that she correctly guesses the key k for large l (the probability is $1/|\mathcal{K}|\sim 2^{-l}$), but she can guess that $\overline{k}$ belongs to a certain subset $\mathcal{K}_1$. The probability of such success is equal to $|\mathcal{K}_1|/|\mathcal{K}|$. Then, she can choose not arbitrary $n=\gamma N$ positions to attack, but special positions. Namely, positions i such that the bits $a_i(k')$ coincide with each other for most $k'\in\mathcal{K}_1$ are preferable. Following this way of thinking, we arrive at the optimization problem. If $|\mathcal{K}_1|$ is less or comparable to l, we are able to solve it explicitly. This is done in Theorem 1 and adopted for practical situation in Corollary 1. In this case, we can use explicit formula (C15). If $|\mathcal{K}_1|$ is large, then we can still use formula (C15), but it gives too pessimistic (for Alice and Bob) estimate of $n_{\rm correct}(\gamma)$. A tighter bound can be obtained if we numerically solve the linear programming problem given in formula (C18) (Corollary 2). The linear programming problems are known to have efficient algorithms of solutions.

Of course, $n_{\rm correct}(\gamma)$ increases as $|\mathcal{K}_1|$ decreases (Eve adopted her attack to a tighter set of keys). However, the probability that $k\in\mathcal{K}_1$ is $|\mathcal{K}_1|/|\mathcal{K}|$, i.e., small whenever $|\mathcal{K}_1|$ is small. Thus, both estimates (C15) and (C18) are dependent on the additional parameter $\varepsilon=|\mathcal{K}_1|/|\mathcal{K}|$, i.e., $n_{\rm correct}=n_{\rm correct}(\gamma,\varepsilon)$. The parameter ε can be called the failure probability: the probability that Eve will succeed to guess a more tight set containing the actual key, other words, that she will be more lucky than we expect. The emergence of such (in)security parameter is common for QKD security proofs [20].

In short, we use Eq. (C15) (explicit formula) or Eq. (C18) (linear programming problem which gives a tighter bound) to estimate $n_{\rm correct}(\gamma,\varepsilon)$ from above for given failure probability ε. These estimations can be substituted to Eq. (13) to find the function $q(\gamma)$ and then to Eq. (20) to obtain (numerically) the secret fraction.

It turns out that Eve can guess more elements of pseudorandom sequence than those of truly random sequence (see the end of Appendix C). By this reason, the BB84 protocol with pseudorandom sequence of bases gives higher secret key rates than the usual BB84 protocol (because of the absence of sifting), but lower secret key rates than the asymmetric BB84 protocol. Thus, we do not consider the BB84 protocol with pseudorandom sequence of bases as a real alternative to aBB84 and switch to the multibasis case. In the multibasis case, the larger number of bases that Eve can correctly guess for the pseudorandom sequence is compensated by additional uncertainty for Eve caused by the use of many (instead of two) bases. In Sec. III C, we will compare the results of the multibasis PRB protocol with BB84 and aBB84 and show that the multibasis protocol can give slightly better results than the aBB84 protocol.

B. Multibasis case

Here, we investigate the intercept-resend attack for the multibasis version of the protocol. Denote the difference between the Eve's guess of the ith angle $\varphi_i^{\rm E}$ and the actual angle $\varphi_i(k)$ as Δ_i and let $(b_i^{(1)},\dots,b_i^{(m)})$ be its binary expansion:

$$\Delta_i=\varphi_i^{\rm E}-\varphi_i(k)=\frac{\pi}{2}\sum_{j=1}^{m}b_i^{(j)}2^{-j}. \tag{24}$$

For each register j, the upper bound of the number of bits $b_i^{(j)}$ correctly guessed by Eve is $n_{\rm correct}(\gamma)$ given by either Eq. (C15) or Eq. (C18) from Appendix C [i.e., now, $n_{\rm correct}(\gamma)$ denotes the number of correctly guessed bits in a single register]. Denote $T\subset\{1,\dots,N\}$ the set of pulses intercepted by Eve: $|T|=n=\gamma N$. Let us pick a position from T at random. For each register, consider the event that the corresponding bit is correctly guessed. The probability of this event is (at most) $n_{\rm correct}(\gamma)/(\gamma N)$. Since the keys for different registers are chosen independently, these events are independent. Therefore, one has

$$\begin{aligned}\Pr\left[\Delta_i=\frac{\pi t}{2M}\right]&\equiv p_t(\gamma)\\ &=\prod_{j=0}^{m-1}\Pr\left[b_i^{(j)}=\lfloor 2^{-j}t\rfloor \bmod 2\right]\\ &=\left(\frac{n_{\rm correct}\left(\gamma,\frac{\varepsilon}{m}\right)}{\gamma N}\right)^{\#0(t)}\left(\frac{n_{\rm incorrect}\left(\gamma,\frac{\varepsilon}{m}\right)}{\gamma N}\right)^{\#1(t)},\end{aligned} \tag{25}$$

where $\#0(t)$ and $\#1(t)$ are the numbers of 0's and 1's in the binary expansion of t. Note the argument ε/m of the function $n_{\rm incorrect}$: If the probability that Eve correctly guesses more then a given number of bits in a single register is not greater than ε/m, then the probability that Eve correctly guesses more than a given number of bits in each of m registers is not greater than ε.

Remark 6: For Eve, the correct guessing of the highest-order bit $b_i^{(1)}$ in the binary expansion (24) is of the most importance. Thus, her optimal strategy is to chose positions to intercept which maximize the number of correctly guessed elements in the sequence for the first register $(b_1^{(1)},b_2^{(1)},\dots)$. The maximal number of correct guesses is bounded from above as $n_{\rm correct}(\gamma,\varepsilon)$. Since Eve adjusts attack to optimize the number of correct guesses in the first register, she is not so good in the number of correct guesses in further registers. But, in favor of Eve, we bounded the number of correct guesses for other registers from above also by the same quantity $n_{\rm correct}(\gamma,\varepsilon)$.

Now, let us derive formulas for QBER and Eve's mean information on a raw key bit. For simplicity, let us drop the subscript i. Denote $x \in \{0,1\}$ the bit value transmitted by Alice, $y,z \in \{0,1\}$ the results of Bob's and Eve's measurements. We then have

$$p(z|x,\Delta) = \cos^2\left[\Delta + \frac{\pi}{2}(x-z)\right],$$

$$p(y|z,\Delta) = \cos^2\left[\Delta + \frac{\pi}{2}(z-y)\right], \qquad (26)$$

$$\begin{aligned} p(y \neq x|\Delta) &= p(y \neq x|z = x,\Delta)p(z = x|\Delta) \\ &\quad + p(y \neq x|z \neq x,\Delta)p(z \neq x|\Delta) \\ &= \tfrac{1}{2}\sin^2(2\Delta) = \tfrac{1}{4}[1-\cos(4\Delta)], \end{aligned} \qquad (27)$$

$$\begin{aligned} p(y \neq x) &= \sum_{j=0}^{M-1} p_j(\gamma)p\left(y \neq x|\Delta = \frac{\pi j}{2M}\right) \\ &= \frac{1}{4}\sum_{j=0}^{M-1} p_j(\gamma)\left[1-\cos\left(\frac{2\pi j}{M}\right)\right], \end{aligned} \qquad (28)$$

where $p(y \neq x)$ is the probability of error in Alice's and Bob's bit for an intercepted position.

To obtain the QBER value, one has to multiply this quantity on the fraction of intercepted positions:

$$q = \gamma p(y \neq x) = \frac{\gamma}{4}\sum_{j=0}^{M-1} p_j(\gamma)\left[1-\cos\left(\frac{2\pi j}{M}\right)\right]. \qquad (29)$$

The Eve's information on an intercepted raw key bit x is as follows:

$$I_{\mathrm{E}}^{\text{intercepted}}(\gamma) = 1 - \sum_{j=0}^{M-1} p_j(\gamma)h\left[\cos^2\left(\frac{\pi j}{2M}\right)\right]. \qquad (30)$$

The mean Eve's information on a raw key bit then has the following form:

$$I_{\mathrm{E}}(\gamma) = \gamma I_{\mathrm{E}}^{\text{intercepted}}(\gamma). \qquad (31)$$

From Eq. (29) we can find the inverse function $\gamma(q)$, which expresses the fraction of intercepted positions γ dependent on the measured value of QBER q. Then, we have

$$I_{\mathrm{E}}(q) = \gamma(q)\left\{1 - \sum_{j=0}^{M-1} p_j(\gamma(q))h\left[\cos^2\left(\frac{\pi j}{2M}\right)\right]\right\}. \qquad (32)$$

To calculate the secret fraction, (32) should be substituted into (20).

It is useful to calculate the Eve's information in case $N \to \infty$ (also, $l' \to \infty$ since $N = 2^{l'}$). In this case, $n_{\text{correct}}(\gamma)/n = \frac{1}{2}$ [see the Remark 9 in Appendix C and Eq. (A4) in Appendix A]. Then, we arrive at the following expression:

$$q = \gamma\left[\frac{1}{4} - \frac{1}{M}\sum_{j=0}^{M-1}\cos\left(\frac{2\pi j}{M}\right)\right] = \frac{\gamma}{4}. \qquad (33)$$

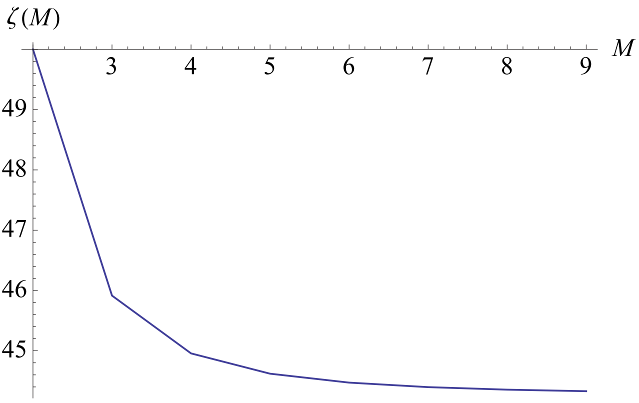

FIG. 2. Function $\zeta(M)$.

The mean Eve's information is then

$$I_{\mathrm{E}}(q) = 4q\left\{1 - \frac{1}{M}\sum_{j=0}^{M-1} h\left[\cos^2\left(\frac{\pi j}{2M}\right)\right]\right\} \equiv 4q\zeta(M), \qquad (34)$$

where $\zeta(M)$ is a decreasing function of M:

$$\lim_{M\to\infty} \zeta(M) \approx 0.4427. \qquad (35)$$

The function $\zeta(M)$ is shown in Fig. 2. Since $\zeta(M) < 0.5$ for $M > 2$, the multibasis version protocol has advantage over the BB84 with pseudorandom sequence of bases.

Finally, we arrive at the following expression for the secret fraction:

$$R = 1 - f(q)h(q) - I_{\mathrm{E}}(q) = 1 - f(q)h(q) - 4q\zeta(M). \qquad (36)$$

C. Numerical comparison

We compare the secret key rates per bit of the raw key (secret fractions) for the multibasis PRB protocol [Eqs. (20) and (32)] with those for the BB84 protocol and the asymmetric BB84. The secret key rate per bit of the raw key $R(q)$ for the BB84 protocol is given by the following expression:

$$R(q) = \tfrac{1}{2}[1 - f(q)h(q) - 2q], \qquad (37)$$

where the factor $\frac{1}{2}$ is due to sifting of a half of positions.

We consider the following variation of the asymmetric BB84 protocol. For each pulse, Alice and Bob choose independently the basis $\{|0\rangle,|\pi/2\rangle\}$ ("the signal basis") with the probability $1-p$ or the basis $\{|\pi/4\rangle,|3\pi/4\rangle\}$ ("the test basis") with the probability p. The first basis is used to establish the secret key, while the second one is used to detect the eavesdropping. The sifting rate is, thus, on average, $1-(1-p)^2$. Alice and Bob announce the bit values for positions encoded using the test basis and calculate the QBER in the test basis $q_\times$. Also, after the error-correction step, they calculate the QBER in the signal basis q_+ [see Sec. II, step (iv) of the protocol]. We compare the performances of QKD protocols in the case of the absence of actual eavesdropping, where the QBER is caused only by natural noise, then, on average, $q_+ = q_\times = q$.

The mean Eve's information per bit of the sifted key is $2q$. But now, for small p, we cannot neglect the statistical fluctuations. Use the formula from Ref. [42] to treat statistical fluctuations:

$$R(q) = (1-p)^2[1 - f(q)h(q) - 2(q+\theta)], \tag{38a}$$

where θ is a minimal positive number such that

$$\frac{\sqrt{N_+ + N_\times}}{\sqrt{q(1-q)N_+N_\times}} 2^{-(N_+ + N_\times)\xi(\theta)} \leqslant \varepsilon, \tag{38b}$$

$$\xi(\theta) = (q + \theta - v\theta) - vh(q) - (1-v)h(q+\theta). \tag{38c}$$

Here, ε is the failure probability, N_+ and $N_\times$ are the numbers of positions where both Alice and Bob have chosen the signal basis and the test basis, respectively, and $v = N_\times/(N_\times + N_+)$.

Thus, a smaller p leads to lower sifting, but also to higher statistical fluctuations of the potential Eve's information that we have to take into account. In the calculations, we optimized (38) over p for each value of q.

In the calculations, we use the following parameters: $L = N = 10^{10} - 33, l' = \log_2 L \approx 16, m = 10\,(M = 1024\,\text{bases})$. To obtain function $\gamma(q)$ in Eq. (32), we have used formula (C18) with the failure probability $\varepsilon \approx 10^{-6}$ (more precisely, $\varepsilon/m = S/L$ for $S = 1000$). The parameter s in Eq. (C18) was set to $s = 12$. The number of pulses sent by Alice for all protocols is equal to L. The number of pulses received by Bob is $N_\text{r} = N$ if the quantum channel is lossless. For a lossy channel we have taken a realistic loss rate $N_\text{r}/N = 0.001$. The failure probability for aBB84 in (38) was also taken as $\varepsilon \approx 10^{-6}$. For N_+ and $N_\times$ in Eq. (38), we use the average values $N_+ = (1-p)^2 N_\text{r}$ and $N_\times = p^2 N_\text{r}$. The results are given in Fig. 3.

It is clearly seen that the PRB protocol gives twice as large secret fraction as the BB84 protocol (due to the absence of sifting). In PRB, the larger number of bases that Eve can correctly guess for the pseudorandom sequence is compensated by additional uncertainty for Eve caused by the use of many (instead of two) bases. This results in approximately the same secret fractions for PRB and asymmetric BB84 for the lossless channels. However, if the channel is lossy, Alice and Bob have to increase p and, hence, sifting rate, to collect large enough statistics for the test basis. In this case, we can see that PRB gives slightly better results.

Note that the losses do not decrease the secret fraction for the PRB protocol since the estimate of the number of bases correctly guessed by Eve is dependent on L and not on N_r. Moreover, if optimal positions to attack are lost (see the end of Sec. III A for general comments on the optimal Eve's attack on the PRB protocol and Appendix C for rigorous analysis), the losses even weaken Eve's attack.

Remark 7: Let us discuss the initial key consumption for the considered protocols. It is well known that the BB84 protocol (both symmetric and asymmetric versions of it) require short initial secret key for authentication purposes, which cannot be reused. Thus, a part of the generated key should be consumed for authentication in the next run of a QKD protocol. This is also true for our protocol. But, additionally, our protocol consumes the key for the seed for PRNG in the next run since this seed also cannot be reused. Thus, for an honest comparison of our protocol with BB84, we should compare the "net generation rates," i.e., secret key generation rates minus key secret key consumption rates. Let us ignore the key consumption for the authentication because there is no difference in the authentication procedures for the BB84 and PRB protocols. Moreover, PRB has less data to authenticate since there is no announcement of bases. If we ignore the authentication problem, then the BB84 protocol does not consume the key at all, while PRB consumes some amount of the key. However, from the data provided above, we can see that the key consumption is negligible for PRB: in a single run of the protocol, we used the initial secret key with the length 160 bits and generate a key with the length of order 10^{10} for lossless case and of order 10^7 for the lossy case. Thus, the key consumption (160 bits) is several orders smaller than the key generation and can be neglected.

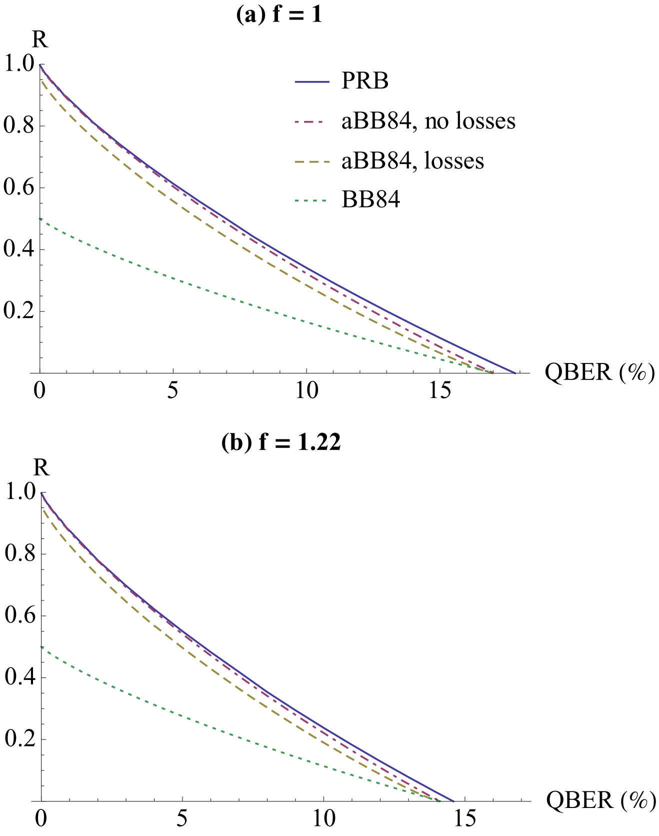

FIG. 3. Secret key rates (per bit of the raw key, before sifting) of the presented pseudorandom bases (PRB) protocol, the BB84 protocol, and the asymmetric BB84 (aBB84) protocol with and without losses in the quantum channel, when the efficiency of error correction achieves the theoretical limit $f = 1$ (a) and for practically achievable efficiency $f = 1.22$ (b). It can be seen that the considered protocol gives better secret key rates than the BB84 protocol and approximately the same rates as the asymmetric BB84 protocol.

Remark 8: Since we consider a protocol with many bases, it is worthwhile to compare it with the known six-state BB84 protocol [12] which uses three bases: $\{|0\rangle, |1\rangle\}$, $\{|\pi/4\rangle, |3\pi/4\rangle\}$, and $\{|0\rangle + i|1\rangle, |0\rangle - i|1\rangle\}$. The use of three bases provides higher uncertainty for Eve comparing to the original (four-state) BB84. However, this protocol is more difficult to implement in practice than the original BB84. In contrast, PRB can be implemented with the same hardware as BB84: PRB requires just another regulation of the angle φ in state (2). This can be done on the software level.

However, if we have a six-state implementation, we can also consider its pseudorandom modification, which also uses many numbers of bases of the form $\{|\mathbf{r}\rangle,|-\mathbf{r}\rangle\}$, where $|\mathbf{r}\rangle$ is a qubit state corresponding to the point $\mathbf{r}$ on the unit sphere (the Bloch sphere). In this case, $|\mathbf{r}\rangle$ should be chosen pseudorandomly from some discrete set of points on the Bloch sphere. Moreover, in the six-state BB84 protocol, the sifting efficiency is not $\frac{1}{2}$ but $\frac{1}{3}$. Hence, the use of a protocol with pseudorandom choices of bases, which allows to avoid sifting at all, may be even more advantageous.

IV. PHOTON-NUMBER SPLITTING PLUS QUANTUM-STATE DISCRIMINATION ATTACK

We performed the analysis for an ideal case, where the light source is assumed to be single photon. Practically, usually weak coherent pulses are used [6]. This gives possibilities to Eve to perform additional attacks, for example, photon-number splitting (PNS) attack [43,44]. In this attack it is assumed that quantum technologies are fully accessible for Eve. Let us describe this attack for the BB84 protocol. Eve measures the number of photons in each pulse and, if the number of photons is at least two, takes one photon and saves it in her quantum memory. After the announcement of the bases, she measures this photon in the known basis and, so, obtains a bit of information about the raw keys without disturbance. This potential Eve's information must be taken into account by Alice and Bob. Eve also can stop the single-photon pulses to increase the fraction of the multiphoton pulses (i.e., the number of pulses about which she can obtain full information without disturbance). To detect such actions, the so-called decoy state method has been proposed and developed [45–49]. Its purpose allows one to obtain tight estimates on the number of single-photon pulses and the number of errors in these pulses.

For the case of the proposed pseudorandom basis protocol, Eve can also perform the PNS attack. While in BB84 she waits for the announcement of bases, here she waits for the moment when she gets full knowledge of the initial key and, hence, bases. To account for this attack, Alice and Bob can also use the decoy state method.

But, now Eve can perform another type of attack, which we will refer to as "photon-number splitting plus quantum-state discrimination" (PNS+QSD) attack. Namely, Eve has a possibility to use multiphoton pulses to get knowledge of the initial key during the transmission of quantum states. Recall that all our analyses above were based on the assumption that Eve has zero information on the initial key during the transmission of quantum states and has to guess the bases. But, now she can perform the following variant of the PNS attack. Again, she measures the number of photons in each pulse. If the number of photons in a pulse is at least three, she sends one photon to Bob (i.e., does not introduce disturbance) and takes two photons to her quantum memory. We have shown that a single photon without the knowledge of the raw key bit leaks no information about the basis [see Eq. (12)]. But, this is not true if Eve has two photons in the same state. Let us analyze this attack.

Suppose that Eve has intercepted n such double photons in positions $i_1,\dots,i_n$. Then, to get knowledge of the initial key x, she has to distinguish between 2^{l+n} states

$$|\psi(k,\mathbf{x})\rangle=\bigotimes_{j=1}^{n}\left|\varphi_{i_j}(k)+\frac{\pi}{2}x_{i_j}\right\rangle^{\otimes 2}, \tag{39}$$

where $k\in\{0,1\}^l$, $\mathbf{x}=\{x_{i_1},\dots,x_{i_n}\}\in\{0,1\}^n$.

Discrimination of quantum states (or hypothesis testing) is a famous problem in quantum information science [50]. We will use the following lower bound on the success probability $p_{\rm succ}$ of guessing the correct quantum state (in our case, correct k and $\mathbf{x}$) [51]:

$$\begin{aligned}p_{\rm succ}&\geqslant\frac{1}{2^{l+n}}\sum_k\sum_{\mathbf{x}}\frac{1}{\sum_{k'}\sum_{\mathbf{y}}\langle\psi(k,\mathbf{x})|\psi(k',\mathbf{y})\rangle^2}\\&=\frac{1}{2^{l+n}}\sum_k\sum_{\mathbf{x}}\frac{1}{1+\sum_{k'\neq k}\sum_{\mathbf{y}}\langle\psi(k,\mathbf{x})|\psi(k',\mathbf{y})\rangle^2},\end{aligned}$$

$$\begin{aligned}&\sum_{\mathbf{y}}\langle\psi(k,\mathbf{x})|\psi(k',\mathbf{y})\rangle^2\\&\quad=\prod_{j=1}^{n}\{\cos^4[\varphi_{i_j}(k)-\varphi_{i_j}(k')]+\sin^4[\varphi_{i_j}(k)-\varphi_{i_j}(k')]\}.\end{aligned}$$

We restrict the analysis to the case of two bases ($M=2$). The analysis of the multibasis case leads to more cumbersome calculations but qualitatively the same conclusions. If $k\neq k'$, then approximately a half of basis choices for the keys k and k' coincide. Hence,

$$\sum_{\mathbf{y}}\langle\psi(k,\mathbf{x})|\psi(k',\mathbf{y})\rangle^2\approx 2^{-n/2} \tag{40}$$

and

$$p_{\rm succ}\geqslant\frac{1}{1+(2^l-1)2^{-n/2}}, \tag{41}$$

i.e., Eve needs approximately $2l\ll N$ three-photon pulses to guess the secret key k with a non-negligible probability. Then she can measure the pulses in correct bases without disturbance. Hence, the protocol crucially requires single-photon sources.

V. DISCUSSIONS AND CONCLUSIONS

In this work, we have analyzed a prepare-and-measure QKD protocol. It uses the pseudorandom sequence of bases generated by the legitimate parties of communications from a common initial secret key (seed). The use of a common pseudorandom sequence of bases allows one to avoid the sifting procedure and, hence, losing the half of the key. Moreover, since the bases of Alice and Bob are always the same, they can use more than two bases.

The main result of this work is the calculation of the secret key rates of the proposed protocol for the intercept-resend attack presented in Fig. 3. The main technical ingredient is

Appendix C, where the mathematical tools of analysis of pseudorandom sequences in the context of quantum cryptography (where the adversary has unlimited computing power) are proposed. The main practical formulas derived in this appendix are (C15) and (C18). They give upper bound on the elements of a pseudorandom sequence that can be correctly guessed by the eavesdropper with unlimited computing power.

We have obtained that, for single-photon sources, the protocol gives twice as large secret key rates as the original BB84 protocol and in some cases gives slightly higher rates than the aBB84 protocol. However, we did some assumptions in favor of Eve (see, for example, Remark 6). More tight analysis, which requires the development of techniques of Appendix C, probably, will lead to even more significant advantage of the proposed pseudorandom multibasis protocol. The protocol strongly requires a single-photon source of light.

The mathematical tools developed in Appendix C can be used in different problems of quantum cryptography, for example, for rigorous estimation of how the use of pseudorandom sequences (instead of truly random ones) influence the security of the conventional prepare-and-measure QKD protocols, such as BB84, asymmetric BB84, etc. First steps in such analysis have been done in [30]. Our approach suggests that the pseudorandomness can be even turned to an advantage.

In general, future investigations of the power of classical pseudorandomness combined with quantum uncertainty in quantum cryptography are required. The fundamental difference with the consideration of pseudorandom sequences in conventional cryptography is that, in the latter case, one typically assumes the boundedness of the eavesdropper's computing power. For example, it is assumed that the eavesdropper cannot use the brute force attack to try all possible seeds for a PRNG. In contrast, in quantum cryptography we assume unlimited computing power of the eavesdropper. But, from the other side, the possibilities of the eavesdropper are limited by the quantum uncertainty principle. Thus, the analysis of quantum pseudorandom sequences may require novel mathematical methods.

ACKNOWLEDGMENTS

We are grateful to A. V. Duplinskiy, V. L. Kurochkin, and V. E. Ustimchik for fruitful discussions, and especially grateful to M. A. Korolev and N. Lütkenhaus for valuable suggestions. The development and description of the PRB protocol as well as analysis of guessing the seed for PRNG (Secs. I and II and Appendix B) were supported by the grant of the Russian Science Foundation (RSF), Project No. 17-71-20146. The security analysis against the intercept-resend attack (Secs. III–V and Appendices A and C) was supported by the grant of the President of the Russian Federation (Project No. MK- 2815.2017.1).

APPENDIX A: PSEUDORANDOMNESS PROPERTY OF THE LEGENDRE SEQUENCES

Here, we formulate the keystone pseudorandomness property of the Legendre sequences which is exploited in the present analysis, mainly in Appendix C. Note that

$$\overline{a}_{i+L} = \overline{a}_i, \tag{A1}$$

i.e., $\{\overline{a}_i\}$ is a periodic sequence. We then denote $\mathbb{Z}_L = \{0,\dots,L-1\}$ the residue ring with respect to integer addition and multiplication modulo L. For distinct elements $i_1,\dots,i_s \in \mathbb{Z}_L$ and binary $b_1,\dots,b_s$, denote $d_{i_1,i_2,\dots,i_s}(b_1,b_2,\dots,b_s)$ the number of $j \in \{0,\dots,L-1\}$ such that

$$\overline{a}_{j+i_1} = b_1,\quad \overline{a}_{j+i_2} = b_2,\quad \dots\quad \overline{a}_{j+i_s} = b_s, \tag{A2}$$

i.e., the number of occurrences of the pattern

$$*\dots*b_1*\dots*b_2*\dots*\dots*b_s \tag{A3}$$

in one period. Here, $*$ are "do-not-care" bits. In other words, we look for patterns with the bit values $b_1,\dots,b_s$ on positions $i_1,\dots,i_s$. Here, we do not care bit values on other positions.

We will use the following bounds [40]: for all distinct $i_1,\dots,i_s \in \mathbb{Z}_P$ and all binary $b_1,\dots,b_s$:

$$d_{i,j}(b_1,b_2) = \begin{cases}(L-3)/4, & (b_1,b_2) = (1,1)\\ (L+1)/4, & (b_1,b_2) \neq (1,1)\end{cases} \tag{A4a}$$

$$-W(s) \leqslant d_{i_1,i_2,\dots,i_s}(b_1,b_2,\dots,b_s) - \frac{L}{2^s} \leqslant W(s) \tag{A4b}$$

for $s \geqslant 3$, where

$$W(s) = \frac{\sqrt{L}[2^{s-1}(s-3)+2] + 2^{s-1}(s+1) - 1}{2^s}. \tag{A4c}$$

For large s, W can be approximated as $\sqrt{L}[(s-3)/2] + (s+1)/2$.

APPENDIX B: GUESSING THE SEED FOR PRNG

In our analysis we assumed that, after all stages of the protocol and after the transmission of a message encrypted with the use of the distributed key, Eve can correctly guess the initial secret key (the seed for the PRNG). In this appendix, we derive bounds on the number of qubits that Eve needs to intercept for correct guessing of the seed. We show that this assumption is not too pessimistic.

First, we specify assumptions on Eve's knowledge. Of course, Eve knows her measurement results of intercepted qubits $z_{i_1},\dots,z_{i_n}$. We further assume that, during the stage of error correction, Eve discovers (along with a syndrome or other messages sent via error correction) the positions where she introduces errors and the values of the bits x_i and y_i in such positions. This is indeed true if the Cascade protocol for error correction is used, but may be too pessimistic in the case of the use of one-way error-correcting codes (for example, LDPC codes). Let us also denote $c_i = x_i \oplus y_i$.

Moreover, we assume that Eve may know a part of the message encoded with the key distributed by the protocol ("known plain-text attack"). Let r bits of the secret key $(u_1,\dots,u_r)$ be distributed; $r < N$ due to the key contraction in the privacy amplification stage. The last l bits from this key are kept for the next session as a new initial key. The first $r-l$ bits are used for encryption of a message, for example, using one-time pad encryption. Eve may know a part of this message (or even the whole message) and, hence, the corresponding bits of the key $(u_1,\dots,u_q)$, $q \leqslant r-l$. She can use this knowledge as well as her results of quantum measurements to guess the unknown part of the distributed key and, in particular, the initial key for the next session. The knowledge of the initial key for the next session gives her a possibility to obtain the key of the next session without introducing errors since she can also construct the sequence $\{\varphi_i(k)\}$ in the next session.

For definiteness, let us assume that Bob is the side that corrects errors, so that after error correction Alice and Bob have the common key $x_1,\dots,x_N$. If the Toeplitz hashing is used for privacy amplification, then u_i are linear combinations (with respect to XOR) of x_j:

$$u_i=\sum_{j=1}^{N}t_{ij}x_j,\quad i=1,\dots,r \tag{B1}$$

where t_{ij} are the elements of a Toeplitz matrix.

Thus, Eve knows the following:

(i) her measurement results $z_{i_1},\dots,z_{i_n}$;

(ii) the syndrome of the used error-correcting code (or parities of certain subsets of positions if an interactive error-correcting procedure like "Cascade" is used);

(iii) whether she has introduced errors in the intercepted positions: $c_{i_1},\dots,c_{i_n}$; also she knows x_{i_j} and y_{i_j} if $c_{i_j}=1$;

(iv) $q\leqslant r-l$ outputs of linear combinations (B1).

For convenience of notations, let Eve attack the first r transmitted state and $i_j=j$. Also, to make the derivations simpler, we do the following modification of the protocol: let the angles $\varphi_i(k)$ in (3) are chosen from the set $\{\frac{\pi j}{2M}\}_{j=0}^{2M-1}$ rather than from $\{\frac{\pi j}{2M}\}_{j=0}^{M-1}$, $M\geqslant 2$. Thus, we add an additional pseudorandom binary register which is responsible for an additional rotation of the angle over $\pi/2$. This does not alter the security properties of the protocol (if the initial key is also added by one bit) since a basis rotated over $\pi/2$ is in fact the same basis as the initial one up to the interchange of the assigned bit values 0 and 1. But, as we said before, the bit value is supposed to be known to Eve after the transmission of an encrypted message. Such modification is useless for practice since we should spend the initial secret key on the additional register, but it is useful for the purposes of the present section: this modification makes the situation more symmetric and simplifies the analysis.

Thus, in expansion (10), we have an additional register:

$$\varphi_i(k)=\frac{\pi}{2}\sum_{j=0}^{m}a_i(k^{(j)})2^{-j}. \tag{B2}$$

Now, we are going to estimate the probability of guessing the seed

$$p_{\rm guess}=\max_k p(k|e_1,\dots,e_n), \tag{B3}$$

where $e_i=(x_i,z_i,c_i)$ is Eve's knowledge on the ith transmitted quantum state. Then, we have

$$p(k|e_1,\dots,e_n)=\frac{p(e_1,\dots,e_n|k)p(k)}{p(e_1,\dots,e_n)}=2^{-l}\frac{p(e_1|\varphi_1(k))\cdots p(e_n|\varphi_r(k))}{p(e_1,\dots,e_n)}. \tag{B4}$$

Let us find an expression for $p(e_i|\varphi_i)$:

$$\begin{aligned}p(e_i|\varphi_i)&=p(x_i|\varphi_i)p(z_i|\varphi_i,x_i)p(c_i|\varphi_i,x_i,z_i)=\frac{1}{2}p(z_i|\varphi_i,x_i)p(c_i|\varphi_i,x_i,z_i)\\&=\frac{1}{2}\cos^2\left[\beta_i-\varphi_i+\frac{\pi}{2}(z_i-x_i)\right]\cos^2\left[\beta_i-\varphi_i+\frac{\pi}{2}(z_i-x_i-c_i)\right]\\&=\frac{1}{8}\left\{\cos\frac{\pi c_i}{2}+\cos\left[2(\beta_i-\varphi_i)+\pi(x_i-z_i)-\frac{\pi c_i}{2}\right]\right\}^2\\&=\frac{1}{8}\left\{1-c_i+\cos\left[2(\beta_i-\varphi_i)+\pi(x_i-z_i)-\frac{\pi c_i}{2}\right]\right\}^2.\end{aligned} \tag{B5}$$

Here, we have used that $\cos\frac{\pi x}{2}=1-x$ if $x\in\{0,1\}$. Therefore, we arrive at the following expressions:

$$p(e_i)=\frac{1}{M}\sum_{j=0}^{M-1}p\left(e_i|\varphi_i=\frac{\pi j}{M}\right)=\begin{cases}3/16, & c_i=0,\\ 1/16, & c_i=1,\end{cases}\quad p(c_i)=\sum_{x_i,z_i\in\{0,1\}}p(e_i)=\begin{cases}3/4, & c_i=0,\\ 1/4, & c_i=1.\end{cases} \tag{B6}$$

Substitution of Eq. (B5) into Eq. (B4) yields

$$p(k|e_1,\dots,e_n)=\frac{2^{-l}}{p(e_1,\dots,e_n)}\prod_{i=1}^{n}\frac{1}{8}\left\{1-c_i+\cos\left[2(\beta_i-\varphi_i)+\pi(x_i-z_i)-\frac{\pi c_i}{2}\right]\right\}^2. \tag{B7}$$

If $n\leqslant l'$ (recall that l' is the length of the initial key for each register, while $l=ml'$ is the whole length of the initial key), then φ_i are approximately independent and uniformly distributed on their domain. Hence, Eve's variables $e_1,\dots,e_n$ are also approximately independent: $p(e_1,\dots,e_n)\approx p(e_1)\dots p(e_n)$. By properties of the Legendre sequences, for every combination of angles φ_i, there exists a key k generating this combination. Then, the maximization of Eq. (B7) is equivalent to maximization of every separate term in the product in its right-hand side, i.e., maximization of (B5).

Let us maximize (B5) over φ_i. We have

$$\max_{\varphi_i} p(e_i|\varphi_i) = \begin{cases} p(e_i|\beta_i) = 1/2, & c_i = 0,\, z_i = x_i, \\ p(e_i|\beta_i + \pi/2) = 1/2, & c_i = 0,\, z_i \neq x_i, \\ p(e_i|\beta_i + \pi/4) = 1/8, & c_i = 1. \end{cases} \tag{B8}$$

In other words, if Eve has not introduced an error and her measurement result z_i has coincided with x_i, then her optimal guess is that Alice has chosen the same basis as Eve: $\varphi_i = \beta_i$. If Eve has not introduced an error, but her measurement result z_i has not coincided with x_i, then her optimal guess is $\varphi_i = \beta_i + \pi/2$, i.e., Alice has chosen the basis different from Eve's basis by $\pi/2$: the basis rotated by $\pi/2$ coincides with the initial basis up to a bit flip. Finally, if Eve has introduced an error, then her optimal guess is $\varphi_i = \beta_i + \pi/4$, which corresponds to a situation that yields an error with the maximal probability (1/2).

Putting it all together, we arrive at the following expression:

$$\max_k p(k|e_1,\dots,e_n) = 2^{-l}\left(\frac{8}{3}\right)^{n_0} 2^{n_1} = 2^{-l+n_0\log\frac{8}{3}+n_1}$$
$$= 2^{-l+3n_0+n_1-n_0\log 3}, \tag{B9}$$

where n_0 and n_1 are number of positions where $c_i = 0$ and 1, respectively. Indeed, $n_0 + n_1 = n$. Since $p(c_i = 0) = 3/4$ and $p(c_i = 1) = 1/4$, if n is large, then $n_0 \approx 3n/4$, $n_1 = n/4$. Then,

$$\max_k p(k|e_1,\dots,e_n) \approx 2^{-l+n[\frac{3}{4}\log\frac{8}{3}+\frac{1}{4}]} = 2^{-l+n[\frac{5}{2}-\frac{3\log 3}{4}]}$$
$$\approx 2^{-l+1.3n}. \tag{B10}$$

Recall that this derivation is valid if $n \leqslant l'$. But, we are interested in the inverse case. Then, the analysis is more complicated. First, not every sequence of angles $\{\varphi_i\}$ is possible: different angles are not independent, hence, maximization of the numerator in (B7) is not reduced to maximization of its separate factors. Second, $e_1,\dots,e_n$ are also not independent (as the measurement results of dependent quantum states).

Nevertheless, we will still use formula (B9) as an upper bound for the guessing probability. The arguments are as follows. The most advantageous situation for Eve is when the current angle φ_i does not depend on the previous angles. In this case, the measurement gives Eve more information than the measurement in the case when Eve already has partial information on φ_i. Thus, in favor of Eve, we treat φ_i independent from each other even if $r > l'$ and use formulas (B9) and (B10). Numerical experiments confirm that the validity of these formulas as upper bounds. Then, Eve can guess a key in approximately (lower bound)

$$n = \frac{l}{\frac{5}{2}-\frac{3\log 3}{4}} \approx 0.76\,l. \tag{B11}$$

Thus, Eve needs order l intercepted positions to correctly guess the initial secret key. Our numerical experiments with short enough keys (up to $l' = 10$, which allows one to explicitly implement the proposed maximum likelihood method) and $m \leqslant 8$ show that formula (B11) is adequate as a rough estimate at least for the subkeys $k^{(0)}$ and $k^{(1)}$ that govern the highest-order (i.e., most important) bits in the binary expansion of angles (B2). One needs much more iterations to correctly guess the lowest-order bits since close quantum states are hard to distinguish. From the other side, lowest-order registers are less important. Hence, the assumption that Eve gets a knowledge of the initial key after the transmission of the cipher text (provided that she knows the plain text) seems to be not too pessimistic.

APPENDIX C: GUESSING IN PSEUDORANDOM BINARY SEQUENCES

In this Appendix we obtain an upper bound for the number of correctly guessed bits in a certain class of binary pseudorandom sequences. We assume that Eve has access to unlimited computing power, and we design an optimal attack for Eve.

Let us introduce assumptions about PRNG. Let $\{a_i(k)\}_{i=1}^{\infty}$ be a periodic sequence with the period L for any k, i.e., $a_{i+L}(k) = a_i(k)$. The set of keys is $\mathcal{K} = \mathbb{Z}_L = \{0,\dots,L-1\}$, the residue ring with respect to integer addition and multiplication modulo L. For distinct keys $k_1,\dots,k_s$ and binary $b_1,\dots,b_s$, denote

$$A_{b_1\dots b_s}(k_1,\dots,k_s) = \{i \in \mathbb{Z}_L \,|\, a_i(k_1) = b_1,\dots,a_i(k_s) = b_s\}. \tag{C1}$$

Let us assume that there exists $S \geqslant 2$ such that, for all distinct keys $k_1,\dots,k_S$ and for all binary $b_1,\dots,b_S$,

$$|A_{b_1,\dots,b_S}(k_1,\dots,k_S)| = \frac{L}{2^S}. \tag{C2}$$

It is also assumed that L is divided by 2^S.

Eve chooses the fraction $0 < \gamma \leqslant 1$ of positions that she will try to guess, i.e., she will try to guess γL positions in a period. Her aim is to choose the positions to maximize the fraction of the guessed outcomes. If Eve attacks all L positions, then, due to Eq. (C2), she guesses exactly a half of positions, which is expected when the sequence is truly random. We are going to prove the upper bound for the case $0 < \gamma < 1$.

Theorem 1: Let the pattern distribution satisfy (C2) and we try to guess $n = \gamma L$ positions in a period. Then the number of correctly guessed bits does not exceed

$$n_{\text{correct}}(\gamma) = L\left\{P_{s-1}(r) + \frac{s-r-1}{s}[\gamma - 2P_s(r)]\right\} \tag{C3}$$

with the probability at least $1 - s/|\mathcal{K}|$, for every $2 \leqslant s \leqslant S$. Here, $P_s(t) = \Pr[X_s \leqslant t]$, where X_s is a binomially distributed random variable with the number of experiments s and the success probability in one experiment 1/2 [i.e., $P_s(t)$ is a cumulative distribution function], and r is the integer such that $P_s(r) \leqslant \gamma/2$ but $P_s(r+1) > \gamma/2$.

Proof. Let us consider s arbitrary distinct keys $k_1,\dots,k_s$. Let $T \subset \mathbb{Z}_L$ be the set of positions chosen by Eve, $|T| = \gamma L$. Denote $n_{b_1\dots b_s} = |A_{b_1\dots b_s}(k_1,\dots,k_s) \cap T|$. Let it be known that the actual key k is one of the keys $k_1,\dots,k_s$. Let $n_{\text{correct}}^{(i)}$ be the number of our correct guesses provided that the actual key is k_i, $i = 0,\dots,s-1$. We try to maximize $\min(n_{\text{correct}}^{(1)},\dots,n_{\text{correct}}^{(s)})$, i.e., the guaranteed number of correct guesses.

Let the set T be fixed. For each position $j \in T$, if the majority of the values $a_j(k_1),\dots,a_j(k_s)$ is 0 (1), then the optimal guess a_j^{guess} is also equal to 0 (1), i.e.,

$$a_j^{\text{guess}} = \begin{cases} 0, & \mathrm{HW}(a_j(k_1),\dots,a_j(k_s)) \leqslant s/2, \\ 1, & \text{otherwise}, \end{cases} \tag{C4}$$

where $\mathrm{HW}(b_1,\dots,b_s)$ is the Hamming weight of the vector $(b_1,\dots,b_s)$. Then,

$$n_{\text{correct}}^{(i)} = \sum_{\substack{(b_1,\dots,b_s):\\ \mathrm{HW}(b_1,\dots,b_s)\leqslant s/2,\\ b_i=0}} n_{b_1\dots b_s} + \sum_{\substack{(b_1,\dots,b_s):\\ \mathrm{HW}(b_1,\dots,b_s)> s/2,\\ b_i=1}} n_{b_1\dots b_s} \tag{C5}$$

for $i=1,\dots,s$.

Thus, we have the following optimization problem with respect to 2^s integers $n_{b_1\dots b_s}$:

$$\min\left(n_{\text{correct}}^{(1)},\dots,n_{\text{correct}}^{(s)}\right) \to \max\,, \quad \sum_{(b_1,\dots,b_s)} n_{b_1\dots b_s} = \gamma L, n_{b_1\dots b_s} \leqslant 2^{-s}L, \quad \forall\,(b_1,\dots,b_s), \tag{C6}$$

where the last condition is a consequence of (C2). From the symmetry of the problem we can put

$$n_{b_1\dots b_s} = n_{\mathrm{HW}(b_1,\dots,b_s)} \tag{C7}$$

[that is, only the number of keys k_i such that $a_j(k_i)=a_j^{\text{guess}}$ for a certain position j matters]. Thus, $n_{\text{correct}}^{(1)}=\dots=n_{\text{correct}}^{(s)}=n_{\text{correct}}$. Denote also $\nu_t=n_t/L$ and $\nu_{\text{correct}}=n_{\text{correct}}/L$.

The number of vectors $(b_1,\dots,b_s)$ with the Hamming weight t is equal to $\binom{s}{t}$ (a binomial coefficient). If we have a constraint $b_i=1$ for fixed i [like in the summation in (C5)], then the number of vectors with the Hamming weight t is equal to $\binom{s-1}{t-1}$. If we have a constraint $b_i=0$ for fixed i, then the number of vectors with the Hamming weight t is equal to $\binom{s-1}{t}$. Thus,

$$\nu_{\text{correct}} = \sum_{t=0}^{\lfloor s/2\rfloor}\binom{s-1}{t}\nu_t + \sum_{t=\lfloor s/2\rfloor+1}^{s}\binom{s-1}{t-1}\nu_t, \tag{C8}$$

and the optimization problem (C6) is reduced to

$$\nu_{\text{correct}} = \sum_{t=0}^{\lfloor s/2\rfloor}\binom{s-1}{t}\nu_t + \sum_{t=\lfloor s/2\rfloor+1}^{s}\binom{s-1}{t-1}\nu_t \to \max\,, \quad \sum_{t=0}^{s}\binom{s}{t}\nu_t=\gamma, \quad \nu_t\leqslant 2^{-s}, \quad t=0,\dots,s. \tag{C9}$$

Obviously, an optimal choice is to assign the maximally possible value 2^{-s} to ν_t with t close to 0 or s. This means that we prefer positions where the guessed value is true for large number of keys. In other words, we primarily try to maximize ν_0 and ν_s, then try to maximize ν_1 and ν_{s-1}, and so on. The restriction of this process is the first constraint in (C9). Denote r the minimal integer such that

$$\sum_{t=0}^{r}\binom{s}{t}2^{-s} = \sum_{t=s-r}^{s}\binom{s}{t}2^{-s} \leqslant \frac{\gamma}{2}, \text{ but} \tag{C10}$$

$$\sum_{t=0}^{r+1}\binom{s}{t}2^{-s} = \sum_{t=s-r-1}^{s}\binom{s}{t}2^{-s} > \frac{\gamma}{2}. \tag{C11}$$

Equations (C10) and (C11) can be rewritten as $P_s(r)\leqslant\gamma/2$ and $P_s(r+1)>\gamma/2$.

Thus, ν_t are set to maximally possible value 2^{-s} for $t\leqslant r$ and $t\geqslant s-r$. Then, ν_{r+1} and ν_{s-r-1} are assigned by as large a value as possible:

$$\nu_{r+1}=\nu_{s-r-1}=\left[\frac{\gamma}{2}-P_s(r)\right]\binom{s}{r+1}^{-1}. \tag{C12}$$

Other ν_t (i.e., for $r+2<t<s-r-2$) are set to zero. The optimal value of the target function ν_{correct} is

$$\begin{aligned}\nu_{\text{correct}} &= \sum_{t=0}^{r}\binom{s-1}{t}2^{-s} + \binom{s-1}{r+1}\binom{s}{r+1}^{-1}\left[\frac{\gamma}{2}-P_s(r)\right]\\ &\quad+\sum_{t=s-r}^{s}\binom{s-1}{t-1}2^{-s} + \binom{s-1}{s-r-1}\binom{s-r-1}{r+1}^{-1}\left[\frac{\gamma}{2}-P_s(s-r)\right]\\ &= P_{s-1}(r)+\frac{s-r-1}{s}[\gamma-2P_s(r)].\end{aligned} \tag{C13}$$

This result means that, for any s keys, we cannot choose γL positions such that more than n_{correct} guesses given by formula (C3) are correct for all keys. This means that the number of correct guesses cannot be larger than n_{correct} with the probability at least $1 - s/|\mathcal{K}|$. ■

Remark 9: If, in (C3), $s \to \infty$, then, by definition of r, we have $P_s(r) \to \gamma/2$ and $P_{s-1}(r) \to \gamma/2$, so, $n_{\mathrm{correct}} \to L\gamma/2$. Recall that $L\gamma$ is the number of terms we try to guess, i.e., the fraction of correct guesses tends to a $1/2$, as in the case of truly random sequences.

Now let us relax condition (C2).

Corollary 1: Let the pattern distribution satisfy

$$|A_{b_1,\dots,b_s}(k_1,\dots,k_s)| = \frac{L}{2^s} + W(s) \tag{C14}$$

for all s from some range, where $W(s)$ is some function. We try to guess $n = \gamma L$ positions in a period. Then, with the probability at least $1 - \varepsilon$, the number of correctly guess bits does not exceed

$$n_{\mathrm{correct}}(\gamma,\varepsilon) = L'\left\{P_{s-1}(r) + \frac{s-r-1}{s}[\gamma' - 2P_s(r)]\right\}, \tag{C15}$$

where $s = \varepsilon|\mathcal{K}|$. Here,

$$L' = L'(s) = L\left[1 + \frac{2^s W(s)}{L}\right], \quad \gamma' = \gamma'(s) = \gamma\left[1 + \frac{2^s W(s)}{L}\right]^{-1} \tag{C16}$$

and r is the integer such that $P_s(r) \leqslant \gamma'/2$ but $P_s(r+1) > \gamma'/2$.

Note that, for the PRNG based on the Legendre symbol, (C14) is satisfied due to (A4) and (8).

Proof. The proof is the same, but optimization problem (C6) is modified into

$$\min\left(n^{(1)}_{\mathrm{correct}},\dots,n^{(s)}_{\mathrm{correct}}\right) \to \max, \quad \sum_{(b_1,\dots,b_s)} n_{b_1\dots b_s} = \gamma L = \gamma' L', \quad n_{b_1\dots b_s} \leqslant 2^{-s}L + W(s) = 2^{-s}L', \quad \forall\,(b_1,\dots,b_s). \tag{C17}$$

Thus, formula (C15) holds with the substitutions of L and γ by L' and γ'. ■

If $s > \lceil L \rceil = l$, formulas (C3) and (C15) may be too optimistic for Eve. But, we can obtain more tight bounds.

Corollary 2: Let (C14) be satisfied for some s. Other conditions are as in Corollary 1. Then, with the probability at least $1-\varepsilon$, the number of correctly guess bits does not exceed

$$n_{\mathrm{correct}}(\gamma,\varepsilon) = \nu^*_{\mathrm{correct}} L'. \tag{C18}$$

Here, ν^*_{correct} is the solution of the following linear programming problem for $S = \varepsilon|\mathcal{K}|$:

$$\nu_{\mathrm{correct}} = \sum_{t=0}^{\lfloor S/2\rfloor}\binom{S-1}{t}\nu_t + \sum_{t=\lfloor S/2\rfloor+1}^{S-1}\binom{S-1}{t-1}\nu_t \to \max, \quad \sum_{t=0}^{S}\binom{S}{t}\nu_t = \gamma', \quad \sum_{t=0}^{S-s}\binom{S-s}{t}\nu_{h+t} \leqslant 2^{-s}, \quad h = 0,\dots,s \tag{C19}$$

with the agreement $\nu_t = \nu_{S-t}$, i.e., the actual number of variables in the optimization problem is $\lfloor (S+1)/2\rfloor$.

Proof. Let us consider S arbitrary keys $k_1,\dots,k_S$, but condition (C14) is satisfied for $s \leqslant S$. A generalization of optimization problem (C17) to this case is

$$\min\left(n^{(1)}_{\mathrm{correct}},\dots,n^{(s)}_{\mathrm{correct}}\right) \to \max, \quad \sum_{(b_1,\dots,b_S)} n_{b_1\dots b_S} = \gamma' L', \quad \sum_{\substack{(b_1,\dots,b_S):\\ b_{i_1}=c_1,\dots,b_{i_s}=c_s}} n_{b_1\dots b_S} \leqslant 2^{-s}L', \quad \forall\,(c_1,\dots,c_s),\ 1 \leqslant i_1 < \dots < i_s \leqslant S, \tag{C20}$$

where

$$n^{(i)}_{\mathrm{correct}} = \sum_{\substack{(b_1,\dots,b_S):\\ \mathrm{HW}(b_1,\dots,b_S)\leqslant S/2,\\ b_i=0}} n_{b_1\dots b_S} + \sum_{\substack{(b_1,\dots,b_S):\\ \mathrm{HW}(b_1,\dots,b_S)> S/2,\\ b_i=1}} n_{b_1\dots b_S}. \tag{C21}$$

Again, by the symmetry, we can put

$$n_{b_1\dots b_S} = n_{\mathrm{HW}(b_1,\dots,b_S)} \tag{C22}$$

and $n_t = n_{S-t}$, so, the problem is reduced to (C19) (where $\nu_t = n_t/L'$ and $\nu_{\mathrm{correct}} = n_{\mathrm{correct}}/L'$). Let us comment the last set of constraints. If $\mathrm{HW}(c_1,\dots,c_s) = h$, then the left-hand side of the last constraint in (C20) is reduced to

$$\sum_{t=h}^{h+S-s}\binom{S-s}{t-h}\nu_t = \sum_{t=0}^{S-s}\binom{S-s}{t}\nu_{h+t}, \tag{C23}$$

which coincides with that of (C19). ■

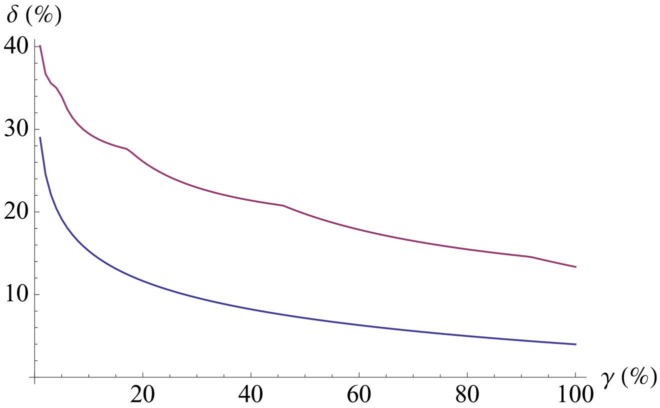

FIG. 4. Results of analysis of the Legendre sequences in terms of $\delta(\gamma) = \nu_{\text{correct}}(\gamma) - 1/2$ by formulas (C15) (top curve) and (C18) (bottom curve) for $L = 10^{10} - 33$, $\log_2 L \approx 16$, $s = 12$, and $S = 1000$.

Remark 10: The parameter s in (C19) is, in fact, an optimization parameter. If we use the Legendre sequences as PRNG, then, according to (A4), s should be taken between $\sqrt{l}$ and l: smaller s leads to less tight bounds, while larger s lead to large deviations $W(s)$ in (A4), which also lead to less tight bounds.

The comparison of the results of formulas (C15) and (C18) for Legendre sequences (see Sec. II A and Appendix A) is given in Fig. 4. We took $L = 10^{10} - 33$, $\log_2 L \approx 16$, $s = 12$ [for both (C15) and (C18)], and $S = 1000$, $\varepsilon = S/L \approx 10^{-7}$. We calculate the quantity $\delta(\gamma) = \nu_{\text{correct}}(\gamma) - 1/2$, i.e., deviation of ν_{correct} from the mean value $1/2$ in the case of random guessing.

It is clearly seen that the results of (C18) are significantly better. But, the picture will be incomplete if we do not compare these results with the corresponding estimates for truly random sequences. To estimate the number of correct guesses for truly random sequences, we can use the Hoeffding's inequality [52]: if X is a binomially distributed random variable with γL trials and the probability of success $1/2$ in one trial, then

$$\Pr[X \geqslant (1/2 + \delta)\gamma L] \leqslant e^{-2\delta^2 \gamma L} = \varepsilon \tag{C24}$$

or

$$\delta = \delta(\gamma) = \sqrt{\frac{1}{2\gamma L} \ln \frac{1}{\varepsilon}}. \tag{C25}$$

For $\varepsilon = 10^{-7}$, we have $\delta(\gamma) \approx 2.8 \times 10^{-4} = 0.028\%$ whenever $\gamma \geqslant 0.01 = 1\%$, which is of several orders of magnitude smaller than the results obtained for the pseudorandom sequences. Thus, from the cryptanalyst's point of view (if it has enough computing power), an optimal guesses of elements of pseudorandom sequences gives much better results than the simple random guessing.

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[2] P. W. Shor, SIAM J. Comput. **26**, 1484 (1997).
[3] G. S. Vernam, J. Am. Inst. Electr. Eng. **45**, 109 (1926).
[4] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).
[5] B. Schneier, *Applied Cryptography* (Wiley, New York, 1996).
[6] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002); V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dusek, N. Lütkenhaus, and M. Peev, *ibid.* **81**, 1301 (2009).
[7] H.-K. Lo, M. Curty, and K. Tamaki, Nat. Photon. **8**, 595 (2014); E. Diamanti, H.-K. Lo, and Z. Yuan, npj Quantum Inf. **2**, 16025 (2016).
[8] C. H. Bennet and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India* (IEEE, New York, 1984), p. 175.
[9] S. Wiesner, SIGACT News **15**, 78 (1983).
[10] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[11] C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).
[12] H. Bechmann-Pasquinucci and N. Gisin, Phys. Rev. A **59**, 4238 (1999).
[13] H.-K. Lo, H. F. Chau, and M. Ardehali, J. Cryptol. **18**, 133 (2005).
[14] V. Scarani, A. Acín, G. Ribordy, and N. Gisin, Phys. Rev. Lett. **92**, 057901 (2004).
[15] K. Inoue, E. Waks, and Y. Yamamoto, Phys. Rev. Lett. **89**, 037902 (2002); Phys. Rev. A **68**, 022317 (2003).
[16] T. Moroder, M. Curty, C. C. W. Lim, L. P. Thinh, H. Zbinden, and N. Gisin, Phys. Rev. Lett **109**, 260501 (2012).
[17] K. Tamaki, M. Koashi, and G. Kato, arXiv:1208.1995.
[18] D. Stucki, N. Brunner, N. Gisin, V. Scarani, and H. Zbinden, Appl. Phys. Lett. **87**, 194108 (2005).
[19] F. Grosshans, G. Van Assche, J. Wenger, R. Brouri, N. J. Cerf, and Ph. Grangier, Nature (London) **421**, 238 (2003).
[20] M. Tomamichel, C. C. W. Lim, N. Gisin, and R. Renner, Nat. Commun. **3**, 634 (2012).
[21] Y. V. Kurochkin, SPIE Proc. **5833**, 213 (2005).
[22] V. L. Kurochkin and Yu. V. Kurochkin, Phys. Part. Nucl. Lett. **6**, 605 (2009).
[23] Y. V. Kurochkin, A. K. Fedorov, and V. L. Kurochkin, *QCrypt 2015 5th International Conference on Quantum Cryptography, Tokyo, Japan* (unpublished).
[24] H. P. Yuen, arXiv:quant-ph/0311061.
[25] O. Hirota, K. Kato, M. Sohma, T. Usuda, and K. Harasawa, SPIE Proc. **5551**, 206 (2004).
[26] O. Hirota, M. Sohma, M. Fuse, and K. Kato, Phys. Rev. A **72**, 022335 (2005).
[27] D. J. Lum, J. C. Howell, M. S. Allman, T. Gerrits, V. B. Verma, S. W. Nam, C. Lupo, and S. Lloyd, Phys. Rev. A **94**, 022315 (2016).
[28] Y. Liu, Z. Cao, C. Wu, D. Fukuda, L. You, J. Zhong, T. Numata, S. Chen, W. Zhang, S.-C. Shi, C.-Y. Lu, Z. Wang, X. Ma, J. Fan, Q. Zhang, and J.-W. Pan, Phys. Rev. A **94**, 020301 (2016).
[29] N. Walenta, A. Burg, D. Caselunghe, J. Constantin, N. Gisin, O. Guinnard, R. Houlmann, P. Junod, B. Korzh, N. Kulesza,

M. Legré, C. C. W. Lim, T. Lunghi, L. Monat, C. Portmann, M. Soucarros, P. Trinkler, G. Trolliet, F. Vannel, and H. Zbinden, New J. Phys. **16**, 013047 (2014).

[30] J. Bouda, M. Pivoluska, M. Plesch, and C. Wilmott, Phys. Rev. A **86**, 062308 (2012).

[31] E. O. Kiktenko, A. S. Trushechkin, C. C. W. Lim, Y. V. Kurochkin, and A. K. Fedorov, Phys. Rev. Appl. **8**, 044017 (2017).

[32] We note that this stage typically precedes the information reconciliation stage: the QBER value is estimated by random sampling from the sifted keys (of course, being publicly announced this sample is discarded from the sifted keys). However, recently proposed schemes suggest estimation of the QBER after the information reconciliation and verification stages (see Refs. [29,31,33]).

[33] E. O. Kiktenko, A. S. Trushechkin, Y. V. Kurochkin, and A. K. Fedorov, J. Phys: Conf. Ser. **741**, 012081 (2016).

[34] M. Christandl, A. Ekert, M. Horodecki, P. Horodecki, J. Oppenheim, and R. Renner, Lect. Notes Comput. Sci. **4392**, 456 (2007).

[35] W. Y. Hwang, I. G. Koh, and Y. D. Han, Phys. Lett. A **244**, 489 (1998); W.-Y. Hwang, X.-B. Wang, K. Matsumoto, J. Kim, and H.-W. Lee, Phys. Rev. A **67**, 012302 (2003).

[36] N. S. Aladov, Math. Sbor. **18**, 61 (1896).

[37] H. Davenport, J. London Math. Soc. **s1-8**, 46 (1933).

[38] B. Z. Moroz, Vestn. Leningr. Univ. Math. **16**, 164 (1961).

[39] R. Peralta, Math. Comput. **58**, 433 (1992).

[40] C. Ding, IEEE Trans. Inf. Theory **44**, 1693 (1998).

[41] I. Cziszár and J. Körner, IEEE Trans. Inf. Theory **24**, 339 (1978).

[42] C.-H. F. Fung, X. Ma, and H. F. Chau, Phys. Rev. A **81**, 012318 (2010).

[43] M. Dušek, M. Jahma, and N. Lütkenhaus, Phys. Rev. A **62**, 022306 (2000).

[44] N. Lütkenhaus and M. Jahma, New J. Phys. **4**, 44 (2002).

[45] H.-K. Lo, X. Ma, and K. Chen, Phys. Rev. Lett. **94**, 230504 (2005).

[46] X.-B. Wang, Phys. Rev. Lett. **94**, 230503 (2005).

[47] X. Ma, B. Qi, Y. Zhao, and H.-K. Lo, Phys. Rev. A **72**, 012326 (2005).

[48] Z. Zhang, Q. Zhao, M. Razavi, and X. Ma, Phys. Rev. A **95**, 012333 (2017).

[49] A. S. Trushechkin, E. O. Kiktenko, and A. K. Fedorov, Phys. Rev. A **96**, 022316 (2017).

[50] A. S. Holevo, *Quantum Systems, Channels, Information: A Mathematical Introduction* (De Gruyter, Berlin, 2012).

[51] A. Montanaro, Commun. Math. Phys. **273**, 619 (2007).

[52] W. Hoeffding, J. Am. Stat. Assoc. **58**, 13 (1963).