PHYSICAL REVIEW A **84**, 010302(R) (2011)

Semi-device-independent security of one-way quantum key distribution

Marcin Pawłowski[1,2] and Nicolas Brunner[3]

[1]*Department of Mathematics, University of Bristol, Bristol BS8 1TW, United Kingdom*
[2]*Institute of Theoretical Physics and Astrophysics, University of Gdansk, PL-80-952 Gdansk, Poland*
[3]*H. H. Wills Physics Laboratory, University of Bristol, Bristol BS8 1TL, United Kingdom*

(Received 29 March 2011; revised manuscript received 11 June 2011; published 15 July 2011)

By testing nonlocality, the security of entanglement-based quantum key distribution (QKD) can be enhanced to being "device-independent." Here we ask whether such a strong form of security could also be established for one-way (prepare and measure) QKD. While fully device-independent security is impossible, we show that security can be guaranteed against individual attacks in a semi-device-independent scenario. In the latter, the devices used by the trusted parties are noncharacterized, but the dimensionality of the quantum systems used in the protocol is assumed to be bounded. Our security proof relies on the analogies between one-way QKD, dimension witnesses, and random-access codes.

DOI: 10.1103/PhysRevA.84.010302 PACS number(s): 03.67.Dd, 03.67.Hk

The aim of quantum cryptography [1] is to warrant security against an eavesdropper solely limited by the laws of quantum mechanics. However, any quantum key distribution (QKD) scheme relies on an additional assumption which concerns information leakage out of the laboratories of Alice and Bob. Specifically, both parties must be free to choose which measurement they perform in each run of the protocol, and this choice of measurement, as well as the outcome of this measurement, should remain unknown to the eavesdropper. Indeed, if the eavesdropper has access to the laboratory of Alice or Bob, then security cannot be guaranteed.

Apart from these basic requirements, standard security proofs of QKD [2] also assume that Alice and Bob have excellent control on the quantum states and measurements used in the protocol. This assumption is, however, hard to justify in practice, where devices always feature some level of imperfection. Moreover, this assumption turns out to be crucial, as nicely illustrated in Ref. [3]. There it was shown that the security of the Bennett-Brassard (BB84) protocol [4] is entirely compromised if Alice and Bob use four-dimensional states instead of qubits, as usual security proofs always assume. It is, however, possible to avoid this requirement by basing the security on nonlocality. Specifically, by checking for the violation of a Bell inequality [5], Alice and Bob can ensure that they share nonlocal correlations, in which case security can be guaranteed without having any detailed knowledge on the functioning of the cryptographical devices [6,7]. This is "device-independent" (DI) QKD [8] (see also [9]).

The promise of a higher level of security, as well as the recently demonstrated attacks on actual QKD systems [10], have motivated research toward the practical implementation of DI-QKD. Despite recent progress [11], this remains a challenging problem. Moreover, the fact that DI-QKD is based on nonlocality strongly suggests that only entanglement-based protocols are suitable for obtaining this stronger notion of security. However, almost none of the practical QKD systems, in particular, none of the commercially available ones, use entanglement; they all operate in a one-way configuration, in which Alice prepares a quantum state and sends it to Bob, who then performs a measurement on it (hence, often called "prepare and measure").

Here we argue that a form of DI security—thus, stronger than usual security proofs—can nevertheless be obtained for QKD protocols which do not make use of entanglement. Specifically, we see that in a *semi-device-independent* scenario, in which the devices are noncharacterized but only assumed to produce quantum systems of a given dimension, security of one-way QKD against individual attacks can be demonstrated. In particular, our proof will make use of the analogy between one-way QKD protocols, dimension witnesses [12], and random-access codes [13]. Our work shows that security of QKD can be proven directly for the one-way configuration.

We start by presenting the semi-DI scenario we consider, stating clearly all assumptions we make. Then we consider the BB84 protocol and show that it becomes completely insecure in this context. This will also make clear that dimension witnesses are suitable tools for tackling this problem. Next we discuss the intimate relation existing between dimension witnesses and random-access codes [14]. Finally, we describe a specific QKD protocol and derive, via its associated dimension witness (or random-access code), a security proof.

Preliminaries. In a one-way QKD scheme Alice encodes classical information in a quantum system, which she sends to Bob via a quantum channel. Bob then performs a measurement on the system, from which he decodes some information. After repeating these operations many times, Alice and Bob estimate the error rate (by revealing randomly chosen bits from the raw key), which leads to an upper bound on Eve's information. Finally, Alice and Bob perform classical postprocessing—error correction, privacy amplification—to extract the sifted key on which Eve has arbitrarily small information.

Here we work in a semi-DI scenario. That is, we assume that the (relevant) Hilbert space dimension of the quantum systems is known,[1] but that the quantum preparations and measurements are noncharacterized. It will thus be convenient to describe the devices of Alice and Bob by black boxes.

[1] Here we mean the dimension of the relevant part of the Hilbert space, that is, the degrees of freedom that are correlated with the classical information encoded by Alice.

©2011 American Physical Society

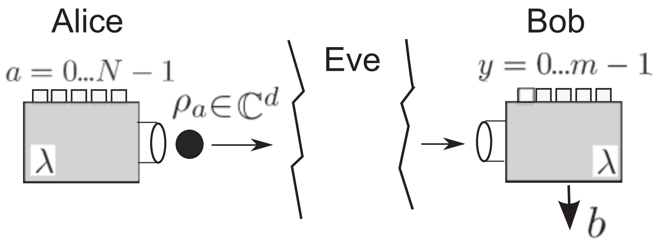

FIG. 1. Semi-device-independent one-way QKD.

Specifically Alice's black box is a "state preparator." Alice has the freedom to choose among a certain set of preparations $\rho_a \in \mathbb{C}^d$ with $a \in \{0,\ldots,N-1\}$, but knows nothing about these quantum states apart from their dimensionality d. We also assume that Alice's preparations ρ_a are unentangled from Eve; note that if Alice's preparations were entangled with Eve's system, then the communication capacity would be effectively doubled using dense coding. Bob's device is a measurement black box. He can choose to perform a measurement M_y with $y \in \{0,\ldots,m-1\}$ and gets the outcome $b \in \{0,\ldots,k-1\}$. The measurement operators M_y are noncharacterized; note that Eve could, in principle, send a system of arbitrary dimension to Bob. The boxes may also feature shared classical variables λ, known to Eve, but uncorrelated from the choice of preparation made by Alice and the choice of measurement made by Bob.

After repeating this procedure many times, Alice and Bob can estimate the probability distribution (or data table [15])

$$P(b|a,y) = \mathrm{tr}\big(\rho_a M_y^b\big), \tag{1}$$

which denotes the probability of Bob finding outcome b when he performed measurement M_y and Alice prepared ρ_a. Our goal will be to show that, in some cases, the security of a given protocol against a quantum eavesdropper can be guaranteed solely from its associated data table $P(b|a,y)$. The security is thus semi-DI in the sense that we do not require any knowledge on how the data table $P(b|a,y)$ was obtained, except from the fact that the device of Alice emits quantum systems of a given dimension.

Here we restrict ourselves to individual attacks, in which Eve attacks independently each system sent by Alice (using the same strategy) and measures her system before the classical postprocessing [1]. Indeed, we also need to make the basic assumption about information leakage from the devices. That is, no information about the inputs and output (i.e., a, y, and b) leaks out of the boxes to Eve.

Dimension witnesses. At this point one can already see a first requirement for obtaining semi-DI security for a given protocol. Suppose Alice's device prepares d-dimensional quantum systems. Then it must be impossible to reproduce the quantum data table with classical systems of dimension d. If not, then it could have been the case that Alice's device emits orthogonal quantum states (or equivalently classical states) from which Eve can get full information. Thus, full DI security, that is, where no assumption is made on the Hilbert space dimension, is impossible, since every data table can be reproduced using classical systems of sufficient dimension.

It turns out that a simple method for establishing lower bounds on the dimension of classical systems necessary to reproduce a given data table was recently developed in Ref. [12]. More precisely, the authors devised "dimension witnesses" of the form

$$\sum_{a,y,b} w_{aby} P(b|a,y) \leqslant C_d, \tag{2}$$

which can be thought of as Bell-type inequalities for data tables. Here the bound C_d denotes the maximal value of the left-hand-side polynomial obtainable when Alice's device prepares classical d-dimensional systems. Interestingly, d-dimensional classical dimension witnesses can be violated by d-dimensional quantum systems, thus indicating that certain quantum data tables cannot be reproduced using classical states of the same dimension. Below, we make use of this "quantum advantage." We consider a simple dimension witness which provides a separation between qubits and bits. In particular, we show how this witness can be naturally understood as a random-access code, which allows us to prove semi-DI security of the corresponding QKD protocol.

From now we now focus on the case where Alice's device prepares two-dimensional quantum systems and restrict to four preparations ($N = 4$) indexed by two bits a_0,a_1. Bob's device can perform two binary measurements ($m = 2$, $k = 2$). For the rest of the Rapid Communication it is convenient to use expectation values of the form

$$E_{a_0a_1,y} = P(b = 0|a_0a_1,y). \tag{3}$$

Thus, every experiment corresponds to a data table, given by a vector $\vec{E} = \{E_{a_0a_1,y}\}_{a_0a_1,y}$ of $Nm = 8$ correlators.

First, we would like to characterize the set of data tables (i.e., the set of vectors $\vec{E}$) which can be obtained when Alice's box emits classical bits. We follow the geometrical methods of Ref. [12]. The set of interest to us is a polytope (in an eight-dimensional space). Its facets are (tight) two-dimensional classical witnesses; that is, inequalities of the form (2) with $d = 2$ (note that here probabilities are simply replaced by correlators). It turns out that there are only two types of witnesses in the case. The first is a straightforward extension[2] of the witness I_3 of Ref. [12]. The second is of the following form:

$$\begin{aligned} S = &+E_{00,0} + E_{00,1} + E_{01,0} - E_{01,1} \\ &-E_{10,0} + E_{10,1} - E_{11,0} - E_{11,1} \leqslant 2. \end{aligned} \tag{4}$$

This witness will be our main tool to assess the security of one-way QKD protocols.

BB84 is not secure. As a warmup, it is instructive to consider first the case of the BB84 protocol. In this case, the four preparations of Alice are given by

$$\begin{aligned} \rho_{00} &= |0\rangle\langle 0|, \quad \rho_{11} = |1\rangle\langle 1|, \\ \rho_{10} &= |+\rangle\langle +|, \quad \rho_{01} = |-\rangle\langle -|. \end{aligned} \tag{5}$$

Here $|0\rangle$ and $|1\rangle$ are the eigenstates of the Pauli matrix σ_z, and $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ are the eigenstates of σ_x. The two measurements of Bob are given by $M_0 = \sigma_z$ and $M_1 = \sigma_x$.

[2] Note that the witness I_3 of [12] uses only $N = 3$ preparations. Here we obtain the lifting of I_3 to the case $N = 4$; the fourth preparation does simply not appear into the expression.

Thus, the corresponding data table is given by $E_{00,0} = \mathrm{tr}(\rho_{00}\sigma_z) = 1$, and similarly $E_{10,1} = 1$, $E_{01,1} = E_{11,0} = 0$, and $E_{00,1} = E_{01,0} = E_{10,0} = E_{11,1} = 1/2$. Thus, the BB84 data table achieves $S = 2$ and thus does not violate the witness (4)—note that it also satisfies the witness I_3 of [12] as well as all symmetries—which indicates that it can be reproduced by sending one classical bit when the boxes of Alice and Bob share randomness. Note that this is a peculiarity of the BB84 data table.[3] Indeed, this result also applies to any protocol using the same states and measurements as BB84, for instance, the Scarani-Acin-Ribordy-Gisin (SARG) protocol [16].

A possible strategy is the following. Alice and Bob share one random bit λ. Considering Alice's preparations, note that the bit $a_0 \oplus a_1$ denotes the basis, while the bit a_1 denotes the encoded bit. When $\lambda = 0$, Alice sends to Bob the (one-bit) message $m = a_0$. Bob, upon getting his input y and the message from Alice m, outputs $b = m \oplus y = a_0 \oplus y$. Thus, $b = a_1$ whenever Alice and Bob choose the same basis ($a_0 \oplus a_1 = y$), and $b \neq a_1$ when they choose a different basis ($a_0 \oplus a_1 \neq y$). When $\lambda = 1$, Alice sends the message $m = a_1$, and Bob outputs $b = m = a_1$. Thus, we have that $a_1 = b$ for any pair of basis $a_0 \oplus a_1, y$. Since the shared variable λ is unbiased, Alice and Bob reproduce the BB84 data table.

Connection to random-access codes. To devise a secure QKD protocol in the semi-DI setup, we need to consider data tables which violate (at least) one of the dimension witness I_3 or S. Here we focus on the latter, which is useful to think of in terms of a random-access code.

Specifically, let us imagine that Alice receives two (uniformly distributed) bits a_0 and a_1. She is then allowed to send a physical system to Bob, which encodes information about her input bit string. Bob is asked to guess the yth bit of Alice (y is uniformly distributed as well), and thus performs a measurement on the system he received from Alice to extract this information. This is a 2-to-1 random-access code. When Alice sends one bit of classical communication, the optimal average probability for Bob to succeed is 3/4 [13].

The witness S (4) represents a 2-to-1 random-access code. For each of her four possible input bit strings $\{a_0,a_1\}$, Alice associates a preparation $\rho_{a_0a_1}$. Upon being asked to guess bit y, Bob performs measurement M_y. The outcome of the measurement b is then his guess for a_y.

From inspection of S, we see that $w_{a_0a_1,y} = (-1)^{a_y}$ (where $w_{a_0a_1,y}$ is the coefficient of the term $E_{a_0a_1,y}$), which implies that

$$S = \sum_{a_0,a_1,y} P(b = a_y|a_0a_1,y) - 4. \tag{6}$$

Thus, for a given data table, Bob's success probability,

$$P_B = \frac{1}{8}\sum_{a_0,a_1,y} P(b = a_y|a_0a_1,y) = \frac{S+4}{8}, \tag{7}$$

is determined by the value of the dimension witness S, and inversely. Indeed, the inequality $P_B \leqslant 3/4$ corresponds to $S \leqslant 2$. Note that the relation between dimension witnesses and random-access codes can be generalized (see also [14] for a related approach).

It turns out that Alice and Bob can perform better at this task when using qubits. The optimal set of preparations are, for instance, obtained by having preparations (5), but changing Bob's measurements to $M_0 = (\sigma_z + \sigma_x)/\sqrt{2}$ and $M_1 = (\sigma_z - \sigma_x)/\sqrt{2}$. This choice of preparations and measurements leads to $S = 8\cos^2(\pi/8) - 4$ or, equivalently,

$$P_B = \cos^2(\pi/8) \approx 0.8536. \tag{8}$$

Note that this set of preparations and measurements is intimately related to the Clauser-Horne-Shimony-Holt Bell inequality (see also [14]).

Security of one-way QKD. The protocol is based on the preparations and measurements achieving the optimal violation of S for qubits. Alice generates two random bits a_0,a_1 and sends the corresponding preparations $\rho_{a_0a_1}$ to Bob. Bob generates a random bit y and performs measurement M_y and guesses bit a_y. After repeating these operations a large number of times (we consider here only the asymptotic limit), Alice and Bob can estimate the data table by revealing part of their data on a public channel. By computing the value of S they obtain P_B. Below we show that if $P_B > \frac{5+\sqrt{3}}{8} \approx 0.8415$—a value slightly lower than the optimal value using qubits (8)—security is obtained.

Proof. Csiszar and Körner [17] showed that Alice and Bob can obtain a secret key if $I(A:B) > I(A:E)$, where the mutual information is given by

$$I(A:X) = \sum_j 1 - h(P_X(a_{y_j})). \tag{9}$$

Here y_j denotes the choice of basis (or, equivalently, which bit of Alice party X chose to guess) in the jth run of the protocol, and $h(p)$ is the Shannon binary entropy. From this, one can get a sufficient condition for security given by

$$P_B > P_E, \tag{10}$$

where $P_X = \frac{1}{2}[P_X(a_0) + P_X(a_1)]$ denotes the average probability of guessing correctly for party X.

Our main ingredient is a result derived by König [18]. Consider the set F_n of all (Boolean) balanced functions on n-bit strings, that is, which return 0 for exactly half of the 2^n strings. Alice gets as input the n-bit string and Bob is asked to guess the value of a randomly (and uniformly) chosen function in F_n after receiving from Alice s qubits. Then the average probability for Bob to succeed is upper bounded as follows:

$$P_n \leqslant \frac{1}{2}\left(1 + \sqrt{\frac{2^s - 1}{2^n - 1}}\right). \tag{11}$$

For the case of interest to us, that is, $n = 2$ bits, the set of all balanced functions is a_0, a_1, $a_0 \oplus a_1$, and their negations. Clearly, the optimal probability of guessing a function or its negation are equal. Thus, when Alice sends a single qubit to Bob ($s = 1$), we have that

$$P_B(a_0) + P_B(a_1) + P_B(a_0 \oplus a_1) \leqslant \frac{3}{2}\left(1 + \frac{1}{\sqrt{3}}\right). \tag{12}$$

[3] It turns out that any deviation from the BB84 data table, obtained from modifying one (or more) preparations and/or one measurement, will lead to a value of $S > 2$. Thus, BB84 appears to be worst case in this context.

Clearly, the previous inequality holds also when Bob and Eve collaborate—the index B is then simply replaced by BE—and we make use of it in this case. Using the relations $P_{BE}(a_i) \geqslant P_B(a_i)$ and $P_{BE}(a_i) \geqslant P_E(a_i)$ and

$$\begin{aligned} P_{BE}(a_0 \oplus a_1) &\geqslant P_{BE}(a_0,a_1) \\ &\geqslant P_{BE}(a_0) + P_{BE}(a_1) - 1, \end{aligned} \tag{13}$$

where the second inequality follows from the sum rule, we get

$$\begin{aligned} P_{BE}(a_0) + P_{BE}(a_1) + P_{BE}(a_0 \oplus a_1) \\ \geqslant 2P_B(a_0) + 2P_E(a_1) - 1. \end{aligned} \tag{14}$$

Using (12) we get that

$$P_B(a_0) + P_E(a_1) \leqslant \frac{5+\sqrt{3}}{4} \tag{15}$$

and an analogous inequality with a_0 and a_1 interchanged. This shows that when Eve tries to guess a different bit than Bob (i.e., she measures in the wrong basis) she will necessarily disturb the statistics of Bob. From inequality (15) and its symmetry with respect to a_0 and a_1, we get that

$$P_B + P_E \leqslant \frac{5+\sqrt{3}}{4}. \tag{16}$$

This implies that $P_B > P_E$ as long as

$$P_B > \frac{5+\sqrt{3}}{8} \approx 0.8415, \tag{17}$$

as announced. For the optimal qubit preparations and measurements achieving (8), the key rate is found to be

$$r = I(A:B) - I(A:E) \approx 0.0581. \qquad \blacksquare$$

Discussion. We have discussed the security of one-way QKD in a semi-device-independent context. By making links to dimension witnesses and random-access codes, we showed that security against individual attacks is possible.

It is natural to ask whether this concept is relevant from a practical viewpoint. Since semi-DI QKD represents a relaxation of the assumption of standard QKD proofs, it offers several advantages, notably that no assumptions on the devices are required (apart from the fact that Alice's device emits preparations of bounded dimension), and that it can be applied directly to the one-way configuration. At this stage, our result should, however, be understood as a proof of principle. A next step would be to study robustness to imperfections (such as losses or detection efficiency) as well as against more general attacks. It would also be interesting to improve on our bound for security (which is likely to be suboptimal) and to see whether all data tables violating a classical dimension witness could offer security. In this context it might also be relevant to consider entropic quantities [14,19].

A comparison to full DI QKD is also worthwhile. Arguably, the main drawback of our approach is the assumption of bounded dimensionality, as it forces us to assume that Alice's device features no side channels from which Eve could extract information. This requirement could, however, be partly lifted by finding protocols where qubits offer security under the assumption that the preparations are arbitrary quantum states of higher dimensions—note that this would require protocols using more preparations.

Finally, from a more foundational point of view, it would be interesting to study the connection between semi-DI one-way QKD and DI entanglement-based QKD in the light of the strong link that exists between nonlocality and random-access codes [20].

Acknowledgments. M.P. thanks Rob Spekkens and Marek Żukowski for useful discussions. This work was supported by the UK EPSRC, EU QESSENCE, and the Leverhulme Trust.

[1] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dusek, N. Lütkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).

[2] P. W. Shor and J. Preskill, Phys. Rev. Lett. **85**, 441 (2000); D. Mayers, JACM **48**, 351 (2001).

[3] A. Acin, N. Gisin, and L. Masanes, Phys. Rev. Lett. **97**, 120405 (2006).

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

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

[6] D. Mayers and A. Yao, Quantum Inf. Comput. **4**, 273 (2004).

[7] J. Barrett, L. Hardy, and A. Kent, Phys. Rev. Lett. **95**, 010503 (2005).

[8] A. Acin, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007); S. Pironio, A. Acín, N. Brunner, N. Gisin, S. Massar, and V. Scarani, New J. Phys. **11**, 045021 (2009).

[9] L. Masanes, S. Pironio, and A. Acin, Nat. Commun. **2**, 238 (2011); E. Hänggi and R. Renner, e-print arXiv:1009.1833; M. McKague, Ph.D. thesis, e-print arXiv:1006.2352.

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

[11] N. Gisin, S. Pironio, and N. Sangouard, Phys. Rev. Lett. **105**, 070501 (2010).

[12] R. Gallego, N. Brunner, C. Hadley, and A. Acin, Phys. Rev. Lett. **105**, 230501 (2010).

[13] A. Nayak, in *Proceedings of 40th IEEE FOCS* (1999), pp. 369–376.

[14] S. Wehner, M. Christandl, and A. C. Doherty, Phys. Rev. A **78**, 062112 (2008).

[15] N. Harrigan, T. Rudolph, and S. Aaronson, e-print arXiv:0709.1149.

[16] V. Scarani, A. Acin, G. Ribordy, and N. Gisin, Phys. Rev. Lett. **92**, 057901 (2004).

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

[18] R. König, Ph.D. thesis, University of Cambridge (see Corrolary 5.2.3).

[19] M. Tomamichel and R. Renner, Phys. Rev. Lett. **106**, 110506 (2011).

[20] M. Pawlowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Zukowski, Nature (London) **461**, 1101 (2009); J. Oppenheim and S. Wehner, Science **330**, 1072 (2010).