Small Accessible Quantum Information Does Not Imply Security

Robert König,[1] Renato Renner,[1] Andor Bariska,[2] and Ueli Maurer[2]
[1]*Centre for Quantum Computation, University of Cambridge, United Kingdom*
[2]*Institute of Theoretical Computer Science, ETH Zurich, Switzerland*
(Received 13 December 2005; published 3 April 2007)

The security of quantum key distribution is typically defined in terms of the mutual information between the distributed key S and the outcome of an optimal measurement applied to the adversary's system. We show that even if this so-called *accessible information* is small, the key S might not be secure enough to be used in applications such as one-time pad encryption. This flaw is due to a *locking* property of the accessible information: one additional (physical) bit of information can increase the accessible information by more than one bit.

DOI: 10.1103/PhysRevLett.98.140502 PACS numbers: 03.67.Dd

Secret keys are the basis for various cryptographic tasks such as message encryption or authentication. In order to guarantee security of these tasks, it is hence crucial that the underlying key provides a sufficiently high *level of security*. The strongest and thus most desirable notion of security for a secret key S is called *perfect security* and is characterized by two conditions: (i) any value of S is equally likely (i.e., the distribution P_S of S is uniform on a key space $\mathcal{S}$) and (ii) an adversary has no information about S (i.e., the state of any system controlled by an adversary is independent of the value of S).

Using such a perfectly secure key enables the realization of highly secure cryptographic schemes. For example, S could be used for one-time pad encryption [1], where the ciphertext C of a message M is obtained by computing the bitwise addition (modulo 2) of M and S. It is easy to verify that, from the viewpoint of an adversary who does not know S, the ciphertext C is independent of the message M and thus completely useless.

Unfortunately, it is generally impossible to generate perfectly secure keys—even with the help of quantum mechanics. For this reason, one usually considers slightly weakened security definitions. For example, condition (ii) might be substituted by a bound on the information that the adversary has about S. This, however, raises a basic question: what is an appropriate measure to quantify the adversary's information about S?

In the context of classical information-theoretic cryptography [2], the adversary's knowledge about a key S is most generally characterized by a classical random variable Z. An n-bit key S is then said to be secure [4] if, for some small security parameter $\varepsilon \geq 0$,

$$H(S) \geq n - \varepsilon \tag{1}$$

$$I(S;Z) \leq \varepsilon, \tag{2}$$

where $H(S)$ denotes the *Shannon entropy* of S and $I(S;Z) := H(S) - H(S|Z)$ is the *mutual information between* S and Z. Inequality (1) implies that S is almost uniformly distributed; it is thus an approximation of condition (i) above. Similarly, (2) is an approximation of (ii).

In *quantum cryptography*, the knowledge of an adversary about a (classical) key S is described by the state of a quantum system E instead of a classical random variable Z. Accordingly, the mutual information occurring in criterion (2) is usually generalized to the accessible information $I_{\mathrm{acc}}(S;E)$, which is defined as the mutual information between S and the outcome Z of an optimal measurement applied to E (see below for a formal definition). The quantum version of (2) then reads

$$I_{\mathrm{acc}}(S;E) \leq \varepsilon. \tag{2'}$$

Inequality (2′) seems to be a natural formalization of the requirement that an adversary has almost no information about S and is in fact commonly used in the standard literature on quantum cryptography and, in particular, quantum key distribution (QKD) [8].

In this Letter, we show that the accessible information has the following property.

Proposition 1.—For any $\varepsilon > 0$ there exists a quantum state on a system E which depends on a classical random variable V and an n-bit string W, where n is linear in $\log 1/\varepsilon$, such that the following holds: (a) $I_{\mathrm{acc}}(V;E) \leq \varepsilon$ and (b) $I_{\mathrm{acc}}(V;WE) \geq H(W) + 1$.

This result implies that the accessible information is lockable [17]; i.e., $H(W)$ bits of additional information (encoded in W) might increase the accessible information by more than $H(W)$ bits. We use the locking property given in Proposition 1 to argue that the security definition for QKD described above has a flaw which is relevant for practical applications.

Proposition 2.—There exists a setting consisting of a classical key S and a quantum system E controlled by an adversary such that security criterion (2′) is satisfied (for some security parameter ε which is exponentially small in the key size), but S cannot securely be used for one-time pad encryption.

In the sequel, we give an explicit construction of a quantum state that proves Proposition 1. We then show that Proposition 2 follows from a simple argument based on this construction. Finally, in the last part of this Letter, we discuss an alternative security definition which overcomes the problem exposed by Proposition 2. Moreover, we argue that a key satisfying this alternative definition can be obtained from standard QKD protocols (by modifying the security proofs).

Consider a ''hybrid'' setting consisting of both a classical random variable V and a quantum system E. This setting can be described by the so-called enlarged Hilbert space representation, where one thinks of the classical value V as being encoded into the state of a quantum system with respect to an orthonormal basis $\{|v|\}_{v\in\nu}$ as follows:

$$\rho_{VE} := \sum_{v\in\nu} P_V(v)|v\rangle\langle v| \otimes \rho_{E|V=v},$$

where $\rho_{E|V=v}$ is the state of E conditioned on $V = v$. We will refer to a density operator of this form as a cq state. We will also use generalizations of this convention to tripartite systems with two classical parts and call the corresponding density operators ccq states.

For any cq state ρ_{VE}, the accessible information (of E about V) is defined as [18]

$$I_{\mathrm{acc}}(V;E) := \max_{\mathcal{M}} I(V;\mathcal{M}[E]),$$

where the maximum is over all local measurements given by a positive operator-valued measure $\mathcal{M}$ on E and where $I(V;\mathcal{M}[E])$ denotes the mutual information between V and the measurement outcome $\mathcal{M}[E]$. The accessible information $I_{\mathrm{acc}}(V;E)$ thus quantifies the amount of information about the classical value V that can be obtained by an optimal measurement applied to the quantum system E.

Consider now an extended setting involving an additional random variable W, described by a ccq state ρ_{VWE}. Let [19]

$$\Delta := I_{\mathrm{acc}}(V;WE) - I_{\mathrm{acc}}(V;E)$$

be the amount by which the accessible information about V increases when W is appended to E. Hence, given access to the quantum system E, the quantity Δ measures by how much the knowledge about V increases if one learns W. Saying that I_{acc} is lockable [17] then means that Δ can generally be larger than the size of the so-called unlocking information W, i.e., the number of bits which are needed to represent its value. It should be emphasized that locking is a purely nonclassical property. In fact, if the quantum system E is substituted by a classical random variable Z, we have $\Delta = I(V;W|Z) \leq H(W)$ [20], that is, Δ cannot be larger than the size of W.

Let us now describe a ccq state ρ_{VWE} which satisfies properties (a) and (b) of Proposition 1 and, hence, proves the proposition. It is motivated by previous constructions [17,21,22], where a random variable X is encoded into a quantum system E using a basis randomly chosen from a family of bases [23]. The unlocking information W specifies the basis. Given this information, the value of $V = (X, W)$ can be obtained by a measurement of E. Intuitively, the locking effect arises because it is impossible to perform the ''right'' measurement without knowing the basis in which X is encoded. More formally, it is shown that the accessible information $I_{\mathrm{acc}}(XW;E)$ is limited. This is achieved by replacing the minimization over POVMs in its definition by a minimization over states, which establishes a link to so-called entropy uncertainty relations.

We use the following notational conventions: $\sigma_1, \sigma_2, \sigma_3$ are the Pauli matrices on the Hilbert space $\mathbb{C}^2$. For any m-tuple $w = (w_1, \ldots, w_m) \in \{1, 2, 3\}^m$, we denote the m-fold tensor product $\sigma_{w_1} \otimes \cdots \otimes \sigma_{w_m}$ by σ_w. Let X and W be random variables on the set $\mathcal{X} := \{0, 1\}$ and the set of m-tuples $\mathcal{W} := \{1, 2, 3\}^m$, respectively, such that the joint probability distribution P_{XW} is uniform. Then, for any $x \in \mathcal{X}$ and $w \in \mathcal{W}$, we define

$$\rho_{E|(X,W)=(x,w)} := 2^{-m}(\mathrm{id}_{(\mathbb{C}^2)^{\otimes m}} + (-1)^x \sigma_w), \qquad (3)$$

which is an operator on $(\mathbb{C}^2)^{\otimes m}$ representing the state of a quantum system E conditioned on $X = x$ and $W = w$. It is straightforward to check that this is a consistent description of a ccq state ρ_{XWE} [24]. The relevant properties of this state are expressed by the following.

Lemma 1.—Let ρ_{XWE} be as described. Then (i) for any fixed value $w \in \mathcal{W}$ of the random variable W, there exists a measurement of the quantum system E with output equal to X. In particular, $I_{\mathrm{acc}}(XW;WE) = H(XW)$. (ii) $I_{\mathrm{acc}}(XW;E) \leq (\frac{2}{3})^{m/2}$.

Proof.—Property (i) directly follows from the fact that the conditional states $\rho_{E|(X,W)=(0,w)}$ and $\rho_{E|(X,W)=(1,w)}$ are orthogonal for any fixed $w \in \mathcal{W}$. To prove (ii), we show that for any measurement $\mathcal{M}$ applied to the quantum part E of ρ_{XWE}, the entropy of the pair (X, W) conditioned on the outcome $\mathcal{M}[E]$ is bounded by

$$H(XW|\mathcal{M}[E]) \geq H(XW) - (\tfrac{2}{3})^{m/2}. \qquad (4)$$

The assertion then follows because $I_{\mathrm{acc}}(XW;E) = H(XW) - \min_{\mathcal{M}} H(XW|\mathcal{M}[E])$.

The family $\mathcal{N}_w := \{2^m P_{X|W=w}(x)\rho_{E|(X,W)=(x,w)}\}_{x\in\mathcal{X}}$ is a positive operator-valued measure for every $w \in \mathcal{W}$ because $\rho_{WE} = \rho_W \otimes \rho_E$, where ρ_E is the fully mixed state. As in [17], it can be shown that the conditional entropy of interest can be bounded in terms of a sum of entropies minimized over all states σ on $(\mathbb{C}^2)^{\otimes m}$, i.e.,

$$H(XW|\mathcal{M}[E]) \geq H(W) + \min_{\sigma} \sum_{w\in\mathcal{W}} P_W(w) H(\mathcal{N}_w[\sigma]).$$

Lower bounds on the expression on the right-hand side are called entropy uncertainty relations as they express the average uncertainty about the outcome $\mathcal{N}_w[\sigma]$ when measuring a state σ using different measurements $\{\mathcal{N}_w\}_{w\in\mathcal{W}}$. Using the fact that the POVMs are binary valued, it can be

shown [25] that in our case, this average uncertainty is at least $1 - (2/3)^{m/2}$, which implies (4) as desired. □

Note that W can be represented by a string of length roughly $H(W)$. Proposition 1 therefore follows directly from Lemma 1 by setting $V := (X, W)$ and choosing m such that $(2/3)^{m/2} \leq \varepsilon$.

To prove Proposition 2, we similarly represent W by n bits $S_1, \ldots, S_n$ (for $n \approx m\log_2 3$). We then consider the $(n+1)$-bit key $S = (S_1, \ldots, S_{n+1})$ consisting of these n bits and, additionally, the bit $S_{n+1} := X$ [26]. Moreover, we assume that the adversary controls the system E.

It is an immediate consequence of Lemma 1 (ii) that the key S satisfies the security criterion (2′), i.e.,

$$I_{\mathrm{acc}}(S; E) = I_{\mathrm{acc}}(XW; E) \leq \varepsilon, \tag{5}$$

where $\varepsilon := 2^{-(n-2/6)}$ decreases exponentially fast in the key length. However, as illustrated by the following example, this is not sufficient if S is used for one-time pad encryption.

Let $M = (M_1, \ldots, M_{n+1})$ be an $(n+1)$-bit message and let $C = (C_1, \ldots, C_{n+1})$ be the ciphertext given by $C_i \equiv M_i + S_i \mathrm{mod} 2$. Moreover, assume that an adversary has some *a priori* knowledge which determines the first n message bits $M_1, \ldots, M_n$, but does not know the bit M_{n+1} [27]. Upon receiving the ciphertext bits $C_1, \ldots, C_n$, the adversary can thus immediately infer the first n key bits $S_1, \ldots, S_n$. Hence, by Lemma 1 (i), the adversary now is in a position to choose an appropriate measurement of her quantum system E which reveals the $(n+1)$th key bit S_{n+1} with certainty. The encryption of the $(n+1)$th message bit M_{n+1} is thus completely insecure. This concludes the proof of Proposition 2.

According to this result, defining secrecy in terms of the accessible information is problematic in a quantum world. This raises the question whether there are stronger security definitions which, e.g., imply that a secret key can safely be used for one-time pad encryption. As shown recently [15,16,29], the answer to this question is positive [30].

Let S be a classical key which takes values from $\mathcal{S}$ and let E be a quantum system controlled by an adversary. As described above, this situation can be represented by a cq state $\rho_{SE} = \sum_{s\in\mathcal{S}} P_S(s)|s\rangle\langle s| \otimes \rho_{E|S=s}$.

Definition 1.([16,29,31])—A random variable S is called an ε-secure key with respect to E if [32]

$$\tfrac{1}{2}\|\rho_{SE} - \rho_U \otimes \rho_E\|_1 \leq \varepsilon,$$

where $\rho_U := \sum_{s\in\mathcal{S}} \frac{1}{|\mathcal{S}|}|s\rangle\langle s|$ is the completely mixed state.

As discussed in [16,29], ε security has an intuitive interpretation: with probability at least $1 - \varepsilon$, the key S can be considered identical to a perfectly secure key U, i.e., U is uniformly distributed and independent of the adversary's information. In other words, Definition 1 guarantees that the key S is perfectly secure except with probability ε. Because this is still true if S is used in any application, the above definition is said to be composable.

Interestingly, composable security can be achieved quite easily. For example, it has been shown [16] that the key computed by applying a two-universal hash function to a random string with sufficient entropy satisfies Definition 1 [33]. Security proofs of QKD which are based on this result (see, e.g., [29,34]) are thus not affected by the problem discussed above.

Strongly secure keys can also be obtained by measuring predistributed Bell states $|\Phi^+\rangle$ (or approximations thereof), as first observed in [15] (see [35]).

Lemma 2.—Let $\varepsilon \geq 0$ and let ρ_{AB} be a bipartite quantum state such that $F(\rho_{AB}, |\Phi^+\rangle^{\otimes n}) \geq \sqrt{1-\varepsilon^2}$. Then the two n-bit strings resulting from local measurements of ρ_{AB} in the computational basis are ε-secure keys (with respect to an adversary who controls a purifying system of ρ_{AB}).

The statement implies that security proofs based on entanglement purification (where the entanglement is usually measured in terms of the fidelity $F(\rho_{AB}, |\Phi^+\rangle^{\otimes n})$ to a fully entangled state, as, e.g., in [9,10]) can easily be adapted to meet Definition 1 [36].

In conclusion, we have presented a novel example of the lockability of classical correlations in quantum states. This example reveals a weakness of security definitions based on the accessible information. In particular, a secret key which is secure according to such a definition might become completely insecure when used in certain applications. A solution to this problem is provided by the stronger yet still achievable notion of ε security (Definition 1) which is composable. An ε-secure key can safely be used in any application except with some (small) probability ε. Our result thus does not imply that actual QKD schemes are insecure, but shows that statements about their security (as well as corresponding proofs) that are phrased in terms of the accessible information are not satisfactory.

This work has been supported by the EC under projects PROSECCO of the IST-FET programme, SECOQC, QAP, SCALA, and by HP Labs Bristol.

[1] G. S. Vernam, J. Am. Inst. Electr. Eng. **45**, 109 (1926).
[2] For an introduction to classical information-theoretic key agreement, see, e.g., [3].
[3] U. M. Maurer, IEEE Trans. Inf. Theory **39**, 733 (1993).
[4] For classical security definitions, see, e.g., [3,5–7].
[5] A. D. Wyner, Bell Syst. Tech. J. **54**, 1355 (1975).
[6] C. H. Bennett, G. Brassard, and J.-M. Robert, SIAM J. Comput. **17**, 210 (1988).
[7] I. Csiszár and J. Körner, IEEE Trans. Inf. Theory **24**, 339 (1978).
[8] See, e.g., [9–14] and also the discussion in [15,16].
[9] H.-K. Lo and H. F. Chau, Science **283**, 2050 (1999).
[10] P. W. Shor and J. Preskill, Phys. Rev. Lett. **85**, 441 (2000).
[11] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
[12] D. Gottesman and H.-K. Lo, IEEE Trans. Inf. Theory **49**, 457 (2003).

[13] H.-K. Lo, H. F. Chau, and M. Ardehali, J. Cryptology **18**, 133 (2005).

[14] E. Biham, M. Boyer, P. O. Boykin, T. Mor, and V. Roychowdhury, J. Cryptology **19**, 381 (2006).

[15] M. Ben-Or, M. Horodecki, D. W. Leung, D. Mayers, and J. Oppenheim, *Second Theory of Cryptography Conference TCC*, Lecture Notes in Computer Science Vol. 3378 (Springer, New York, 2005), pp. 386–406.

[16] R. Renner and R. König, *Second Theory of Cryptography Conference TCC*, Lecture Notes in Computer Science Vol. 3378 (Springer, New York, 2005), pp. 407–425.

[17] D. P. DiVincenzo, M. Horodecki, D. W. Leung, J. A. Smolin, and B. M. Terhal, Phys. Rev. Lett. **92**, 067902 (2004).

[18] In the literature, the accessible information is often defined in terms of ensembles. It is easy to verify that such a definition is equivalent to the one given here.

[19] $I_{\mathrm{acc}}(V;WE)$ denotes the accessible information of the cq state $\rho_{V(WE)}$ which is obtained from ρ_{VWE} by combining the systems W and E.

[20] $I(V;W|Z) := H(V|Z) + H(W|Z) - H(VW|Z)$ is the mutual information between V and W given Z.

[21] P. Hayden, D. Leung, P. W. Shor, and A. Winter, Commun. Math. Phys. **250**, 371 (2004).

[22] H. Buhrman, M. Christandl, P. Hayden, H. W. Lo, and S. Wehner, Phys. Rev. Lett. **97**, 250501 (2006).

[23] We cannot directly use the constructions [17,21,22] because they do not satisfy property (a) of Proposition 1. This property is, however, needed for our considerations related to cryptography.

[24] An alternative description of the state ρ_{XWE} which clarifies the relation to the construction of [17] is the following. For $w \in \{1,2,3\}$, let $\{[0]_w, [1]_w\}$ denote the projectors onto the eigenspaces of σ_w. Let $R = (R_1, \ldots, R_m)$ and $W = (W_1, \ldots, W_m)$ be independent and uniformly distributed random variables on $\{0,1\}^m$ and $\{1,2,3\}^m$. For $r \in \{0,1\}^m$ and $w \in \{1,2,3\}^m$, let $\rho_{E|(R,W)=(r,w)} := [r_1]_{w_1} \otimes \cdots \otimes [r_m]_{w_m}$. Finally, let X be the random variable on $\mathcal{X}$ defined by $X := \sum_{i=1}^m R_i \mathrm{mod} 2$. It is then straightforward to check that the resulting conditional states $\rho_{E|(X,W)=(x,w)}$ are given by (3). Note that the locking scheme described in [17] uses the conditional states $\rho_{E|(X,W)=(x,w)} = [x]_w$ on $\mathbb{C}^2$, with W uniformly distributed on $\{1,2,3\}$. Our construction therefore individually encodes each bit of the string R (whose parity X we are interested in) using the locking scheme of [17].

[25] More precisely, we apply the bound $h(p) \geq 1 - |p - (1-p)|$ for $p \in [0,1]$ on the binary entropy function $h(p) = -p\log_2 p - (1-p)\log_2(1-p)$. With $|\mathrm{Tr}[\sigma(\rho_{E|(X,W)=(0,w)} - \rho_{E|(X,W)=(1,w)})]| = 2^{-m+1}|\mathrm{Tr}(\sigma_w\sigma)|$ for every $w \in \mathcal{W}$ and every state σ, we obtain $\sum_{w\in\mathcal{W}} P_W(w) H(\mathcal{N}_w[\sigma]) \geq 1 - \frac{1}{|\mathcal{W}|}\sum_{w\in\mathcal{W}} |\mathrm{Tr}(\sigma_w\sigma)|$. The latter term can be bounded by use of the Cauchy-Schwarz inequality, i.e., $\frac{1}{|\mathcal{W}|}\sum_{w\in\mathcal{W}} |\mathrm{Tr}(\sigma_w\sigma)| \leq \frac{1}{\sqrt{|\mathcal{W}|}} \times \sqrt{\sum_{w\in\mathcal{W}} \mathrm{Tr}(\sigma_w\sigma)^2}$ and the fact that $\sum_{w\in\mathcal{W}} \mathrm{Tr}(\sigma_w\sigma)^2 \leq 2^m$ for every state σ since $\mathrm{Tr}(\sigma^2) \leq 1$. We thus get $\sum_{w\in\mathcal{W}} P_W(w) H(\mathcal{N}_w[\sigma]) \geq 1 - (2/3)^{m/2}$ for all states σ on $(\mathbb{C}^2)^{\otimes m}$.

[26] As we will consider one-time pad encryption with the key S, we assume for simplicity that S is a bit string. Because the cardinality of the range of (X, W) (i.e., $\{0,1\} \times \{1,2,3\}^m$) and S (i.e., $\{0,1\}^{n+1}$) do not match exactly, S is not perfectly uniformly distributed on the key space. However, a qualitatively identical statement with a perfectly uniformly distributed key can be obtained by adapting the one-time pad to keys and messages on the space $\mathcal{X} \times \mathcal{W} = \{0,1\} \times \{1,2,3\}^m$.

[27] For example, the first n bits of the message might be some redundant header information, which is known to the adversary. Note that the assumption that an adversary has *a priori* knowledge about the message is standard in cryptography, and usually referred to as a *known-plaintext attack* [28]. It is indeed a historic fact that many successful attacks on realistic cryptosystems have been based on partial knowledge of the message.

[28] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography* (CRC Press, Boca Raton, FL, 1996).

[29] R. Renner, Ph.D. thesis, Swiss Federal Institute of Technology (ETH) Zurich, 2005, quant-ph/0512258.

[30] It has been shown in [15] that a key S is sufficiently secure to be used in applications if (5) holds for a security parameter ε which is exponentially small in the key size. Our example shows that this exponential dependence is in fact necessary, thus answering an open question in [15]. Note, however, that making the security parameter ε in (5) exponentially small comes at the cost of reducing the key rate substantially.

[31] A qualitatively equivalent security definition has been proposed independently in [15].

[32] $\|A\|_1$ denotes the L_1 norm of the operator A.

[33] We refer to [16,29] for a detailed description of privacy amplification in the context of quantum adversaries.

[34] B. Kraus, N. Gisin, and R. Renner, Phys. Rev. Lett. **95**, 080501 (2005).

[35] By Uhlmann's theorem, there exists a density operator σ_E such that $F(\rho_{ABE}, |\Phi^+\rangle^{\otimes n} \otimes \sigma_E) = F(\rho_{AB}, |\Phi^+\rangle^{\otimes n})$. The assertion of Lemma 2 then follows from the general inequality $\frac{1}{2}\|\rho - \sigma\|_1 \leq \sqrt{1 - F(\rho,\sigma)^2}$ and the fact that $\|\cdot\|_1$ cannot increase when taking the partial trace over A or B.

[36] These security proofs usually make use of a similar relation between the fidelity and the accessible information (see, e.g., Lemma 1 and 2 given in the supplementary material of [9] and the discussion in footnote 28 of [9]). Substituting this relation by Lemma 2 thus turns these arguments into proofs of security according to Definition 1 (see also [15]).