Algorithms **2009**, *2*, 790-807; doi:10.3390/a2020790

OPEN ACCESS

algorithms

ISSN 1999-4893

www.mdpi.com/journal/algorithms

Article

Security of the Bennett-Brassard Quantum Key Distribution Protocol against Collective Attacks

Michel Boyer [1,†], Ran Gelles [2,†,⋆] and Tal Mor [2,†]

[1] Département IRO, Université de Montréal, CP 6128 Succ. Centre-Ville, Montréal H3C 3J7, Canada
[2] Computer Science Department, Technion, Haifa 32000, Israel
E-mails: (boyer@iro.umontreal.ca); (talmo@cs.technion.ac.il)
[†] Alphabetical authors order.

[⋆] Author to whom correspondence should be addressed; E-mail: gelles@cs.technion.ac.il

Received: 16 October 2008; in revised form: 17 April 2009 / Accepted: 7 May 2009 / Published: 3 June 2009

Abstract: The theoretical Quantum Key-Distribution scheme of Bennett and Brassard (BB84) has been proven secure against very strong attacks including the collective attacks and the joint attacks. Though the latter are the most general attacks, collective attacks are much easier to analyze, yet, they are conjectured to be as informative to the eavesdropper. Thus, collective attacks are likely to be useful in the analysis of many theoretical and practical schemes that are still lacking a proof of security, including practical BB84 schemes. We show how powerful tools developed in previous works for proving security against the joint attack, are simplified when applied to the security of BB84 against collective attacks whilst providing the same bounds on leaked information and the same error threshold.

Keywords: QKD; security; collective attacks; error-rate threshold

1. Introduction

Quantum Theory allows us to have new cryptographic protocols of which we can prove security. Those protocols are secure against adversaries with unlimited power.[*] One of those protocols is the Quantum Key Distribution (QKD) protocol which is named BB84 after its inventors Bennett and Bras-

[*] The only limitations are the laws of physics as we currently know them.

sard [1]. In this protocol, two users (conventionally named Alice and Bob) wish to set up a common random key, using a quantum channel and a classical (insecure) authenticated channel. Their adversary (named Eve) is trying to eavesdrop on both of those channels in order to have as much information as possible about the agreed key.

The goal of Alice and Bob is to use a protocol that can be proven secure, potentially even unconditionally secure, against powerful eavesdropping. In this paper we discuss the security of the BB84 protocol against the *collective attacks* [2–4], that form a subclass of the joint attacks (which are the most powerful theoretic attacks). This subclass is conjectured[†] to contain the strongest joint attacks and therefore, to be as informative to Eve as the joint attack [2, 3]. In addition, analyzing the collective attack is much simpler than analyzing the joint attack. Thus, analyzing the collective attack might be highly relevant for practical setups of QKD where proving security is still a hard task.

We improve the analysis done in [4] to the BB84 scheme against all collective attacks. The analysis shown in [4] bounds the information in a non-optimal way which adds a factor of $2^r 2^m$ to the information bound, where r is the amount of error-correction bits revealed during the protocol, and m is the final-key length. Our proof uses methods that are used in [5] for the joint attack, in order to achieve an optimized bound for the collective attack.

Let $\mathscr{H}_2$ be the 2-dimensional Hilbert space with standard (or computational) basis $|0^0\rangle, |1^0\rangle$. Let $|0^1\rangle = \frac{1}{\sqrt{2}}[|0^0\rangle + |1^0\rangle]$ and $|1^1\rangle = \frac{1}{\sqrt{2}}[|0^0\rangle - |1^0\rangle]$; it is clear that $|0^1\rangle$, $|1^1\rangle$ is an orthonormal basis;[‡] it is called the Hadamard basis. The unitary map H such that $H|0^0\rangle = |0^1\rangle$ and $H|1^0\rangle = |1^1\rangle$ is called the Hadamard transform. Due to linearity, $H|0^1\rangle = \frac{1}{\sqrt{2}}[H|0^0\rangle + H|1^0\rangle] = |0^0\rangle$ and similarly, $H|1^1\rangle = |1^0\rangle$ i.e. $H \cdot H = I$ (the identity). Those bases are used for measurements in the BB84 scheme; measuring a state represented as the density matrix ρ in the b basis returns output 0 with probability $\langle 0^b|\rho|0^b\rangle$ and 1 with probability $\langle 1^b|\rho|1^b\rangle$. Thus if the state $|0^b\rangle$ (or $|1^b\rangle$) is measured in the b basis, it results with output 0 (1) with certainty. Yet, when $|0^b\rangle$ or $|1^b\rangle$ is measured in the $\bar{b} = 1 - b$ basis, the output is random, i.e. 0 with probability $1/2$ and 1 with probability $1/2$. This is the principle underlying the BB84 quantum key distribution protocol [1]. Alice sends Bob qubits (2-dimensional systems), each qubit in one of the four state $|i^b\rangle$ with $i, b \in \{0, 1\}$. In order to send a bitstring $\mathbf{i} = i_1 \dots i_t$ to Bob, Alice first draws randomly a bitstring $\mathbf{b} = b_1 \dots b_t$ and then sends the state $|\mathbf{i}^{\mathbf{b}}\rangle = |i_1^{b_1}\rangle \dots |i_t^{b_t}\rangle = H^{\mathbf{b}}|\mathbf{i}\rangle$ where $H^{\mathbf{b}} = H^{b_1} \otimes \dots \otimes H^{b_t}$ and $|\mathbf{i}\rangle = |i_1 \dots i_t\rangle$, with $H^0 = I$ and $H^1 = H$. In the conventional setting, Bob measures each qubit in one of those two bases, and whenever they used the same basis they obtain the same bit i. Using classical error correction and privacy amplification protocols, Alice and Bob reach a final key of length $m < t$ bits. In this paper, bitstrings of (an arbitrary) length t are denoted by a bold letter (e.g. we use below the $2n$ bits string $\mathbf{i} = i_1 \dots i_{2n}$ with $i_1, \dots, i_{2n} \in \{0, 1\}$) and are identified to elements of the t-dimensional $\mathbf{F}_2$-vector space $\mathbf{F}_2^t$.

1.1. A Formal Description of the BB84 Protocol

Let us describe the BB84 protocol we shall use in this paper.

1. Alice and Bob agree on a large number n, an error threshold p_a and on a linear error-correction

[†] See discussion in Section 4.

[‡] The notations we use match the physicists "spin-notations" where $|0^0\rangle = |0\rangle_z$ and $|1^0\rangle = |1\rangle_z$ is the standard basis, and $|0^1\rangle = |0\rangle_x$ and $|1^1\rangle = |1\rangle_x$ is the Hadamard basis.

code C with parity check matrix P_C of order $r \times n$. They agree as well on a linear key-generation function (privacy amplification) represented by a matrix P_K of order $m \times n$. Those matrices can be publicly known beforehand or they can be determined during the protocol and sent over the classical channel. The $(r+m) \times n$ matrix whose rows are those of P_C and P_K put together is required to be of rank $r+m$.

2. Alice randomly chooses $2n$-bit strings $\mathbf{i}, \mathbf{b} \in \mathbf{F}_2^{2n}$, where $\mathbf{F}_2$ denotes the two element field, with elements $\{0,1\}$, i.e. the field of integers modulo 2. Alice encodes the state $|\mathbf{i}^{\mathbf{b}}\rangle = |i_1^{b_1}\rangle \dots |i_{2n}^{b_{2n}}\rangle$ and sends it to Bob over the quantum channel, one qubit at a time. Each time Bob receives a qubit he informs Alice, yet he doesn't measure it.[§]

3. Alice publicly sends Bob the string $\mathbf{b}$. Bob applies $H^{\mathbf{b}} = H^{b_1} \otimes \dots \otimes H^{b_{2n}}$ to his state, so that if Bob had the state $|\mathbf{i}^{\mathbf{b}}\rangle$, once he performs $H^{\mathbf{b}}$ he possesses the state $|\mathbf{i}\rangle = |i_1 \dots i_{2n}\rangle$. Bob then measures these qubits in the computation basis.

   We denote by $\mathbf{i}^B$ the string measured by Bob. If there is no noise and no eavesdropping, he gets exactly the bitstring $\mathbf{i}$ sent by Alice.

4. Alice randomly chooses n-bits that will be used to detect eavesdropping. This is done by choosing a $2n$-bit string that has exactly n ones. Formally, Alice chooses $\mathbf{s} \in \mathbf{F}_2^{2n}$ such that $|\mathbf{s}| = n$. Alice publicly sends Bob $\mathbf{s}$.

   The bits indexed by $j \in [1 \,..\, 2n]$ such that $s_j = 0$ are used for testing, while the rest are used for generating the final key (via error correction and privacy amplification). We denote the appropriate substrings of $\mathbf{i}, \mathbf{b}$ that are relevant for the testing by $\mathbf{i}_{\bar{\mathbf{s}}}$ and $\mathbf{b}_{\bar{\mathbf{s}}}$, while the substrings relevant for creating the key are denoted $\mathbf{i}_{\mathbf{s}}$ and $\mathbf{b}_{\mathbf{s}}$.

5. For each $j \in [1 \,..\, 2n]$ such that $s_j = 0$, Alice and Bob publish the value of the jth-bit. Bob and Alice compare those bit values, and if more than np_a bits mismatch, they abort the protocol. The pre-fixed protocol parameter p_a is actually the ratio of allowed bit-flips on the testing bits.

6. Alice and Bob keep the values of the remaining n bits secret. Alice's string is denoted $\mathbf{x} = \mathbf{i}_{\mathbf{s}}$ and named the *information string*. The corresponding bitstring on Bob's side is denoted $\mathbf{x}^B$.

7. Alice sends Bob the r-bit error-correction string $\boldsymbol{\xi} = \mathbf{x}P_C^{\mathrm{T}}$ (where P_C^{T} is the transpose of the parity check matrix). Bob uses $\boldsymbol{\xi}$ to correct his string $\mathbf{x}^B$. The string $\boldsymbol{\xi}$ is called the *syndrome* of the string $\mathbf{x}$ (with regard to P_C).

8. Alice and Bob compute the m-bit final key $\mathbf{k} = \mathbf{x}P_K^{\mathrm{T}}$.

[§] Here we assume that Bob delays measuring each qubit till after learning its basis. In the more realistic case in which Bob cannot wait with his measurement, or in case some qubits are lost, Alice needs to send more qubits to make sure that $2n$ qubits are obtained (in Alice's bases) as required.

2. Description of Eve's attack and its properties

To each qubit $|\phi_j\rangle$ ($j \in [1\ ..\ 2n]$) sent by Alice, Eve attaches a separate probe that we assume to be in a pure state $|0_j^E\rangle$ and applies a unitary transform U_j to the composite system $|0_j^E\rangle|\phi_j\rangle$. She then keeps her probes in a quantum memory for subsequent measurement and sends Bob his part of the system. For each qubit there is thus a particular Hilbert probe space, and a particular U_j; they are decided beforehand by Eve and are thus fixed for all possible choices of $\mathbf{i}$, $\mathbf{b}$ and $\mathbf{s}$.

2.1. Eve's attack on a single qubit

Since the attack is bitwise, we now concentrate the analysis on some fixed qubit, drop momentarily the subindex j, and express the global effect of Eve's action on this particular qubit with respect to the basis $|0^b\rangle$, $|1^b\rangle$:

$$U|0^E\rangle|0^b\rangle = |E_{00}^b\rangle|0^b\rangle + |E_{01}^b\rangle|1^b\rangle = |\phi_0^b\rangle \tag{1}$$

$$U|0^E\rangle|1^b\rangle = |E_{10}^b\rangle|0^b\rangle + |E_{11}^b\rangle|1^b\rangle = |\phi_1^b\rangle; \tag{2}$$

$|E_{00}^b\rangle$, $|E_{01}^b\rangle$, $|E_{10}^b\rangle$ and $|E_{11}^b\rangle$ are vectors ("non normalized states") in Eve's Hilbert probe space corresponding to this particular qubit. Since U is unitary, $|\phi_0^b\rangle$ and $|\phi_1^b\rangle$ are of norm 1 and orthogonal. This means that

$$\langle E_{00}^b|E_{00}^b\rangle + \langle E_{01}^b|E_{01}^b\rangle = 1 \tag{3}$$

$$\langle E_{10}^b|E_{10}^b\rangle + \langle E_{11}^b|E_{11}^b\rangle = 1 \tag{4}$$

$$\langle E_{00}^b|E_{10}^b\rangle + \langle E_{01}^b|E_{11}^b\rangle = 0 \qquad \langle E_{10}^b|E_{00}^b\rangle + \langle E_{11}^b|E_{01}^b\rangle = 0 \tag{5}$$

2.2. Extending the attack to multiple qubits — the collective attack

For each qubit $j \in [1\ ..\ 2n]$, Eve applies the unitary U_j on the space $\mathscr{H}_j^E \otimes \mathscr{H}_2$ where $\mathscr{H}_j^E$ is her probe space and $\mathscr{H}_2$ is the qubit space. Eve's view expressed with respect to basis b_j is obtained by tracing out Bob from the states $|\phi_0^{b_j}\rangle_j$ and $|\phi_1^{b_j}\rangle_j$, resulting with the respective density matrices

$$(\rho_0^{b_j})_j = |E_{00}^{b_j}\rangle_j\langle E_{00}^{b_j}| + |E_{01}^{b_j}\rangle_j\langle E_{01}^{b_j}| \tag{6}$$

$$(\rho_1^{b_j})_j = |E_{10}^{b_j}\rangle_j\langle E_{10}^{b_j}| + |E_{11}^{b_j}\rangle_j\langle E_{11}^{b_j}|. \tag{7}$$

If Alice sends the string $\mathbf{i}$ using bases $\mathbf{b}$, then Eve's global state is the tensor product of all those states $(\rho_{i_j}^{b_j})_j$. After the test bits are revealed, Eve needs only those $(\rho_{i_j}^{b_j})_j$ for which $s_j = 1$. The set $\{j \mid s_j = 1\}$ has n elements; let us denote it $\{j_1, \ldots, j_n\}$, so that $s_{j_l} = 1$ for $1 \leq l \leq n$. Eve's global state corresponding to $\mathbf{s}$, $\mathbf{b}$ and $\mathbf{x}$ can now be written

$$\rho_{\mathbf{x}}^{\mathbf{b_s}} = (\rho_{i_{j_1}}^{b_{j_1}})_{j_1} \otimes \ldots \otimes (\rho_{i_{j_n}}^{b_{j_n}})_{j_n} = \bigotimes_{l=1}^{n} \left(\rho_{i_{j_l}}^{b_{j_l}}\right)_{j_l}. \tag{8}$$

We can rewrite Eq. (8) using the n-bit strings $\mathbf{x}$ and $\mathbf{b}' = \mathbf{b_s}$ with the index l running from 1 to n (instead of the $2n$ strings $\mathbf{i}$ and $\mathbf{b}$ indexed by $\{j_1, \ldots, j_n\}$),

$$\rho_{\mathbf{x}}^{\mathbf{b_s}} = \rho_{\mathbf{x}}^{\mathbf{b}'} = \rho_{x_1}^{b'_1} \otimes \ldots \otimes \rho_{x_n}^{b'_n} = \bigotimes_{l=1}^{n} \rho_{x_l}^{b'_l}. \tag{9}$$

It is the state $\rho_{\mathbf{x}}^{\mathbf{b}_\mathbf{s}}$ (or a mixture of such states) that Eve measures collectively to guess the string $\mathbf{x}$ (or directly the final key $\mathbf{k}$) once $\mathbf{b}$, $\mathbf{s}$ and the information for error correction and privacy amplification is known to her.

2.3. *The probability of error*

Assuming a qubit is attacked by U as defined by (1) and (2), an error occurs if Alice sends 0 and Bob measures 1 or if Alice sends 1 and Bob measures 0. Let k be the value measured by Bob, i the value sent by Alice for a specific qubit, and b the basis used by Alice to encode i. The probability of Bob measuring an error is then given by

$$p(k=1 \mid i=0)p(i=0)+p(k=0 \mid i=1)p(i=1)=\langle E_{01}^b|E_{01}^b\rangle\frac{1}{2}+\langle E_{10}^b|E_{10}^b\rangle\frac{1}{2},$$

and we denote

$$p_e^b \triangleq \frac{1}{2}\left[\langle E_{01}^b|E_{01}^b\rangle+\langle E_{10}^b|E_{10}^b\rangle\right]. \tag{10}$$

2.4. *The probability of error in the conjugate basis*

We are now interested in $p_e^{\bar{b}}$ where $\bar{b}=1-b$ (i.e. $\bar{0}=1$ and $\bar{1}=0$) corresponds to the basis conjugate to that given by b. The attack U is always the one described by (1) and (2) in the b basis but, in order to calculate the probability or error when Alice encodes i_j as $|i_j^{\bar{b}}\rangle$ instead of $|i_j^b\rangle$, we now need to express U in the $\bar{b}$ basis. From (10), we know that the probability of error for this situation is given by

$$p_e^{\bar{b}}=\frac{1}{2}\left[\langle E_{01}^{\bar{b}}|E_{01}^{\bar{b}}\rangle+\langle E_{10}^{\bar{b}}|E_{10}^{\bar{b}}\rangle\right].$$

Using the fact that

$$|0\rangle^{\bar{b}}=\frac{1}{\sqrt{2}}[|0^b\rangle+|1^b\rangle], \qquad |1\rangle^{\bar{b}}=\frac{1}{\sqrt{2}}[|0^b\rangle-|1^b\rangle]$$

and using the linearity of U, we deduce directly from (1) and (2) that

$$U|0^E\rangle|0^{\bar{b}}\rangle=\frac{1}{\sqrt{2}}\left(|E_{00}^b\rangle+|E_{10}^b\rangle\right)|0^b\rangle+\frac{1}{\sqrt{2}}\left(|E_{01}^b\rangle+|E_{11}^b\rangle\right)|1^b\rangle, \tag{11}$$

$$U|0^E\rangle|1^{\bar{b}}\rangle=\frac{1}{\sqrt{2}}\left(|E_{00}^b\rangle-|E_{10}^b\rangle\right)|0^b\rangle+\frac{1}{\sqrt{2}}\left(|E_{01}^b\rangle-|E_{11}^b\rangle\right)|1^b\rangle. \tag{12}$$

Replacing $|0^b\rangle$ and $|1^b\rangle$ on the right-hand sides with their values in terms of $|0^{\bar{b}}\rangle$ and $|1^{\bar{b}}\rangle$ i.e. $|0^b\rangle = \frac{1}{\sqrt{2}}[|0^{\bar{b}}\rangle+|1^{\bar{b}}\rangle]$ and $|1^b\rangle=\frac{1}{\sqrt{2}}[|0^{\bar{b}}\rangle-|1^{\bar{b}}\rangle]$ we obtain

$$\begin{aligned} U|0^E\rangle|0^{\bar{b}}\rangle = {} & \frac{1}{2}\left[|E_{00}^b\rangle+|E_{10}^b\rangle+|E_{01}^b\rangle+|E_{11}^b\rangle\right]|0^{\bar{b}}\rangle \\ & +\frac{1}{2}\left[\left(|E_{00}^b\rangle-|E_{11}^b\rangle\right)+\left(|E_{10}^b\rangle-|E_{01}^b\rangle\right)\right]|1^{\bar{b}}\rangle \end{aligned} \tag{13}$$

$$\begin{aligned} U|0^E\rangle|1^{\bar{b}}\rangle = {} & \frac{1}{2}\left[\left(|E_{00}^b\rangle-|E_{11}^b\rangle\right)-\left(|E_{10}^b\rangle-|E_{01}^b\rangle\right)\right]|0^{\bar{b}}\rangle \\ & +\frac{1}{2}\left[|E_{00}^b\rangle-|E_{10}^b\rangle-|E_{01}^b\rangle+|E_{11}^b\rangle\right]|1^{\bar{b}}\rangle \end{aligned} \tag{14}$$

where the terms for $|E^{\bar{b}}_{01}\rangle$ and $|E^{\bar{b}}_{10}\rangle$ are parenthesized so that we can easily see that

$$\begin{aligned} p^{\bar{b}}_e &= \frac{1}{2}\left[\langle E^{\bar{b}}_{01}|E^{\bar{b}}_{01}\rangle + \langle E^{\bar{b}}_{10}|E^{\bar{b}}_{10}\rangle\right] \\ &= \frac{1}{4}\left[(\langle E^{b}_{00}| - \langle E^{b}_{11}|)(|E^{b}_{00}\rangle - |E^{b}_{11}\rangle) + (\langle E^{b}_{10}| - \langle E^{b}_{01}|)(|E^{b}_{10}\rangle - |E^{b}_{01}\rangle)\right]. \end{aligned}$$

We expand this result by using the identities $\langle\phi|\psi\rangle = \overline{\langle\psi|\phi\rangle}$ and $z + \overline{z} = 2\mathrm{Re}(z)$ for $z \in \mathbf{C}$ (here the overline indicates the complex conjugate). Using equalities (3) and (4) we get

$$\begin{aligned} p^{\bar{b}}_e &= \frac{1}{4}\left[2 - \langle E^{b}_{00}|E^{b}_{11}\rangle - \langle E^{b}_{11}|E^{b}_{00}\rangle - \langle E^{b}_{01}|E^{b}_{10}\rangle - \langle E^{b}_{10}|E^{b}_{01}\rangle\right] \\ p^{\bar{b}}_e &= \frac{1}{2}\left[1 - \mathrm{Re}\left(\langle E^{b}_{00}|E^{b}_{11}\rangle + \langle E^{b}_{01}|E^{b}_{10}\rangle\right)\right]. \end{aligned} \tag{15}$$

This Formula will be used to connect the disturbance induced by Eve when Alice encodes in the $\overline{b}_j$ basis bits i_j such that $s_j = 1$ to the information Eve can get when Alice encodes them in the b_j basis. Following the "Information versus Disturbance" [6] principle we will show that the more information Eve gets when the encoding is in the b basis, the more disturbance she causes when the bits are encoded and tested in the conjugate basis. Hence, we can bound Eve's knowledge about the key by bounding the allowed error-rate in the protocol.

2.5. *Flat attacks with respect to basis b*

Assume now that Eve's attack U is fixed, and that $P^{\bar{b}}_e$ is given by Eq. (15). We will present a virtual attack that is proven to be better for Eve, as it induces a smaller error-rate. This virtual attack cannot be executed by Eve since it requires knowledge of the basis b used by Alice (a knowledge that, of course, Eve does not have at the stage in which she chooses her transformation U). Still, the existence of such an attack that is proven to be better for Eve, allows us to derive bounds on Eve's knowledge when the original attack (actually used by Eve) is applied.

Proposition 1. *For each attack U with ρ^b_0, ρ^b_1 and p^b_e given by (6), (7) and (10), that satisfy*

$$\langle E^{b}_{00}|E^{b}_{11}\rangle + \langle E^{b}_{01}|E^{b}_{10}\rangle = e^{i\theta} r \qquad \textit{for } r \in \mathbf{R}_+ \tag{16}$$

there exists U'_b with the same ρ^b_0, ρ^b_1 and p^b_e as U, which satisfy

$$\langle E'^{b}_{00}|E'^{b}_{11}\rangle + \langle E'^{b}_{01}|E'^{b}_{10}\rangle = r. \tag{17}$$

Proof. Let $S^b_\theta : \mathcal{H}_2 \to \mathcal{H}_2$ be defined by $S^b_\theta|0^b\rangle = |0^b\rangle$ and $S^b_\theta|1^b\rangle = e^{i\theta}|1^b\rangle$. S^b_θ is clearly unitary and consequently so is $\mathbf{1}^E_j \otimes S_\theta : \mathcal{H}^E_j \otimes \mathcal{H}_2 \to \mathcal{H}^E_j \otimes \mathcal{H}_2$ where $\mathbf{1}^E_j$ is the identity on $\mathcal{H}^E_j$. Let $U'_b = U(\mathbf{1}^E_j \otimes S_{-\theta})$. U'_b is such that

$$U'_b|0\rangle|0^b\rangle = \quad |E^{b}_{00}\rangle|0^b\rangle + \quad |E^{b}_{01}\rangle|1^b\rangle = |\phi'^b_0\rangle, \tag{18}$$

$$U'_b|0\rangle|1^b\rangle = e^{-i\theta}|E^{b}_{10}\rangle|0^b\rangle + e^{-i\theta}|E^{b}_{11}\rangle|1^b\rangle = |\phi'^b_1\rangle. \tag{19}$$

From those equalities it follows that ρ_0^b and ρ_1^b are left unchanged as can be seen from equations (6) and (7). In the same way, the right hand side of (10) is also clearly left unchanged and so p_e^b is left unchanged. Finally

$$\begin{aligned}\langle E'^b_{00}|E'^b_{11}\rangle + \langle E'^b_{01}|E'^b_{10}\rangle &= \langle E^b_{00}|e^{-i\theta}E^b_{11}\rangle + \langle E^b_{01}|e^{-i\theta}E^b_{10}\rangle\\ &= e^{-i\theta}e^{i\theta}r \qquad \text{by (16)}\\ &= r.\end{aligned}$$

□

The attack U'_b provides the same "view" ρ_0^b, ρ_1^b to Eve, and the same probability of being detected if the b basis is used. However, from Eq. (15) we see that it reduces $p_e^{\bar{b}}$ to the minimum value (15) can take, because $\text{Re}(z) \leq |z|$ for any $z \in \mathbf{C}$. This means that by replacing U by U'_b Eve's probability of being detected had the other basis been chosen can only decrease; U'_b is thus better for Eve, since she needs to take into account all possible bases used by Alice. U'_b will be coined the "flat" attack associated to U with respect to basis b. Since Eve is not aware of the basis b used, the flat attack is merely a mathematical tool. Moreover it depends on b. However, by bounding Eve's information when that basis is used we will eventually get a bound on Eve's information under the original attack.

In the more general case of bitstrings, since Eve's view comes from the tensor product of density matrices on individual qubits, using the flat attacks on all qubits does not change Eve's global view, nor the probability of error in the b basis. A flat attack will thus be flat for each qubit. In a flat attack (one qubit case), there exist $r \in \mathbf{R}_+$ such that

$$\langle E^b_{00}|E^b_{11}\rangle + \langle E^b_{01}|E^b_{10}\rangle = r, \tag{20}$$

$$p_e^{\bar{b}} = \frac{1}{2}(1-r). \tag{21}$$

A short summary: we consider two possible cases for a specific qubit sent by Alice to Bob that is attacked by Eve with a flat unitary transform U:

1. Alice and Bob use the b basis. Eve's attack causes a bit-flip with probability $p_e^b = \frac{1}{2}\left[\langle E^b_{01}|E^b_{01}\rangle + \langle E^b_{10}|E^b_{10}\rangle\right]$.
2. However, if Alice and Bob use the $\bar{b}$ basis, Eve's attack causes a bit-flip with probability $p_e^{\bar{b}} = \frac{1}{2}\left[1 - \text{Re}\left(\langle E^b_{00}|E^b_{11}\rangle + \langle E^b_{01}|E^b_{10}\rangle\right)\right] = \frac{1}{2}(1-r)$.

2.6. A purification

We now assume the attack is flat, i.e. it satisfies equations (3)–(5), (20), and also, as a result, equation (21). Still working on a single qubit let us now define $|\psi_0^b\rangle$ and $|\psi_1^b\rangle$ as

$$|\psi_0^b\rangle = |E^b_{00}\rangle|0\rangle + |E^b_{01}\rangle|1\rangle; \qquad |\psi_1^b\rangle = |E^b_{11}\rangle|0\rangle + |E^b_{10}\rangle|1\rangle. \tag{22}$$

where the (normalized and orthogonal) states $|0\rangle$ and $|1\rangle$ live in some Hilbert space $\mathscr{H}$ that need not correspond to any physical reality (they are useful mathematical entities). If we trace states $|\psi_0^b\rangle\langle\psi_0^b|$ and $|\psi_1^b\rangle\langle\psi_1^b|$ over the span of $|0\rangle$ and $|1\rangle$ in $\mathscr{H}$, we get the states ρ_0 and ρ_1 respectively. The states $|\psi_0^b\rangle$ and

$|\psi_1^b\rangle$ are thus called *lift-ups* of ρ_0 and ρ_1. Since they are also pure, they are said to be *purifications* of ρ_0 and ρ_1. Moreover they are normalized and by Eq. (20) their overlap is

$$\langle\psi_0^b|\psi_1^b\rangle = \langle E_{00}^b|E_{11}^b\rangle + \langle E_{01}^b|E_{10}^b\rangle = r. \tag{23}$$

This establishes a direct relation between the overlap of $|\psi_0^b\rangle$ and $|\psi_1^b\rangle$ and the probability of error $p_e^{\bar{b}}$. Since the overlap r is real and positive, with $0 \le r \le 1$, there is an angle α such that

$$\cos(2\alpha) = r = \langle\psi_0^b|\psi_1^b\rangle \qquad\qquad 0 \le \alpha \le \pi/4.$$

As a consequence, we get

$$p_e^{\bar{b}} = \frac{1}{2}[1 - \cos(2\alpha)] = \sin^2(\alpha) \quad \text{or} \quad \sin(\alpha) = (p_e^{\bar{b}})^{1/2}. \tag{24}$$

Since $\langle\psi_0^b|\psi_1^b\rangle$ is real, it is equal to $\langle\psi_1^b|\psi_0^b\rangle$ and consequently the (non normalized) states $|\psi_0^b\rangle + |\psi_1^b\rangle$ and $|\psi_0^b\rangle - |\psi_1^b\rangle$ are orthogonal and their norms are $\sqrt{2+2\cos(2\alpha)} = 2\cos(\alpha)$ and $\sqrt{2-2\cos(2\alpha)} = 2\sin(\alpha)$ respectively. We thus let

$$|0_H^b\rangle = \frac{1}{2\cos(\alpha)}[|\psi_0^b\rangle + |\psi_1^b\rangle]; \qquad\qquad |1_H^b\rangle = \frac{1}{2\sin(\alpha)}[|\psi_0^b\rangle - |\psi_1^b\rangle]. \tag{25}$$

Using this basis, we can re-write the purification for $x \in \{0,1\}$, as

$$|\psi_x^b\rangle = \cos(\alpha)|0_H^b\rangle + (-1)^x \sin(\alpha)|1_H^b\rangle. \tag{26}$$

3. Proof of security of BB84 against collective attacks

3.1. Parity strings for the code and the key

We recall that bitstrings of length n are identified with elements of $\mathbf{F}_2^n$. Vector addition thus corresponds to component-wise sum modulo 2 and thus to the eXclusive-OR of the corresponding bitstrings. We denote $\mathbf{a}\cdot\mathbf{b}$ the scalar product (modulo 2) of the two strings $\mathbf{a}$ and $\mathbf{b}$ of the same length, e.g., for n-bit strings, $\mathbf{a}\cdot\mathbf{b} = \sum_{i=1}^n a_ib_i = a_1b_1 + \ldots + a_nb_n$. Let $\{v_1,\ldots,v_n\}$ be a basis of $\mathbf{F}_2^n$. For any r' let $V_{r'}$ denote the span of $\{v_1,\ldots,v_{r'}\}$ and $V_{r'}^c$ the span of $\{v_{r'+1},\ldots,v_n\}$; it is clear that $V_{r'} + V_{r'}^c = \mathbf{F}_2^n$; moreover, if we let $v, w \in V_{r'}$ and $v', w' \in V_{r'}^c$ then

$$v + v' = w + w' \quad \Longrightarrow \quad v = w \ \text{ and } \ v' = w'. \tag{27}$$

This property is normally summarized by saying that $\mathbf{F}_2^n$ is the *direct sum* of $V_{r'}$ and $V_{r'}^c$, i.e., $V_{r'} + V_{r'}^c = \mathbf{F}_2^n$ and $V_{r'} \cap V_{r'}^c = \{0\}$.

The vectors $v_1,\ldots,v_r$ are used as the rows of P_C, the parity check matrix for the error correcting code which yields the syndrome $\boldsymbol{\xi} = \mathbf{x}P_C^{\mathrm{T}}$; the vectors $v_{r+1},\ldots,v_{r+m}$ are used as the rows of a privacy amplification matrix P_K such that if $\mathbf{x}$ is the string sent by Alice, then the m-bit key is $\mathbf{x}P_K^{\mathrm{T}}$. Let

$$d_{r,m} \triangleq \min_{r \le r' < r+m} d_H(v_{r'+1}, V_{r'}) = \min_{r \le r' < r+m} d_{r',1}. \tag{28}$$

This parameter on which security depends relates in terms of Hamming distance the parity strings used to generate the key $\mathbf{k}$ to the parity strings used to generate the error correction information $\boldsymbol{\xi}$. A large value of $d_{r,m}$ will be shown to imply little information for Eve on the key $\mathbf{k}$, given she knows $\boldsymbol{\xi}$ (Theorem 8).

3.2. *The Shannon distinguishability*

We shall use $SD(\alpha, \beta)$ as it is defined in [4, 5] to denote the *Shannon Distinguishability* between the state (or density matrix) α and the state (or density matrix) β. Consider the following protocol: Sam chooses '0' or '1', randomly with equal probability. If Sam chooses '0', he sends the state α over to Rachel. Otherwise, he sends β. $SD(\alpha, \beta)$ is by definition Rachel's accessible information i.e. the maximum mutual information between Sam's encoded bit and Rachel's measurement of the state she received. Notice that when α and β are orthogonal (thus they form a basis), Rachel can always distinguish between them, and has information of exactly 1 bit about Sam's chosen bit. On the other hand, if $\alpha = \beta$, Rachel can never distinguish between those states, and she has 0 bits of information. Important result of the SD function are summarized in the following lemma:

Lemma 2. *(a) If $\widetilde{\rho}_x$ is a lift-up of ρ_x (where $x \in \{0, 1\}$), then $SD(\rho_0, \rho_1) \leq SD(\widetilde{\rho}_0, \widetilde{\rho}_1)$; (b) The Shannon Distinguishability of two states can be bounded by half the Trace Norm of their difference: $SD(\rho_0, \rho_1) \leq \frac{1}{2} \operatorname{tr} |\rho_0 - \rho_1|$*

Proof. See [4, Theorem 1 and 2]. □

3.3. *Representing states for bitstrings*

Let s be a fixed string of length $2n$ with a 1 in positions $j_1, \ldots, j_n$ corresponding to the n information bits. As in Eq. (9), given the basis string $\mathbf{b}' = b'_1 \ldots b'_n = b_{j_1} \ldots b_{j_n}$ and $\mathbf{x} = x_1 \ldots x_n = i_{j_1} \ldots i_{j_n}$ we define the state $|\psi_{\mathbf{x}}^{\mathbf{b}'}\rangle = \bigotimes_{l=1}^{n} |\psi_{x_l}^{b'_l}\rangle$. Using (26), we write the state as

$$|\psi_{\mathbf{x}}^{\mathbf{b}'}\rangle = \bigotimes_{l=1}^{n} \left[\cos(\alpha_l)\,|0_l\rangle + (-1)^{x_l} \sin(\alpha_l)\,|1_l\rangle\right], \tag{29}$$

where $|0_l\rangle$ and $|1_l\rangle$ represent the vectors $|0_H^{b'_l}\rangle$ and $|1_H^{b'_l}\rangle$ corresponding to the attack U_{j_l} on the j_l-th qubit (the l-th information qubit). If for $\mathbf{c} = c_1 \ldots c_n \in \{0, 1\}^n$ we define

$$d_{\mathbf{c},l} = \begin{cases} \cos(\alpha_l) & \text{if } c_l = 0 \\ \sin(\alpha_l) & \text{if } c_l = 1 \end{cases} \qquad\qquad d_{\mathbf{c}} = d_{\mathbf{c},1} \ldots d_{\mathbf{c},n}$$

then

$$|\psi_{\mathbf{x}}^{\mathbf{b}'}\rangle = \sum_{\mathbf{c} \in \{0,1\}^n} d_{\mathbf{c}} (-1)^{\mathbf{x} \cdot \mathbf{c}}\,|\mathbf{c}\rangle \tag{30}$$

where $|\mathbf{c}\rangle$ stands for $|(c_1)_1 \ldots (c_n)_n\rangle$; for instance if $\mathbf{c} = 0100$ then $|\mathbf{c}\rangle$ is $|0_1\rangle|1_2\rangle|0_3\rangle|0_4\rangle$ with $|0_l\rangle$ and $|1_l\rangle$ as defined above, and $d_{\mathbf{c}} = \cos(\alpha_1)\sin(\alpha_2)\cos(\alpha_3)\cos(\alpha_4)$. We notice that the factors of $d_{\mathbf{c}}^2$ can be interpreted as probabilities, and from (24) we deduce

$$d_{\mathbf{c},l}^2 = \begin{cases} \cos^2(\alpha_l) = q_l^{\overline{b'_l}} & \text{if } c_l = 0 \\ \sin^2(\alpha_l) = p_l^{\overline{b'_l}} & \text{if } c_l = 1 \end{cases}$$

where $p_l^{\overline{b'_l}}$ is the probability of an error on the bit of index j_l (the l-th information bit) when encoded and measured in the conjuguate basis and $q_l^{\overline{b'_l}} = 1 - p_l^{\overline{b'_l}}$ is the probability of no error on the same bit under the same conditions.

Due to the above, $d_{\mathbf{c}}^2$ is the probability of having exactly the error string $\mathbf{c}$ on the bits i_j such that $s_j = 1$ when those bits are encoded and measured in the other basis. Since, according to the protocol, the bits such that $s_j = 1$ are the "information bits", we will say, by abuse of language, that this is the probability of error on information bits. If we represent by $\mathbf{C}_I$ the random variable corresponding to the error in Bob's measurement of the information bits, and by $\mathbf{B}_I$ the random variable giving the corresponding basis string chosen by Alice then we can write, for $\mathbf{c} \in \{0,1\}^n$, $\mathbf{b} \in \{0,1\}^{2n}$ and $\mathbf{s} \in \{0,1\}^{2n}$ such that $|\mathbf{s}| = n$,

$$d_{\mathbf{c}}^2 = P[\mathbf{C}_I = \mathbf{c} \mid \mathbf{B}_I = \overline{\mathbf{b_s}}, \mathbf{s}] \tag{31}$$

where $\overline{\mathbf{b_s}} = \overline{\mathbf{b}'} = \overline{b'_1} \ldots \overline{b'_n}$. This probability is not conditional on the syndrome $\boldsymbol{\xi}$; all possible errors are taken into account here, even with values of $\mathbf{x}$ inconsistent with $\boldsymbol{\xi}$.

3.4. *Case of a one-bit key*

We begin with proving the security of a 1-bit key, and then extend our proof to an arbitrary m-bit length key. This case corresponds to $m = 1$ and the key (when not discarded) is $\mathbf{x} \cdot v_{r+1}$ where $\mathbf{x}$ is the string sent by Alice (that is, P_K has only one row, which equals v_{r+1}). Let $\boldsymbol{\xi} = \mathbf{x}P_C^T$ be the r bit syndrome announced publicly by Alice and let us denote $\widehat{\rho}_0$ and $\widehat{\rho}_1$ Eve's states corresponding to key 0 and key 1 respectively. Those states are obtained by normalizing the operators[¶]

$$\rho_k = \sum_{\mathbf{x} \,\middle|\, \substack{\mathbf{x}P_C^T = \boldsymbol{\xi} \\ \mathbf{x} \cdot v_{r+1} = k}} \rho_{\mathbf{x}}^{\mathbf{b}'}$$

and, since $\operatorname{tr}(\rho_0) = \operatorname{tr}(\rho_1) = 2^{n-r-1}$, $\widehat{\rho}_0$ and $\widehat{\rho}_1$ are equally likely, and

$$\widehat{\rho}_k = \frac{1}{2^{n-r-1}} \sum_{\mathbf{x} \,\middle|\, \substack{\mathbf{x}P_C^T = \boldsymbol{\xi} \\ \mathbf{x} \cdot v_{r+1} = k}} \rho_{\mathbf{x}}^{\mathbf{b}'}. \tag{32}$$

Changing the attack to a flat one, does not change $\rho_{x_l}^{b'_l}$, and therefore does not change $\widehat{\rho}_k$. Moreover, since $|\psi_{x_l}^{b'_l}\rangle\langle\psi_{x_l}^{b'_l}|$ as defined in Equation (26) is a purification of $\rho_{x_l}^{b'_l}$, it follows that

$$\widetilde{\rho}_k = \frac{1}{2^{n-r-1}} \sum_{\mathbf{x} \,\middle|\, \substack{\mathbf{x}P_C^T = \boldsymbol{\xi} \\ \mathbf{x} \cdot v_{r+1} = k}} |\psi_{\mathbf{x}}^{\mathbf{b}'}\rangle\langle\psi_{\mathbf{x}}^{\mathbf{b}'}| \tag{33}$$

is a lift-up of $\widehat{\rho}_k$. According to lemma 2, $SD(\widehat{\rho}_0, \widehat{\rho}_1) \leq SD(\widetilde{\rho}_0, \widetilde{\rho}_1)$ and $SD(\widetilde{\rho}_0, \widetilde{\rho}_1) \leq \frac{1}{2} \operatorname{tr} |\widetilde{\rho}_0 - \widetilde{\rho}_1|$ and thus

$$SD(\widehat{\rho}_0, \widehat{\rho}_1) \leq \frac{1}{2} \operatorname{tr} |\widetilde{\rho}_0 - \widetilde{\rho}_1|. \tag{34}$$

3.5. *Calculating and bounding the trace norm for one bit: the Biham basis.*

We now wish to bound $\frac{1}{2} \operatorname{tr} |\widetilde{\rho}_0 - \widetilde{\rho}_1|$ according to the specific attack Eve has performed. Taking advantage of the fact that $V_{r'} + V_{r'}^c = \mathbf{F}_2^n$ and $V_{r'} \cap V_{r'}^c = \{0\}$ (i.e. the sum is "direct"), equation (30)

[¶] State $\rho_{\mathbf{x}}^{\mathbf{b}'}$ is defined by (9) and (8).

rewrites as

$$|\psi_{\mathbf{x}}^{\mathbf{b}'}\rangle = \sum_{v\in V_r^c}(-1)^{\mathbf{x}\cdot v}\sum_{v'\in V_r}(-1)^{\mathbf{x}\cdot v'}d_{v+v'}|v+v'\rangle. \tag{35}$$

For each $\boldsymbol{\xi} \in \{0,1\}^r$, let i_ξ be a fixed n-bit string such that $i_\xi P_C^T = \boldsymbol{\xi}$. By definition of the syndrome, $\boldsymbol{\xi} = \mathbf{x}P_C^T$ and thus $(\mathbf{x} - \boldsymbol{\xi})P_C^T = 0$, i.e. $(\mathbf{x} - i_\xi)$ is a code word of C. Since every string v' in the dual code $C^\perp = V_r$ is orthogonal to every code word, we get that $v'(\mathbf{x} - i_\xi) = 0$ and thus $v'\mathbf{x} = v'i_\xi$. It follows that

$$|\psi_{\mathbf{x}}^{\mathbf{b}'}\rangle = \sum_{v\in V_r^c}(-1)^{\mathbf{x}\cdot v}\sum_{v'\in V_r}(-1)^{i_\xi\cdot v'}d_{v+v'}|v+v'\rangle.$$

If we define $|\eta_v\rangle = \sum_{v'\in V_r}(-1)^{i_\xi\cdot v'}d_{v+v'}|v+v'\rangle$, we conclude with

$$|\psi_{\mathbf{x}}^{\mathbf{b}'}\rangle = \sum_{v\in V_r^c}(-1)^{\mathbf{x}\cdot v}|\eta_v\rangle. \tag{36}$$

Lemma 3. *The non normalized states $|\eta_v\rangle$ for $v \in V_r^c$ are orthogonal.*

Proof.

$$\langle\eta_{v_1}|\eta_{v_2}\rangle = \sum_{v_1'\in V_r}(-1)^{i_\xi\cdot v_1'}\overline{d_{v+v'}}\langle v_1+v_1'|\sum_{v_2'\in V_r}(-1)^{i_\xi\cdot v_2'}d_{v_2+v_2'}|v_2+v_2'\rangle.$$

If $\langle v_1+v_1'|v_2+v_2'\rangle \neq 0$, then $v_1+v_1' = v_2+v_2'$ which, by (27), implies $v_1 = v_2$ (and $v_1' = v_2'$). □

The $|\eta_v\rangle$ thus provide an orthogonal (but not orthonormal) basis with which we can simply represent $|\psi_{\mathbf{x}}^{\mathbf{b}'}\rangle$, as shown in (36).

Using (33) we get

$$\widetilde{\rho}_0 - \widetilde{\rho}_1 = \frac{1}{2^{n-r-1}}\sum_{\mathbf{x}\,\Big|\substack{\mathbf{x}P_C^T = \boldsymbol{\xi}\\ \mathbf{x}\cdot v_{r+1} = 0}}|\psi_{\mathbf{x}}^{\mathbf{b}'}\rangle\langle\psi_{\mathbf{x}}^{\mathbf{b}'}| - \frac{1}{2^{n-r-1}}\sum_{\mathbf{x}\,\Big|\substack{\mathbf{x}P_C^T = \boldsymbol{\xi}\\ \mathbf{x}\cdot v_{r+1} = 1}}|\psi_{\mathbf{x}}^{\mathbf{b}'}\rangle\langle\psi_{\mathbf{x}}^{\mathbf{b}'}|.$$

The set of elements $\{\mathbf{x} \mid \mathbf{x}P_C^T = \boldsymbol{\xi}\}$ is the code coset containing the string i_ξ, namely, $\{c + i_\xi \mid c \in C\}$, where for every different element c, the string $c+i_\xi$ represents a different possible $\mathbf{x}$. Moreover, the final key bit k can be written as $(c+i_\xi)\cdot v_{r+1}$ and using (36), we get

$$\begin{aligned}\widetilde{\rho}_0 - \widetilde{\rho}_1 &= \frac{1}{2^{n-r-1}}\sum_{c\in C}(-1)^{(c+i_\xi)\cdot v_{r+1}}|\psi_{c+i_\xi}^{\mathbf{b}'}\rangle\langle\psi_{c+i_\xi}^{\mathbf{b}'}| \\ &= \frac{1}{2^{n-r-1}}\sum_{c\in C}(-1)^{(c+i_\xi)\cdot v_{r+1}}\sum_{m\in V_r^c}(-1)^{(c+i_\xi)\cdot m}|\eta_m\rangle\sum_{m'\in V_r^c}(-1)^{(c+i_\xi)\cdot m'}\langle\eta_{m'}|\end{aligned}$$

which can be written as

$$\widetilde{\rho}_0 - \widetilde{\rho}_1 = \frac{1}{2^{n-r-1}}\sum_{m,m'\in V_r^c}(-1)^{(m+m'+v_{r+1})\cdot i_\xi}\left(\sum_{c\in C}(-1)^{(m+m'+v_{r+1})\cdot c}\right)|\eta_m\rangle\langle\eta_{m'}|. \tag{37}$$

Lemma 4. *For every Linear Code C,*

$$\sum_{c\in C}(-1)^{c\cdot w} = \begin{cases}|C| & w \in C^\perp \\ 0 & else,\end{cases}$$

Proof. If $w \in C^\perp$ then $c \cdot w = 0$ for every $c \in C$ by the definition of $C^\perp$. Otherwise, let $\{\beta_1 \dots \beta_k\}$ be a basis of C over $\mathbf{F}_2$. Every codeword $c \in C$ can be written in a unique way as a linear combination $c = \alpha_1\beta_1 + \dots + \alpha_k\beta_k$ with $(\alpha_1, \dots, \alpha_k) \in \mathbf{F}_2^k$. Since $w \notin C^\perp$ there is i such that $\beta_i \cdot w \neq 0$. Assume *wlg* that $\beta_1 \cdot w = 1$; then

$$\begin{aligned}\sum_{c\in C}(-1)^{c\cdot w} &= \sum_{(\alpha_1,\dots,\alpha_k)\in\mathbf{F}_2^k} (-1)^{(\alpha_1\beta_1+\dots+\alpha_k\beta_k)\cdot w} \\ &= [(-1)^0 + (-1)^1] \sum_{(\alpha_2,\dots,\alpha_k)\in\mathbf{F}_2^{k-1}} (-1)^{(\alpha_2\beta_2+\dots+\alpha_k\beta_k)\cdot w} = 0.\end{aligned}$$ □

By Lemma 4, the parenthesized factor in the right-hand side of (37) is zero unless $m + m' + v_{r+1} \in C^\perp = V_r$, however, $m, m', v_{r+1} \in V_r^c$, and so is their sum. Thus, when the parenthesized factor is not zero, $m + m' + v_{r+1}$ must equal 0, since $V_r \cap V_r^c = \{0\}$. The resulting sum must equal $|C| = 2^{n-r}$. The equality $m + m' + v_{r+1} = 0$ rewrites as $m' = m + v_{r+1}$ and (37) reduces to

$$\begin{aligned}\widetilde{\rho}_0 - \widetilde{\rho}_1 &= 2 \sum_{m\in V_r^c} (-1)^{(m+(m+v_{r+1})+v_{r+1})\cdot i_\xi} |\eta_m\rangle\langle\eta_{m+v_{r+1}}| \\ &= 2 \sum_{m\in V_r^c} |\eta_m\rangle\langle\eta_{m+v_{r+1}}|.\end{aligned}$$

Therefore we conclude that

$$\frac{1}{2}\,\mathrm{tr}\,|\widetilde{\rho}_0 - \widetilde{\rho}_1| = \mathrm{tr}\left|\sum_{m\in V_r^c} |\eta_m\rangle\langle\eta_{m+v_{r+1}}|\right|. \tag{38}$$

By Lemma 3, $\langle\eta_m|\eta_n\rangle = 0$ if $m \neq n$ with $m, n \in V_r^c$. If we let $\langle\eta_m|\eta_m\rangle = d_{\eta_m}^2$ we get, $\sum_{m\in V_r^c} d_{\eta_m}^2 = 1$ by (36). Let us rewrite the $|\eta_m\rangle$ for $m \in V_r^c$ as $|\eta_m\rangle = d_{\eta_m}|\hat{\eta}_m\rangle$ with $\langle\hat{\eta}_m|\hat{\eta}_n\rangle = \delta_{m,n}$ for $m, n \in V_r^c$. It is known that for any operator A, $|A| = \sqrt{A^\dagger A}$ and thus[‖]

$$\begin{aligned}\mathrm{tr}\left|\sum_{m\in V_r^c} |\eta_m\rangle\langle\eta_{m+v_{r+1}}|\right| &= \mathrm{tr}\sqrt{\sum_{m\in V_r^c} |\eta_m\rangle\langle\eta_{m+v_{r+1}}| \sum_{m'\in V_r^c} |\eta_{m'+v_{r+1}}\rangle\langle\eta_{m'}|} \\ &= \mathrm{tr}\sqrt{\sum_{m,m'\in V_r^c} |\eta_m\rangle\langle\eta_{m+v_{r+1}}|\eta_{m'+v_{r+1}}\rangle\langle\eta_{m'}|} \\ &= \mathrm{tr}\sqrt{\sum_{m\in V_r^c} d_{\eta_{m+v_{r+1}}}^2 d_{\eta_m}^2 |\hat{\eta}_m\rangle\langle\hat{\eta}_m|} \\ &= \sum_{m\in V_r^c} d_{\eta_{m+v_{r+1}}} d_{\eta_m}\end{aligned}$$

where the last equation follows directly from the spectral decomposition that figures under the square root. Using the fact that $V_r^c = V_{r+1}^c \cup \left(v_{r+1} + V_{r+1}^c\right)$ and that this union is disjoint, we deduce

$$\frac{1}{2}\,\mathrm{tr}\,|\widetilde{\rho}_0 - \widetilde{\rho}_1| = 2 \sum_{m\in V_{r+1}^c} d_{\eta_m} d_{\eta_{m+v_{r+1}}}. \tag{39}$$

In order to bound this result we use the following Lemma,

[‖] Here A is Hermitian, therefore $|A| = \sqrt{AA^\dagger}$.

Lemma 5. *Let I be any set, $s : I \to I$ be such that $s^2 = \mathbf{1}_I$ and $p_i \geq 0$ with $\sum_{i \in I} p_i \leq 1$. Let $I' \subseteq I$ and $E \subseteq I$ such that $I' \cap s(I') = \emptyset$ and $I' \subseteq E \cup s(E)$; then*

$$\sum_{i \in I'} \sqrt{p_i p_{s(i)}} \leq \sqrt{\sum_{i \in E} p_i} \, .$$

Proof. For $i \in I'$, if $i \notin E$ let $h(i) = s(i) \in E$ and $h(s(i)) = i$, else let $h(i) = i \in E$ and $h(s(i)) = s(i)$. This function is well defined because i and $s(i)$ cannot be both in I'. Moreover $h(h(i)) = i$ and h is thus 1–1 on I'.

$$\sum_{i \in I'} \sqrt{p_i p_{s(i)}} = \sum_{i \in I'} \sqrt{p_{h(i)}} \sqrt{p_{s(h(i))}} \leq \sqrt{\sum_{i \in I'} p_{h(i)}} \sqrt{\sum_{i \in I'} p_{s(h(i))}} \leq \sqrt{\sum_{i \in E} p_i}$$

the first inequality being justified by Schwartz inequality. □

We now use the lemma. Let $I = V_r^c$, $I' = V_{r+1}^c$, $s(m) = m + v_{r+1}$; clearly $I' \cap s(I') = \emptyset$ and $s^2 = \mathbf{1}_I$. Let also $E = \{m \in I \mid d_H(m, V_r) \geq d_{r,1}/2\}$ where $d_{r,1}$ was defined as the smallest Hamming distance between v_{r+1} and the elements of V_r. For the lemma to apply, we need to show that $I' \subseteq E \cup s(E)$. If $m \in I'$ was such that $m \notin E$ and $m \notin s(E)$ then $s(m) \notin E$, $d_H(m, V_r) < d_{r,1}/2$ and $d_H(m + v_{r+1}, V_r) < d_{r,1}/2$; this implies $d_H(v_{r+1}, V_r) < d_{r,1}$, contrary to the definition of $d_{r,1}$. By the definition of E, if $c = m + v'$ for $m \in E$ and $v' \in V_r$ then $|c| \geq d_{r,1}/2$. Consequently, letting $p_m = d^2_{\eta_m}$ for $m \in I$, $\sum_{m \in I} p_m = 1$ and

$$\begin{aligned} \left(\frac{1}{2} \operatorname{tr} |\widetilde{\rho}_0 - \widetilde{\rho}_1| \right)^2 &= 4 \left(\sum_{m \in V_{r+1}^c} d_{\eta_m} d_{\eta_{m+v_{r+1}}} \right)^2 && \text{By (39)} \\ &\leq 4 \left(\sqrt{\sum_{m \in E} d^2_{\eta_m}} \right)^2 && \text{By Lemma 5} \\ &= 4 \sum_{\substack{m \in V_r^c \\ d_H(m, V_r) \geq d_{r,1}/2}} \sum_{v' \in V_r} d^2_{m+v'} \leq 4 \sum_{|c| \geq d_{r,1}/2} d^2_c . \end{aligned}$$

Using Lemma 2 (Eq. 34), we get

$$SD(\widehat{\rho}_0, \widehat{\rho}_1) \leq \frac{1}{2} \operatorname{tr} |\widetilde{\rho}_0 - \widetilde{\rho}_1| \leq 2 \sqrt{\sum_{|c| \geq d_{r,1}/2} d^2_c} . \qquad (40)$$

Note that this result is identical to the bound derived in [5, Lemma 4.5 (Eq. D.8)]. This result is much better than the loose bound [5, Lemma D.2 (Eq. D.3)] which is based on the methods of [4].

As a consequence, using (31) we get

$$SD(\widehat{\rho}_0, \widehat{\rho}_1) \leq 2 \sqrt{P \left[|\mathbf{C}_I| \geq d_{r,1}/2 \;\mid\; \mathbf{B}_I = \overline{\mathbf{b}'}, \mathbf{s} \right]} . \qquad (41)$$

3.6. Bounding Eve's accessible information

We now rewrite more carefully inequality (41) so as to be able to take into account all the parameters that were fixed and that we will now let vary in order to average Eve's information on the entire range of these parameters.

Let $\mathbf{c} = \mathbf{i} + \mathbf{i}^B$, the exclusive-or of the $2n$-bit string sent by Alice and of the one measured by Bob. Each index $1 \leq l \leq 2n$ such that $c_l = 1$ corresponds to a mismatch in Bob's bit value with respect to the value sent by Alice. If $s_l = 1$ the bit is an "information bit" and if $s_l = 0$ it is a "test bit". The corresponding "error on the information bits" is thus $\mathbf{c}_\mathbf{s}$ and the error on the test bits is $\mathbf{c}_{\bar{\mathbf{s}}}$. The random variable corresponding to $\mathbf{c}_\mathbf{s}$ and $\mathbf{c}_{\bar{\mathbf{s}}}$ are denoted $\mathbf{C}_I$ and $\mathbf{C}_T$ respectively; they depend on $\mathbf{b}$ and $\mathbf{s}$. In order to lighten the notation, we will write $P[\mathbf{C}_I = \mathbf{c}_\mathbf{s} \mid \mathbf{b}, \mathbf{s}]$ to mean the probability that the error string on the bits such that $s_i = 1$ be $\mathbf{c}_\mathbf{s}$ conditional to Alice having used the basis string $\mathbf{b}$ and the selection string $\mathbf{s}$. As a consequence, $P[\mathbf{C}_I = \mathbf{c}_\mathbf{s} \mid \mathbf{b} + \mathbf{s}, \mathbf{s}]$ denotes the probability that the error string on information bits be $\mathbf{c}_\mathbf{s}$ if the selection string is $\mathbf{s}$ and the basis string is $\mathbf{b} + \mathbf{s}$, i.e. is just the same as $\mathbf{b}$ but all the bases corresponding to the positions selected by $\mathbf{s}$ (of the information bits) are replaced by their conjuguates. Equations (31) and (41) can now be rewritten more cleanly as

$$d^2_{\mathbf{c}_\mathbf{s}} = P[\mathbf{C}_I = \mathbf{c}_\mathbf{s} \mid \mathbf{b} + \mathbf{s}, \mathbf{s}] = P[\mathbf{C}_I = \mathbf{c}_\mathbf{s} \mid \mathbf{c}_{\bar{\mathbf{s}}}, \mathbf{b} + \mathbf{s}, \mathbf{s}] \tag{42}$$

$$SD(\widehat{\rho}_0, \widehat{\rho}_1) \leq 2\sqrt{P\left[|\mathbf{C}_I| \geq d_{r,1}/2 \;\mid\; \mathbf{b} + \mathbf{s}, \mathbf{s}\right]} \tag{43}$$

where in the right hand side of (42) we use the fact that qubits are attacked independently, the error on information bits thus being independent of the error $\mathbf{c}_{\bar{\mathbf{s}}}$ on test bits. Equation (43) was derived for a (virtual) flat attack associated to $\mathbf{b}$. That flat attack had the same $\widehat{\rho}_0$ and $\widehat{\rho}_1$ as the original attack, and could only give a lower error rate in the conjuguate bases. As a consequence equation (43) also holds for the original attack U and from now on, the probability of error on the right-hand side will be understood to be the one induced by the original attack $U = U_1 \otimes \ldots \otimes U_{2n}$.

For any such fixed attack U, Eve's information depends only on the syndrome $\boldsymbol{\xi}$, the characteristic string $\mathbf{s}$ for the information bits, and the corresponding bases of the information string $\mathbf{b}_\mathbf{s}$ (yet, as said, we use the entire $2n$-bit string $\mathbf{b}$).

Corollary 6. *For a* 1*-bit key* k*,*

$$I(\mathbf{K}; \mathbf{E} \mid \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}, \mathbf{c}_{\bar{\mathbf{s}}}) = I(\mathbf{K}; \mathbf{E} \mid \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}) \leq 2\sqrt{P\left[\left(|\mathbf{C}_I| \geq \frac{d_{r,1}}{2}\right) \mid \mathbf{b} + \mathbf{s}, \mathbf{s}\right]} \tag{44}$$

where $\mathbf{K}$ *is the random variable giving as output key* k *and* $\mathbf{E}$ *is the random variable corresponding to the outputs of Eve's (optimal) measurement.*

Proof. This follows from the fact that $SD(\widehat{\rho}_0, \widehat{\rho}_1)$ is Eve's accessible information on k if she holds $\widehat{\rho}_k$ given by (32). These states correspond to Eve's state when Alice encodes the key-bit k assuming Eve learns $\boldsymbol{\xi}$, $\mathbf{b}$ and $\mathbf{s}$. Eve's information also depends in principle on $\mathbf{c}_{\bar{\mathbf{s}}}$ but since her attack on a qubit is independent of the other qubits, the bits of $\mathbf{c}_{\bar{\mathbf{s}}}$ have no influence on her state and may be omitted from the parameters on which Eve's information I depends. □

Proposition 7. *For an* m*-bit key* $\mathbf{k}$*,*

$$I(\mathbf{K}; \mathbf{E} \mid \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}, \mathbf{c}_{\bar{\mathbf{s}}}) = I(\mathbf{K}; \mathbf{E} \mid \mathbf{b}, \mathbf{s}, \boldsymbol{\xi}) \leq 2m\sqrt{P\left[\left(|\mathbf{C}_I| \geq \frac{d_{r,m}}{2}\right) \mid \mathbf{b} + \mathbf{s}, \mathbf{s}\right]}. \tag{45}$$

Proof. This follows from Corollary 6 by applying the chain rule for mutual information. Details of the proof can be found in [5, Section 4.5]. □

The value we want to bound is Eve's expected information, assuming Eve gets no information if the test fails, which happens when $\frac{|\mathbf{c}_{\bar{\mathbf{s}}}|}{n} > p_a$. If we let

$$I_{(p_a)}(\mathbf{K};\mathbf{E} \mid \mathbf{b},\mathbf{s},\boldsymbol{\xi},\mathbf{c}_{\bar{\mathbf{s}}}) = \begin{cases} I(\mathbf{K};\mathbf{E} \mid \mathbf{b},\mathbf{s},\boldsymbol{\xi}) & \text{if } \frac{|\mathbf{c}_{\bar{\mathbf{s}}}|}{n} \le p_a \\ 0 & \text{otherwise} \end{cases} \tag{46}$$

then the accessible information to bound, denoted[**]$\langle I_{\text{Eve}}^{(p_a)} \rangle$, is given by

$$\langle I_{\text{Eve}}^{(p_a)} \rangle = \sum_{\mathbf{b},\mathbf{s},\boldsymbol{\xi},\mathbf{c}_{\bar{\mathbf{s}}}} I_{(p_a)}(\mathbf{K};\mathbf{E} \mid \mathbf{b},\mathbf{s},\boldsymbol{\xi},\mathbf{c}_{\bar{\mathbf{s}}}) p(\mathbf{b},\mathbf{s},\boldsymbol{\xi},\mathbf{c}_{\bar{\mathbf{s}}}). \tag{47}$$

Theorem 8.

$$\langle I_{\text{Eve}}^{(p_a)} \rangle \le 2m \sqrt{P\left[\left(\frac{|\mathbf{C}_I|}{n} \ge \frac{d_{r,m}}{2n}\right) \wedge \left(\frac{|\mathbf{C}_T|}{n} \le p_a\right)\right]} \tag{48}$$

where $\frac{|\mathbf{C}_T|}{n}$ *is the random variable corresponding to the error rate on test bits and* $\frac{|\mathbf{C}_I|}{n}$ *is the random variable corresponding to the error rate on the information bits.*

Proof. The function x^2 is convex, i.e. $(\sum_i p_i x_i)^2 \le \sum_i p_i x_i^2$ for $p_i \ge 0$, $\sum_i p_i = 1$. We apply this to the square $\langle I_{\text{Eve}}^{(p_a)} \rangle^2$ of the information we want to bound.

$$\begin{aligned}
\langle I_{\text{Eve}}^{(p_a)} \rangle^2 &= \left[\sum_{\mathbf{b},|\mathbf{s}|=n,\boldsymbol{\xi},\mathbf{c}_{\bar{\mathbf{s}}}} I_{(p_a)}(\mathbf{K};\mathbf{E} \mid \mathbf{b},\mathbf{s},\boldsymbol{\xi},\mathbf{c}_{\bar{\mathbf{s}}}) p(\mathbf{b},\mathbf{s},\boldsymbol{\xi},\mathbf{c}_{\bar{\mathbf{s}}}) \right]^2 && \text{by (47)} \\
&\le \sum_{\mathbf{b},|\mathbf{s}|=n,\boldsymbol{\xi},\mathbf{c}_{\bar{\mathbf{s}}}} I_{(p_a)}^2(\mathbf{K};\mathbf{E} \mid \mathbf{b},\mathbf{s},\boldsymbol{\xi},\mathbf{c}_{\bar{\mathbf{s}}}) p(\mathbf{b},\mathbf{s},\boldsymbol{\xi},\mathbf{c}_{\bar{\mathbf{s}}}) && \text{by convexity of } x^2 \\
&\le \sum_{\mathbf{b},|\mathbf{s}|=n,\boldsymbol{\xi},\frac{|\mathbf{c}_{\bar{\mathbf{s}}}|}{n} \le p_a} I^2(\mathbf{K};\mathbf{E} \mid \mathbf{b},\mathbf{s},\boldsymbol{\xi}) p(\mathbf{b},\mathbf{s},\boldsymbol{\xi},\mathbf{c}_{\bar{\mathbf{s}}}) && \text{by (46)} \\
&\le 4m^2 \sum_{\mathbf{b},|\mathbf{s}|=n,\boldsymbol{\xi},\frac{|\mathbf{c}_{\bar{\mathbf{s}}}|}{n} \le p_a} P\left[\left(|\mathbf{C}_I| \ge \frac{d_{r,m}}{2}\right) \mid \mathbf{c}_{\bar{\mathbf{s}}},\mathbf{b}+\mathbf{s},\mathbf{s}\right] p(\mathbf{b},\mathbf{s},\boldsymbol{\xi},\mathbf{c}_{\bar{\mathbf{s}}}) && \text{by (45) and (42)} \\
&= 4m^2 \sum_{\mathbf{b},|\mathbf{s}|=n,\frac{|\mathbf{c}_{\bar{\mathbf{s}}}|}{n} \le p_a} P\left[\left(|\mathbf{C}_I| \ge \frac{d_{r,m}}{2}\right) \mid \mathbf{c}_{\bar{\mathbf{s}}},\mathbf{b}+\mathbf{s},\mathbf{s}\right] p(\mathbf{b},\mathbf{s},\mathbf{c}_{\bar{\mathbf{s}}}).
\end{aligned}$$

Since the test bits are unaffected by replacing the basis of the information bits:

$$p(\mathbf{b},\mathbf{s},\mathbf{c}_{\bar{\mathbf{s}}}) = p(\mathbf{c}_{\bar{\mathbf{s}}} \mid \mathbf{b},\mathbf{s}) p(\mathbf{b},\mathbf{s}) = p(\mathbf{c}_{\bar{\mathbf{s}}} \mid \mathbf{b}+\mathbf{s},\mathbf{s}) p(\mathbf{b},\mathbf{s}) = p(\mathbf{c}_{\bar{\mathbf{s}}} \mid \mathbf{b}+\mathbf{s},\mathbf{s}) p(\mathbf{b}+\mathbf{s},\mathbf{s}) = p(\mathbf{c}_{\bar{\mathbf{s}}},\mathbf{b}+\mathbf{s},\mathbf{s}),$$

and, letting $\tilde{\mathbf{b}} = \mathbf{b}+\mathbf{s}$,

$$\langle I_{\text{Eve}}^{(p_a)} \rangle^2 \le 4m^2 \sum_{\tilde{\mathbf{b}},|\mathbf{s}|=n,\frac{|\mathbf{c}_{\bar{\mathbf{s}}}|}{n} \le p_a} \left[\left(|\mathbf{C}_I| \ge \frac{d_{r,m}}{2}\right) \mid \mathbf{c}_{\bar{\mathbf{s}}},\tilde{\mathbf{b}},\mathbf{s}\right] p(\mathbf{c}_{\bar{\mathbf{s}}},\tilde{\mathbf{b}},\mathbf{s})$$

[**] The notation in [5] is $\langle I'_{\text{Eve}} \rangle$, the value p_a being fixed.

$$= 4m^2 \sum_{\tilde{\mathbf{b}},|\mathbf{s}|=n} P\left[\left(|\mathbf{C}_I| \geq \frac{d_{r,m}}{2}\right) \wedge \left(\frac{|\mathbf{C}_T|}{n} \leq p_a\right) \mid \tilde{\mathbf{b}}, \mathbf{s}\right] p(\tilde{\mathbf{b}}, \mathbf{s})$$

$$= 4m^2 P\left[\left(|\mathbf{C}_I| \geq \frac{d_{r,m}}{2}\right) \wedge \left(\frac{|\mathbf{C}_T|}{n} \leq p_a\right)\right] \tag{49}$$

□

3.7. *Proof of security*

Following the point of view of [5] we choose a code such that $\dfrac{d_{r,m}}{2n} > p_a + \epsilon$ for some ϵ; the right-hand side of (48) is then less than $P\left[\left(\dfrac{|\mathbf{C}_I|}{n} > p_a + \epsilon\right) \wedge \left(\dfrac{|\mathbf{C}_T|}{n} \leq p_a\right)\right]$ which itself is exponentially small in n. For each particular string $c_1 \ldots c_{2n}$ corresponding to a measurement of all qubits in some admissible basis $\mathbf{b}$ we can apply Hoeffding's sampling (Theorem 10). Let $\overline{X} = \frac{|\mathbf{C}_I|}{n}$ be the average of the sample corresponding to erroneous information bits; $\mu = \frac{|\mathbf{C}_I|+|\mathbf{C}_T|}{2n}$ is the expectancy of $\overline{X}$. $\frac{|\mathbf{C}_T|}{n} \leq p_a$ is equivalent to $2\mu - \overline{X} \leq p_a$, or equivalently, to $\overline{X} - \mu \geq \mu - p_a$. For the population $c_1, \ldots, c_{2n}$ the conditions $\left(\dfrac{|\mathbf{C}_I|}{n} > p_a + \epsilon\right)$ and $\left(\dfrac{|\mathbf{C}_T|}{n} \leq p_a\right)$ then rewrite to

$$\left(\overline{X} - \mu > \epsilon + p_a - \mu\right) \wedge \left(\overline{X} - \mu \geq \mu - p_a\right) \tag{50}$$

which implies $2(\overline{X} - \mu) > \epsilon$ and using Hoeffding's theorem (Theorem 10)

$$P\left[\left(\frac{|\mathbf{C}_I|}{n} > p_a + \epsilon\right) \wedge \left(\frac{|\mathbf{C}_T|}{n} \leq p_a\right)\right] \leq P\left[\overline{X} - \mu > \frac{\epsilon}{2}\right] \leq e^{-\frac{1}{2}n\epsilon^2}. \tag{51}$$

The above discussion gives the following

Theorem 9. *Let us be given $\delta > 0$, $R > 0$ and, for infinitely many values of n, a family $\{v_1^n, \ldots, v_{r_n+m_n}^n\}$ of linearly independent vectors in $\mathbf{F}_2^n$ such that $\delta \leq \frac{d_{r_n,m_n}}{n}$ and $\frac{m_n}{n} \leq R$. Then for any $p_a > 0$ and $\epsilon_{\text{sec}} > 0$ such that $p_a + \epsilon_{\text{sec}} \leq \frac{\delta}{2}$, Eve's accessible information satisfies the following bound*

$$\langle I_{Eve}^{(p_a)} \rangle \leq 2R\, n e^{-\frac{\epsilon_{\text{sec}}^2}{4} n}.$$

All we need to guarantee security is thus vectors $\{v_1^n, \ldots, v_{r_n+m_n}^n\}$ satisfying the conditions of the theorem. Such families were proven to exist in [5].

3.8. *Reliability*

For the key to be reliable, we need to be sure that the error rate on the information bits is less than the maximal rate that the error correcting code can handle. The maximum number of errors for our code will be fixed to $n(p_a + \epsilon_{\text{rel}})$. For the code to be reliable with exponentially small probability of failure, we need

$$P\left[\left(\frac{|\mathbf{C}_I|}{n} > p_a + \epsilon_{\text{rel}}\right) \wedge \left(\frac{|\mathbf{C}_T|}{n} \leq p_a\right)\right] \leq e^{-\frac{1}{2}n\epsilon_{\text{rel}}^2}.$$

For any fixed set of used bits, the test bits and the information bits is a random partition with two subsets of size n and the argument used in the previous section applies. The same requirement figures in [5].

4. Conclusions and Discussion

In this paper we have analyzed the security of the BB84 protocol against any collective attack using the methods and tools used in proving security against the more powerful joint attack. By doing this we maintain the security proof relatively simple, yet we achieve a far more meaningful result than previously achieved for the collective attack [4]. The basic idea of this paper can also be found in a presentation given by one of us (M.B.), at the Technion [7].

The same theorems (8 and 9) proven in this paper, are also obtained by [5] for the joint attack. This result leads to an asymptotic error-rate threshold of 7.56%,[††] the same asymptotic result obtained for the joint attack in [5, 8]. Note that these results are not just asymptotical but also explicit in the sense that for every ϵ and every threshold smaller than $(7.56 - \epsilon)$, a sufficiently large number n can explicitly be calculated such that the final key is reliable and secure. Explicit numbers expressing the reliability and security can also be obtained. To the best of our knowledge, such explicit results were not obtained via the methods shown in [9]. The threshold of 7.56% obtained here and in [5, 8] still has a gap from the asymptotical threshold of 11% reported by [9]. This gap can be explained by the different choice of privacy amplification, see for instance [5, 10, 11].

Other researchers also reached very interesting results regarding the collective attack and its relations to the joint attack, via other methods. See for instance [12, 13] in which it is proven that security against collective attacks implies security against joint attacks. However, their definition of the collective attack is not identical to the definition given in [2], which is used in [3, 4] and in the current paper. Furthermore, the conjecture that the strongest joint attack *is* a collective attack is not addressed by [12, 13] and remain an open problem. We leave the comparison of our result to the results obtained via these other methods for a future research.

References and Notes

1. Bennett, C.H.; Brassard, G. Quantum Cryptography: Public key distribution and coin tossing. *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* **1984**, pages 175–179.
2. Biham, E.; Mor, T. Security of quantum cryptography against collective attacks. *Physical Review Letters* **1997**, *78*, 2256–2259.
3. Biham, E.; Mor, T. Bounds on information and the security of quantum cryptography. *Physical Review Letters* **1997**, *79*, 4034–4037.
4. Biham, E.; Boyer, M.; Brassard, G.; van de Graaf, J.; Mor, T. Security of Quantum Key Distribution Against All Collective Attacks. *Algorithmica* **2002**, *34*, 372–388.
5. Biham, E.; Boyer, M.; Boykin, P.O.; Mor, T.; Roychowdhury, V.P. A proof of the security of quantum key distribution. *J. Cryptology* **2006**, *19*, 381–439.
6. Fuchs, C.A.; Peres, A. Quantum-state disturbance versus information gain: Uncertainty relations for quantum information. *Physical Review A* **1996**, *53*, 2038–2045.
7. Boyer, M. Security of the BB84 QKD protocol, **2005**. personal notes.
8. Mayers, D. Unconditional security in quantum cryptography. *J. ACM* **2001**, *48*, 351–406.

[††] We refer the reader to section 5 of [5] for detailed results and further discussion.

9. Shor, P.W.; Preskill, J. Simple proof of security of the BB84 quantum key distribution protocol. *Physical Review Letters* **2000**, *85*, 441–444.
10. Watanabe, S.; Matsumoto, R.; Uyematsu, T. Noise tolerance of the bb84 protocol with random privacy amplification. *International Journal of Quantum Information* **2006**, *4*, 935–946.
11. Molotkov, S.; Timofeev, A. Explicit attack on the key in quantum cryptography (BB84 protocol) reaching the theoretical error limit $Q_c \approx 11\%$. *JETP Letters* **2007**, *85*, 524–529.
12. Kraus, B.; Gisin, N.; Renner, R. Lower and upper bounds on the secret-key rate for quantum key distribution protocols using one-way classical communication. *Physical Review Letters* **2005**, *95*, 080501.
13. Renner, R. Security of Quantum Key Distribution. *Arxiv preprint quant-ph/0512258* **2005**.
14. Hoeffding, W. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* **1963**, *58*, 13–20.

A Hoeffding's theorem

Theorem 10 (Hoeffding 1963)**.** *Let $X_1, ..., X_n$ be either*

1. *independent random variables with finite first and second moments such that $a_i \leq X_i \leq b_i$ $(1 \leq i \leq n)$*

2. *or a random sample of size n without replacement taken from a population $c_1, ...c_N$ such that $a_i \leq c_i \leq b_i$ $(1 \leq i \leq N)$*

let $\overline{X} = (X_1 + ... + X_n)/n$ and $\mu = E[\overline{X}]$ be the expectancy of $\overline{X}$, then for any $\epsilon > 0$

$$\Pr\left[\overline{X} - \mu \geq \epsilon\right] \leq e^{-2n^2\epsilon^2/\sum_{i=1}^{n}(b_i-a_i)^2}.$$

In the same way $\Pr\left[\mu - \overline{X} \geq \epsilon\right] \leq e^{-2n^2\epsilon^2/\sum_{i=1}^{n}(b_i-a_i)^2}$. In case (2), $\mu = 1/N\sum_{i=1}^{N} c_i$, i.e. the expectancy of a sample mean is equal to the population mean. Theorem 10 can be found in [14].