Has Quantum cryptography been proven secure?

Tassos Nakassis*, J.C. Bienfang, P. Johnson, A. Mink, D. Rogers, X. Tang, C.J. Williams
National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899-8423

ABSTRACT

Quantum cryptography asserts that shared secrets can be established over public channels in such a way that the total information of an eavesdropper can be made arbitrarily small with probability arbitrarily close to 1. As we will show below, the current state of affairs, especially as it pertains to engineering issues, leaves something to be desired.

keywords: Quantum cryptography, Reconciliation, Privacy amplification, BB84, Rényi Entropy

1. INTRODUCTION AND BRIEF HISTORY

The purpose of this paper is to show that we need more rigorous results, more rigorous proofs, and more attention to engineering details. As a result, we will be forced, often by example, to examine the presumed weaknesses that must be addressed so that we may have provably secure systems. This is not a criticism of the papers cited; they are among the best published. But we have collectively failed to develop a precise formulation of the problem. As a result, the papers that focus on issues of Information theory often appear to overlook relevant aspects of the physical implementation and of the Reconciliation algorithm, while the articles focusing on channel implementation issues seem to rely on Information theory results whose assumptions are not demonstrably met.

We assume that the reader is familiar with the basic aspects of the BB84 protocol as described in the bibliography and will not duplicate this material. For a quick, well written description of BB84 he is directed to the easily available8 although, as it was written nearly a decade ago, it can serve only as an introduction.

To stay focused we will disregard all possible hardware/software shortcomings and we will adopt a minimalist view of the problem to be solved:

Alice and Bob have retained a sequence of communicated bits whose values at Alice are \{x[i] | i=1,2,..,N\} and due to interference by an eavesdropper, Eve, were received as \{y[i] | i=1,2,..,N\} by Bob. Alice and Bob will exchange information over a public channel that supports unbreakable integrity and authentication services. This information will allow reconciliation (i.e., flipping select bits of either x or y so that some sub-strings of x and y are, with great probability, identical). Alice and Bob will estimate a probabilistic upper bound on the amount of information Eve has gathered about outcome of reconciliation and will proceed to extract a string about which Eve will know very little with probability close to 1. To flesh out this rather generic description we shall further assume that:

* anakassis@nist.gov; phone 1 301 975-3632; fax 1 301 590 0932; nist.gov
Has Quantum cryptography been proven secure?

National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD, 20899

Approved for public release; distribution unlimited


see report

<table>
<thead>
<tr>
<th>SECURITY CLASSIFICATION OF:</th>
<th>ABSTRACT CLASSIFICATION OF:</th>
<th>LIMITATION OF ABSTRACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. REPORT</td>
<td>b. ABSTRACT</td>
<td>Same as Report (SAR)</td>
</tr>
<tr>
<td>unclassified</td>
<td>unclassified</td>
<td></td>
</tr>
<tr>
<td>c. THIS PAGE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>unclassified</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

unclassified

unclassified

unclassified

9
1) The values of the bit-string at Alice and their encoding can be seen as the repeated tossing of a perfect coin.
2) Eve is capable only of individual attacks on the transmitted qubits.
3) The observed errors allow Alice and Bob to infer meaningful probabilistic upper bounds on the amount of information Eve collected through tampering.
4) Reconciliation is achieved through parity exchanges.
5) The final secret is extracted through a linear transform.

Over time, the notion of an optimal attack by Eve has evolved, roughly following the stages below:

(I) Measure and Resend strategies in which Eve obtains deterministic information².
(II) Measure and Resend strategies in which Eve obtains probabilistic information³.
(III) Entanglement of the qubits and measurements of the collapsed systems after Bob/Alice announce the bases used¹₃, ¹⁴.
(IV) Channel profile modifications engineered by Eve, such as channels that provide better service to multiple photon emissions¹⁰.

If we carefully monitor and estimate the percentage of double-photon transmission events in the retained bit-stream, reconciliation and privacy amplification can proceed as previously, provided that the relevant parameters are treated as time evolving constants. Therefore, we shall pay minimal attention to this particular threat.

2. THE CASE OF DETERMINISTIC INFORMATION

Initially, it was believed that the best Eve could do is to try to guess the base used by Alice, measure the qubit in the chosen base, and forward to Bob the collapsed qubit. Belief in this system did not quite disappear even after the Breidbart base attack became known⁸.

In the context of such a model, privacy amplification is a breeze. Alice and Bob can count the number of errors. The a-priori probability that a measured qubit was read by Eve in the proper base would be 0.5. For all other measured qubits the value obtained by Eve would be totally unrelated to the encoded value and Bob’s measurement of the qubit would be in error with probability 0.5. Therefore, Alice and Bob can derive a probabilistic upper bound $U = U(e)$, such that Eve has measured correctly fewer than $U(e)$ qubits with probability that exceeds $1 - e$. If a similar upper bound $L = L(e)$ can be established for the multi-photon bits in the retained string, then Alice and Bob can observe that:

- The probability that the known bits exceed $U(e) + L(e)$ is less than $2e$.
- If $M$ is the number of the checksums exchanged and if $N - U(e) - L(e) - M - S$ random checksums are formed, then Eve knows nothing about the values of these checksums with probability that exceeds $1 - 2e - 2^{-S}$.

Simply put, each subset of $\{1, 2, \ldots, N\}$ can be identified with a vector of the vector space $V(N) = \{0, 1\}^N$ which consists of $2^N$ vectors. The values and checksums known to Eve can be seen as spanning a subspace of $V[N]$, $V_0$, whose dimension $D$ does not exceed $U(e) + L(e) + M$ with probability exceeding $1 - 2e$. It can be easily
shown that if we randomly choose N-D-S vectors in V, these vectors and the basis of \( V_D \) are linearly independent with probability exceeding \( 1 - 2^{-S} \). Linear independence implies that Eve has zero knowledge about the checksums that correspond to the random vectors in question.

3. BREIDBART BASES AND THE TREATMENT OF PROBABILISTIC INFORMATION

In short order, nevertheless, it was remarked that if one were using the Breidbart basis (essentially Eve used a measuring apparatus at an angle of \( \pi/8 \) with respect to the encoding bases), then the outcome of Eve’s measure-resend attack (in the Breidbart basis) would change the behavior of the measured bits as follows:

- tampered qubits would be in error in Bob’s string with probability 0.25 and
- a tampered qubit’s sent value would be known to Eve with probability \( \cos^2(\pi/8) = (1 + \cos(\pi/4))/2 = (2 + \sqrt{2})/4 \sim 0.85 \).

This type of attack necessitates that the privacy amplification mechanism be changed. Two new approaches were developed:

- The Big Brother (BB) approach which invented an entity that takes over Eve’s measuring apparatus. BB obtains deterministic readings of selected qubits and passes derivative information onto Eve in a way that mimics the Breidbart measurements. Any privacy amplification mechanism that would limit BB’s knowledge would ipso facto limit Eve’s.
- Reliance on existing results on hashing that are based on the properties of the Rényi entropy. In summary, the idea is that if our knowledge is represented by partly compromised N-bit strings, then – under the right conditions – we can map the N-bit strings onto K-bit strings in such a way that the K-bit strings are, more-or-less, equiprobable. More precisely:

1. The prevalent measure of lack of information, the Shannon entropy
   \[ H = -\sum p \log_2 (p), \] is an average; high H values do not necessarily bound individual probabilities. Of and by itself, a high H value cannot guide us on how to map n-bit strings onto highly secure k-bit strings. E.g., if some p-value equals 0.5, every deterministic mapping will produce some specific string with probability at least 0.5.
2. A more useful measure is the Rényi entropy. For simplicity we shall restrict ourselves to the Rényi entropy, of degree 2, by definition equal to
   \[ R(X) = H_2(X) = -\log_2 \left( \sum p^2 \right) \] where X represents some discrete variable and p the probability values of X’s distribution.
3. The Rényi entropy is used in conjunction with the notion of universal mappings. A set of mappings, G, from finite set A to finite set B (of cardinality \(|B|\)) is called Universal iff for any two distinct points in A, a1 and a2,
   \[ \text{Prob}\{g(a1)=g(a2)\mid g \in G\} \leq 1/|B| \] (assuming, of course, the uniform distribution for G).
4. If B consists of \( 2^k \) elements, if X is a random variable in A of Rényi entropy \( R(X) = r \), and if \( H(G(X)\mid G) \) and \( R(G(X)\mid G) \) represent the expected value of \( H(g(X)) \) and \( R(g(X)) \), then
   \[ H(G(X)\mid G) \geq R(G(X)\mid G) \geq k - 2^{k-r}/\ln(2), \] while for all g,
This is a powerful statement because it asserts that if \( K < r \), the mean value of either \( H(g(X)) \) or \( R(g(X)) \) is close to the maximum value either random variable may attain and, therefore, divergence from the mean is a low probability event. Nevertheless, the statement asserts an ensemble property that may or may not apply to a specific \( g \) and offers an a-priori low probability of failure.

To use this statement we need to have some reasonable lower bound on Eve's entropy after Reconciliation. This is addressed by the following two statements:

5. If \( W \) represents a random \( m \)-bit string, then
\[
R(X) - R(X \mid W=w) \leq 2m + s
\]
with probability not lower than \( 1 - 2^{-s} \)

6. If \( X, X \in A \), is the concatenation of \( N \) i.i.d. random bits and \( f \) a linear mapping from \( A \) onto the space of \( K \)-bit strings, then
\[
R(X) - R(X \mid f(X)=w) \leq K + o(N)
\]

The above results, possibly in combination, can be used to prove that a given algorithm will fail to produce shared secrets with small probability. But, the generality of the results cited entails a small price:

- Reliance, possibly unnecessary, on a priori probabilities when a feedback mechanism might have informed us that we had better abort. I.e., once a specific function \( g \) has been chosen out of a set of Universal mappings \( G \), it is the properties of \( g \) that matter, not the ensemble properties of \( G \).
- Asymptotic statements whose mapping to a well engineered system is left to the implementer, and
- A perception of the underlying problem that may not fully match the underlying reality.

Indeed, when the outcome of a measurement is probabilistic in nature, the level of our ignorance about the true value is a function not only of the measured value but, also, of the effects of Eve's actions on Bob. For instance, if the Breidbart attack is used, then the typical set of four tampered-with bits will consist of three bits for which the values of Bob and Alice agree and one for which they differ. Straightforward computations show that if Eve knows that a tampered bit did not cause an error, Eve has measured the correct value with probability \( (6 + 4\sqrt{2})/12 \approx 0.97 \). Indeed, the a priori probabilities that Bob's measurement matches the value sent are \( \cos^2(\pi/8)\cos^2(\pi/8) = 6 + 4\sqrt{2}/16 \) when Eve has also measured correctly and \( \sin^2(\pi/8) = 6 - 4\sqrt{2}/16 \) when Eve has measured incorrectly while the corresponding probabilities that Bob measured incorrectly are \( \cos^2(\pi/8)\sin^2(\pi/8) = 1/8 \) and \( \sin^2(\pi/8)\cos^2(\pi/8) = 1/8 \), both equal to \( 1/8 \). Thus, the a-posteriori probabilities for Eve are \( \{ \frac{6 + 4\sqrt{2}}{12}, \frac{6 - 4\sqrt{2}}{12} \} \) and \( \{0.5, 0.5\} \) so that in the end
\[
\frac{2 + \sqrt{2}}{4} = \frac{3}{4} \times \frac{6 + 4\sqrt{2}}{12} + \frac{1}{4} \times \frac{1}{2}
\]
is nothing but the average over two distinct populations.

The practical results of this situation are the following:
7. The definition of the Big Brother is in error. A truly undetectable BB must take into consideration the measurements of Bob. When such a measurement agrees with the value sent, BB either reads and passes to Eve the value sent with probability $2\sqrt{2}/3 \approx 0.94$ or he passes to Eve a uniformly distributed random value; ditto, if the measurement of Bob does not agree with the value sent by Eve.

8. Several entropy estimates, and consequently the size of the extractable key, are in error. As an example, let us look at a paper whose emphasis is on hardware and adopted the entropy estimates in the then extant bibliography. The paper cites and estimates the extractable bits as $F(N, \mu, \epsilon) = N(R(\mu, \epsilon) - 1.19 f(\epsilon)) - 1.19 f(\epsilon)$ with $R(\mu, \epsilon) = 1 - \mu - 4 \log_2(1.5)$. The term $f(\epsilon)$ is the Shannon entropy, the factor 1.19 reflects the overhead of the Cascade Reconciliation algorithm, $\mu$ is the percentage of multiple photon emissions, and $\epsilon$ the observed error frequency. If $\mu = 0$, then $R(\mu, \epsilon) = (1 - 4 \epsilon) + 4 \log_2(4/3)$, the average Rényi entropy per bit taking into account the information Eve gathered through tampering. Indeed, $(2 + \sqrt{2}/4)^2 + (2 - \sqrt{2}/4)^2 = 3/4$, hence the Rényi entropy equals $-\log_2(3/4) = \log_2(4/3)$ per tampered bit.

Alas, by the time Cascade has completed Eve knows where errors occurred. Instead of assigning to the typical foursome the Rényi entropy of $4 \log_2(4/3) \approx 1.66$, she will take into account the fact that wherever no error was discovered, she knows the transmitted value with probability $p = (6 + 4\sqrt{2})/12$ and that $R(p) = \log_2(18/17)$. So that the Rényi entropy per typical foursome will be $3R(p) + 1 \approx 0.25 + 1 = 1.25$.

Remark: We do not address in this paper the impact of Reconciliation on entropy. One should ask, however, why the results seemingly fail to match the theoretical results described above. One may think that this can be attributed to the information that flows back during reconciliation. Nevertheless, we could consider a gedanken experiment in which Alice is told by BB as to which of Bob’s bits are in error and lets Bob and the world know the correct values for these bits. The resultant entropy loss would exceed all bounds except for the $2(m+s)$ one which is generally considered to be unrealistically high. It would therefore appear that the extant theorems assert properties for some population $\mathcal{P}$ when our actions may be directing us towards sampling a subpopulation $\mathcal{P}^*$. I.e., the interactive algorithms, such as Cascade, inexorably guide the system toward information vectors $W$ within the very low-probability subset that we would rather avoid.

### 3. THE SLUTSKY PROBE AND BOUND.

Boris Slutsky suggested a different more efficient individual attack. If BB84 is inherently secure, we can endow Eve with whatever is physically possible even though it may be currently unavailable. E.g., with a quantum memory that will allow Eve to entangle the transmitted qubits with qubits of her own to be measured after Alice and Bob have matched bases and agreed which string to amplify. A long analysis follows that addresses individual qubit attacks and attempts to characterize Eve’s best strategy, its effects on the measured error rate and on the Rényi entropy, and the appropriate privacy amplification countermeasures (including detailed formulas that should adequately guide an implementer).

Are we done? Possibly, but we are still short of a complete proof. The cited papers assume that for every (expected) target error rate $E$ that Eve can expect Alice and Bob to witness, the best attack is the one that leads to maximum information gains. From this definition it follows that Eve’s optimal strategy is to entangle the qubits in a way that maps the Breidbart basis, $(|e>, |e>)$, onto $((|\Phi_0>, |e>) |\Phi_1>)$ where...
\( (|\Phi_{00}\rangle, |\Phi_{11}\rangle) = (\cos(\phi) |w_1\rangle + \sin(\phi) |w_2\rangle), (\sin(\phi) |w_1\rangle + \cos(\phi) |w_2\rangle) \), and \\
\( (|w_1\rangle, |w_2\rangle) \) is the orthonormal basis of a Hilbert space perpendicular to \((|e_0\rangle, |e_1\rangle)\).

If \((|u\rangle, |\bar{u}\rangle)\) is one of the bases Alice is using (obtained from the Breidbart basis through a counterclockwise rotation \(\alpha = \pi/8\)), straightforward manipulations of the Slutsky formulas show that Eve's meddling will map

\[
|u\rangle \rightarrow (\cos^2(\alpha) \cos(\phi) + \sin^2(\alpha) \sin(\phi)) |u\rangle |w_1\rangle + \\
(\cos^2(\alpha) \sin(\phi) + \sin^2(\alpha) \cos(\phi)) |u\rangle |w_2\rangle + \\
\cos(\alpha) \sin(\alpha) (\cos(\phi) + \sin(\phi)) |\bar{u}\rangle |w_1\rangle + \\
\cos(\alpha) \sin(\alpha) (\cos(\phi) - \sin(\phi)) |\bar{u}\rangle |w_2\rangle
\]

\[
|\bar{u}\rangle \rightarrow (\sin^2(\alpha) \cos(\phi) + \cos^2(\alpha) \sin(\phi)) |\bar{u}\rangle |w_1\rangle + \\
(\cos^2(\alpha) \cos(\phi) + \sin^2(\alpha) \sin(\phi)) |\bar{u}\rangle |w_2\rangle
\]

The entangled system for \(u\) shows that the squares of the coefficients sum up to 1, so that the probability that \(u\) will be measured as \(\bar{u}\) by Bob is

\[
E = 2 \cos^2(\alpha) \sin^2(\alpha) (\cos(\phi) - \sin(\phi))^2 = (1 - \sin(2\phi))/4
\]

In the event Eve learns that no error occurred, Eve concludes the entangled system collapsed, respectively, either onto

\[
(\cos^2(\alpha) \cos(\phi) + \sin^2(\alpha) \sin(\phi)) |w_1\rangle + (\cos^2(\alpha) \sin(\phi) + \sin^2(\alpha) \cos(\phi)) |w_2\rangle, \text{ or }
\]

\[
(\sin^2(\alpha) \cos(\phi) + \cos^2(\alpha) \sin(\phi)) |w_1\rangle + (\cos^2(\alpha) \cos(\phi) + \sin^2(\alpha) \sin(\phi)) |w_2\rangle.
\]

Inspection shows that the collapsed systems, are of the form \(A |w_1\rangle + B |w_2\rangle\) and \(B |w_1\rangle + A |w_2\rangle\), symmetrically located the \(w\)-plane, and forming equal angles, \(\zeta\), with the nearest base vectors. Therefore,

\[
16A^2 = 6 + 4\sqrt{2} \cos(2\phi) + 2\sin(2\phi),
\]

\[
16(A^2 + B^2) = 12 + 4\sin(2\phi), \text{ and }
\]

\[
\cos^2(\zeta) = \frac{A^2}{A^2 + B^2} = 0.5 + \frac{\sqrt{E(1 - 2E)}}{1 - E} \text{ (def) } \equiv 0.5 + d
\]

Thus, for small \(E\), \(d\) grows, roughly, as the square root of \(E\) while Eve's information gain

\[
(1 + \log_2 ((0.5 + d)^2 + (0.5 - d)^2) = \log_2 (1 + 4d^2)) \text{ is, roughly, proportional to } E. \text{ The square root of } E\]

probability increase suggests that for small \(E\) the relative cost of the BB approach will be high and the arithmetic data below show this to be the case. We also note that when Eve cannot distinguish between correct and incorrect measurements, the probability that Eve knows the value of a bit drops to

\[
0.5 + D = 0.5 + (1 - E)d.
\]

The optimal probe having been found \(^{13}\), the companion paper \(^{14}\) shows how to use the number of observed errors so as to derive probabilistic upper bounds on \(E\) and on the Rényi entropies of Eve before and after reconciliation.
Both papers are rigorous, well written. As they are also conservative in their estimates and err on the side of caution, careful examination of the remaining issues very probably will show that their conclusions and derived bounds are correct even if slightly incorrectly reasoned. Nevertheless, there are a couple issues that are not fully addressed. Namely,

- The paper implicitly assumes an interactive reconciliation scheme and at the very outset disregards the informational content of the bits in error (these values will be common knowledge once reconciliation is completed). But, as we saw above, the vector W(checksums released) is not necessarily random and most of the results used to estimate R(X|W) are not directly applicable. We note, however, that by disregarding the bits in error and by applying the a-posteriori probabilities, the papers probably sidestep the issue raised.
- While it is true that some Reconciliation schemes reveal information about the location of errors, others apparently do not. For example, recent work at BBN showed that sparse matrix error correction can be used for effective reconciliation. Thus, the bounds in would appear to unnecessarily penalize non-interactive reconciliation schemes.
- For reasons probably related to analytical tractability, the best strategy for Eve is defined as the one that maximizes Eve’s average gain in Rényi information. It is clear though that the best strategy for Eve is the one that maximizes Eve’s objectives. That is, Eve’s objective might be to maximize the probability that it will collect more than m bits worth of information once privacy amplification is completed. In other words, Eve may wish to beat the probabilistic upper bound of BB84 often and by a large (absolute or relative) margin. If the reconciled strings are large enough, this becomes synonymous with the best mean possible (the papers prove that a mixed strategy with fixed error rate E must lower the information gain). Nevertheless, a mixed strategy would seem to have a higher variance (fatter tails) and, therefore, for relatively short strings, it is not excluded that the gains in variance trump the losses due to a lowered mean value.

Interestingly enough, the convexity of the Defense frontier has been discussed and there is numerical support that the function displays the proper convexity profile.

- Finally, the referenced paper, like all the papers until recently, assumes that the mean number of errors in the transmitted qubits and the mean number of errors in the qubits retained are the same. Another assumption that has luckily bitten the dust with manageable protocol effects.

### 4. METHODS OF ATTACK AND THEIR IMPACT

For the reader to get a rough idea on how the different schemes by Eve and countermeasures affect the output of BB84, let us see the asymptotic information loss per bit as a function of the observed error rate E (assuming of course that it was all induced by Eve). We are giving the formula and its arithmetic value for E=0.03 side by side. The reader should note that,

- Measurements that cause errors impart no information to Eve,
- The probabilities cited are the probabilities that Eve correctly measured the value sent, and
- When BB mimics an apparatus that delivers the correct value with probability 0.5+x, BB must know the correct value with probability 2x.
Table 1. Average Entropy loss per bit (column 2) for different attack schemes and its value for E=0.03

<table>
<thead>
<tr>
<th>Measure Resend</th>
<th>2E</th>
<th>6%</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I) Breidbart, error location unknown</td>
<td>$4E \log_2 (1.5) \ (prob = \frac{2 + \sqrt{2}}{4})$</td>
<td>7%</td>
</tr>
<tr>
<td>(II) Breidbart, error location known</td>
<td>$3E(1 + \log_2 (17/18)) \cong 2.75E$ \ (prob $= \frac{3 + 2\sqrt{2}}{6}$)</td>
<td>8.25%</td>
</tr>
<tr>
<td>(III) Corresponding BB</td>
<td>$4E \frac{\sqrt{2}}{2}$ \ (case I) or $3E \frac{2\sqrt{2}}{3}$ \ case (II)</td>
<td>8.5%</td>
</tr>
<tr>
<td>(IV) Slutsky, error location unknown</td>
<td>$\log_2 (1 + 4D^2)$</td>
<td>15%</td>
</tr>
<tr>
<td>(V) Slutsky, error location known</td>
<td>$(1 - E) \log_2 (1 + 4d^2)$</td>
<td>15.4 %</td>
</tr>
<tr>
<td>(VI) Corresponding BB</td>
<td>$2D$ \ (case IV) or (1-E)2d \ (case V)</td>
<td>33%</td>
</tr>
</tbody>
</table>

In short, while the Breidbart attack delivered a minor blow to the BB approach, the Slutsky attack KO-ed it.

5. DECOY STATES

Apart from the fact that better proofs and more engineering detail are necessary, it is mildly disconcerting that the devotees of the provably secure system have more than once proven security by assuming physical models which were later debunked. Each time, large or small changes in privacy amplification are devised so that we would still have a provably secure system by sacrificing more of the bits previously deemed to be secure.

Recently, e.g. 10, the community discovered that the percentage of multiple-photon in the transmission stream does not necessarily coincide with their percentage in the sub-stream to be reconciled. Hence all the formulas that rely on the knowledge of Alice’s hardware in order to assess the impact of double photons in the stream are, potentially, incorrect and, therefore, link-monitoring protocols are necessary.

6. CONCLUSION

As of today, the quantum generation of shared-secrets has a single strength: unlike the extant schemes, its security cannot be compromised through scientific or engineering advances enabling new attacks on the communicated data. Therefore, it is highly desirable, if not necessary, that our proofs of this feature be as strong as the claim they serve. In particular, it is important that the engineering systems we build base their security on theorems whose assumptions are clearly satisfied by the systems in question.
ACKNOWLEDGEMENTS

This work was partly funded by the DARPA QuIST program.

REFERENCES