quant-ph/9712023 10 Dec 1997

A brief review on the impossibility of quantum bit commitment

Gilles Brassard Claude Crépeau
Université de Montréal *

Dominic Mayers
Princeton University †

Louis Salvail
BRICS ‡

December 10, 1997

Abstract

The desire to obtain an unconditionally secure bit commitment protocol in quantum cryptography was expressed for the first time thirteen years ago. Bit commitment is sufficient in quantum cryptography to realize a variety of applications with unconditional security. In 1993, a quantum bit commitment protocol was proposed together with a security proof. However, a basic flaw in the protocol was discovered by Mayers in 1995 and subsequently by Lo and Chau. Later the result was generalized by Mayers who showed that unconditionally secure bit commitment is impossible. A brief review on quantum bit commitment which focuses on the general impossibility theorem and on recent attempts to bypass this result is provided.

1 Introduction

After that Mayers obtained his general impossibility theorem for bit commitment (see the Appendix and [1, 2]), different kind of ideas were proposed by Brassard, Crépeau and Salvail with the hope to realise unconditionally secure bit commitment [3]. It was then realized by Mayers that these apparently promising ideas were also ruled out by his attack. These attempts contributed to enhance our understanding of what is going on with quantum bit commitment[4]. However, no complete discussion on the subject has ever been provided in the litterature.

Furthermore, two different proofs, each using a different approach, was provided by Mayers. The first approach was used in the original proof (see the Appendix and [1]) whereas the second approach appeared later in [2]. Despite all these results, and the related discussion by Lo and Chau [5], some quantum bit commitment protocols were recently proposed [6, 7] together with a claim of security that is ruled out by the general result. Fortunately, these claims [7] were not published. In fact, the protocols used the same idea previously described in [3, 4]. A brief history of the result with proper references to original work seems appropriate. We will not describe the proofs again (except in the Appendix which contains the original proof of Mayers). Our objective is to create a wholeness for the different papers written on the subject. We will also discuss the general theorem in the context of the specific ideas and schemes [3] which researchers have tried to realize quantum bit commitment despite this general theorem.

Before we proceed, let us briefly explain the notion of bit commitment and its impact in quantum cryptography. Quantum cryptography is often associated with a cryptographic application called key distribution [8, 9] and it has achieved success in this area [3]. However,

* Département IRO, Université de Montréal, C.P. 6128, succursale centre-ville,
Montréal (Québec), Canada H3C 3J7. e-mail: {brassard,crepeau}@iro.umontreal.ca.

†Department of Computer Science, Princeton University,
Princeton, NJ 08544-2087, U.S.A.. e-mail: mayers@cs.princeton.edu.

‡BRICS, Basic Research in Computer Science of the Danish National Research Foundation, Department of Computer Science,University of Århus, Ny Munkegade, building 540, DK-8000 Århus C, Denmark. e-mail: salvail@daimi.aau.dk

other applications of quantum mechanics to cryptography have also been considered and bit commitment was at the basis of most if not all of these other applications [3, 10, 11, 12]. A *bit commitment* scheme allows $\mathcal{A}lice$ to send something to $\mathcal{B}ob$ that commits her to a bit b of her choice in such a way that $\mathcal{B}ob$ cannot tell what b is, but such that $\mathcal{A}lice$ can later prove him what b originally was. You may think of this as $\mathcal{A}lice$ sending a note with the value b written on it in a strongbox to $\mathcal{B}ob$ and later revealing him the combination to the safe.

$\mathcal{A}lice$ can choose the distribution of probability of b during the commit phase. The commitment obtained after the commit phase is binding if $\mathcal{A}lice$ cannot change this distribution of probability and it is concealing if $\mathcal{B}ob$ cannot obtain any information about b without the help of $\mathcal{A}lice$. The commitent is secure if it is binding and concealing. The commitment is unconditionally secure if it is secure against a cheater, either $\mathcal{A}lice$ or $\mathcal{B}ob$, with unlimited technology and computational power. In 1993 a protocol for quantum bit commitment, henceforth referred to as BCJL, was claimed to be *provably secure* [11], that is, the resulting commitments were thought to be unconditionally secure. Because of quantum bit commitment, the future of quantum cryptography was very bright, with new applications such as the identification protocol of Crépeau and Salvail [13] coming up regularly.

The trouble began in October 1995 when Mayers found a subtle flaw in the BCJL protocol. Though Mayers explained his discovery to many researchers interested in quantum bit commitment [14], his result was not made entirely public until after Lo and Chau discovered independently a similar result in March 1996 [15]. The result of Mayers was more general than the one obtained by Lo and Chau, but both used the same basic idea. The result of Lo and Chau did not encompass the BCJL protocol in which $\mathcal{B}ob$ can obtain an exponentially small amount of information. (In practice a protocol is considered secure as long as $\mathcal{B}ob$ cannot obtain more than an exponentially small amount of information on the bit commited by $\mathcal{A}lice$, that is, an amount of information that goes exponentially fast to 0 as the number of photons used in the protocol increases.) However, the final version published by Lo and Chau [15] used the techniques previously used by Mayers [14] to prove the non security of the BCJL protocol and any other protocol published at the time. So, the paper of Lo and Chau [15] is a proper account of these preliminary results.

2 The general impossibility theorem

Now, we review the general theorem [2] (see also the Appendix) which says that a quantum protocol which creates an unconditionally secure bit commitment is simply impossible. The main additionnal difficulty in the general result is that it is easy to think that measurements and classical communication could be used to restrict the behavior of the cheater during the commit phase, and thus obtain a secure bit commitment. In fact, after BCJL was shown not secure, the spontaneous attitude was to try alternative quantum bit commitment protocols by making some clever use of measurements and classical communication [16]. Some of these protocols were proposed after that Mayers obtained the general result in March 1996 (just a little bit after Lo and Chau discovered their restricted result independently). All of these protocols were found not secure by Mayers.

There exists two approaches to deal with measurements and classical communication in quantum bit commitment protocols: an indirect approach and a direct approach. In the first proof written by Mayers (see the Appendix) the indirect approach was used. It was shown that any protocol in which classical information is used is equivalent to another protocol in which no classical information at all is used. Then it was shown that no protocol of the latter kind is unconditionally secure. The first step of this indirect proof is called a reduction in computer science. The advantage of this approach is that, after the reduction is shown, the attack on the new protocol is easy to describe and analyse because there is no classical communication anymore. The disadvantage is that we don't deal directly with the issue of classical communication and measurements, that is, the attack obtained against the new protocol is not the one that applies on the original protocol. The attack in the new protocol does not include any classical communication, whereas in the original protocol the cheater must communicate classically with the honest participant (otherwise this honest participant will wonder what is going on).

We emphasize that the proof of the reduction which is not that hard must nevertheless explain why the cheater can still cheat in the original protocol despite the fact that he is restricted by measurements and decoherence which must occur because of classical communication. Otherwise the overall proof would simply miss the important issue of classical communication – it would not encompass the protocols and ideas that have been proposed recently [3, 4, 6, 7]. Because this issue was somehow confusing, Mayers prefered to use a more direct approach without reduction in [2]. So, the paper [2] directly describes and analyses the real attack that must be executed by the cheater.

Lo and Chau also wrote a paper [5] to discuss the issue of quantum communication and other aspects of Mayers's result. They used a variant of Yao's model for quantum communication. The essence of Yao's model is that a third system is passed back and fourth under the control of each participant at their turn [12]. Mayers's attack works fine in this model, and it is indeed important to verify that the attack works in such a reasonable model. With regard to classical communication, the discussion of Lo and Chau [5] is similar to the indirect approach.

Now, let us consider the attack. Of course, we are interested in the attack on the original protocol. The attack on the new protocol is just a construction in a proof. We emphasize that in both approaches, with a reduction or without a reduction, the attack on the original protocol is the same. Here we focus on the part of the attack which must be executed during the commit phase. (The remainder of the attack which is executed after the commit phase is the same as when there is no classical communication, so it creates no additional difficulty.) One ingredient in the attack is that the cheater keeps every thing at the quantum level except what must be announced classically. Assume that at some given stage of the commit phase, a participant has normally generated a classical random variable R, executed measurements to obtain an overall outcome X, and shared some classical information Y with the other participant as a result of previous communication. Now, assume that this participant is the cheater and that the protocol says he must transmit some classical information $f(X, R, Y)$, which for simplicity we assume is a binary string. One might think that the cheater must have generated X and executed the measurements, or at the least some of them, in order to be able to compute and send $f(X, R, Y)$. However, the cheater does not have to do that. He can do the entire computation of $f(X, R, Y)$, including the computation of X and the measurements, at the quantum level. Only Y needs to be classical. Then he can measure the bits of the string $f(X, R, Y)$ (only these bits) and send them to the other participant. An example is given in section 3. The final result is that every information is kept at the quantum level, except what must be sent classically to the other participant. As explained in [1, 2] (see also the Appendix) this strategy executed during the commit phase either allows $\mathcal{B}ob$ to obtain some information about the bit commited by $\mathcal{A}lice$, without any help from $\mathcal{A}lice$, or else allows $\mathcal{A}lice$ to change her mind after the commit phase (as in the example of section 3).

This is not the end of the story. After that the above argument was understood, Crépeau proposed a quantum protocol [3, 4] that uses a computationally secure classical bit commitment [17, 18] as a subprotocol. The idea was to rely temporarily on the limitation (in speed) on the cheater during the commit phase to force him to execute some measurements (or restrict his behavior in some other way) in order to obtain a secure bit commitment. The hope was that this short-term assumption could be dropped after the commit phase so as to obtain a quantum bit commitment not relying on any long-term assumption. The same idea was recently used by Kent in [7]. Salvail also proposed a protocol in which two participants, $\mathcal{A}lice$ and $\mathcal{A}lbert$ say, want to commit a bit to $\mathcal{B}ob$. $\mathcal{A}lice$ and $\mathcal{A}lbert$ are sufficiently far apart that they cannot communicate during the commit phase. Again the hope was that this temporary restriction on the cheaters during the commit phase would be sufficient to obtain a secure quantum bit commitment not relying on any long-term assumption.

However, after some thoughts, one realize that the cheater in Mayers's attack executes the honest algorithm, the only difference is that he executes this honest algorithm at the quantum level. Therefore, if the cheater has the power to execute the honest protocol (which he must have) and has the technology to store information at the quantum level, then he has the power to cheat during the commit phase, despite the fact that he has not the power to break the computationally secure bit commitment efficiently, or despite the fact that $\mathcal{A}lice$ and $\mathcal{A}lbert$ cannot communicate during the commit phase. After the commit phase, the rule of the game is that we must drop the assumption on the computational power of the cheater, so the fact that a computationally secure bit commitment was used is irrelevant: the proof applies.

3 An Example:How to Break Kent's Protocol

In this section we illustrate the discussion of the previous section by a concrete example. We shall show how to break Kent's proposal [6, 7] for a quantum bit commitment scheme using a time-bounded computational assumption. The paper [7] describes two constructions for such a scheme, one allows $\mathcal{A}lice$ to commit and the other allows $\mathcal{B}ob$ to commit permanently. In this section we break the protocol allowing $\mathcal{A}lice$ to commit permanently. The other version can be broken by a similar attack.

Kent's protocol [7] uses a classical and unconditionally hidding bit commitment scheme. The hope is that this classical scheme will constraint $\mathcal{A}lice$ to transmit q-bits in pure states. The protocol uses the BB84 coding scheme: $\Psi(0,0) = |0\rangle_+, \Psi(0,1) = |1\rangle_+, \Psi(1,0) = |0\rangle_\times, \Psi(1,1) = |1\rangle_\times$. The first bit corresponds to the basis and the second bits to the encoded bit. Here is the essential idea behind Kent's protocol. For each $i = 1, \ldots, N_B$, $\mathcal{A}lice$ picks a random pair $(x, z) = (x_i, z_i) \in \{0,1\}^2$, sends a photon π_i in the BB84 state $\Psi(x, z)$ and execute a classical bit commitment $BC(x, z)$ according to the above classical bit commitment scheme. We denote[1] $\theta_i = \{+, \times\}_{[x]}$ the basis used by $\mathcal{A}lice$ for the photon π_i. $\mathcal{B}ob$ then picks a random sample $X \subseteq \{1, \ldots, N_B\}$ of size $N_B - N$. For each $i \in X$, $\mathcal{B}ob$ asks $\mathcal{A}lice$ to unveil $(x, z) \in \{0,1\}^2$ corresponding to the committed pair of classical bits in $BC(x, z)$. $\mathcal{B}ob$ then measures π_i in basis $\theta_i = \{+, \times\}_{[x]}$ and verifies that the observed outcome is indeed $z \in \{0,1\}$. The idea behind the remainder of Kent's protocol is very similar to the first bit commitment scheme ever proposed by [8]. The difference is that in [8] $\mathcal{A}lice$ picks the same value for all x_i, that is, the string of bases used by $\mathcal{A}lice$ is either $++\ldots+$ or $\times\times\ldots\times$. (See also [2, 15] for a description and analysis of this protocol.) The basic idea is that the committed bit is encoded in the transmission basis for each photon π_i. In Kent's protocol, if $\mathcal{A}lice$ wants to commit bit b, she announces $x_i \oplus b$ for each $i \in Y = \{1, \ldots, N_B\} - X$. So, the bit is commited in the choice of basis used by $\mathcal{A}lice$ for each $i \in Y$.

The scheme is unconditionnally hidding because no information about the transmission basis can be obtained from any photon π_i since the density matrix corresponding to the transmission in rectilinear basis

$$\rho_+ = \frac{1}{2}|0\rangle_+\langle 0|_+ + \frac{1}{2}|1\rangle_+\langle 1|_+$$

and the one corresponding to the transmission in diagonal basis

$$\rho_\times = \frac{1}{2}|0\rangle_\times\langle 0|_\times + \frac{1}{2}|1\rangle_\times\langle 1|_\times$$

are such that $\rho_+ = \rho_\times$.

Clearly, if sends the pure states $\Psi(x_i, z_i)$, she cannot claim that she used the other basis, that is, the one associated with $x_i \oplus 1$, for each $i \in Y$. So, if really $\mathcal{A}lice$ has sent the pure states $\Psi(x_i, z_i)$, the protocol should be binding. $\mathcal{A}lice$ can cheat in the original protocol of [8] by sending EPR pairs rather than a mixture of BB84 quantum states (see [8] for more details). So the resulting commitment is not binding. In Kent's protocol, if $\mathcal{A}lice$ cannot break the computational assumption during this test phase (between the time the commitments have been sent and the time they are opened), it is argued that $\mathcal{B}ob$ gets convinced that almost all

[1] Notation $\{a, b\}_{[s]}$ for $s \in \{0,1\}$ is a if $s = 0$ and b if $s = 1$.

photons π_i in Y are in the pure states $\Psi(x_i, z_i)$. Indeed, if this was true then the protocol would also be binding. However, we show that it is not the case.

3.1 The Classical Commitment

Let us first model the classical and unconditionally hiding commitment scheme by four one-way permutations [2] $f_{00}, f_{01}, f_{10}, f_{11} : \{0,1\}^n \rightarrow \{0,1\}^n$ for any integer n. In the remaining, functions f_{00}, f_{01}, f_{10} and f_{11} need not to be distinct. To commit (x, z), $\mathcal{A}lice$ picks a random uniformly distributed $w \in \{0,1\}^n$ and sends $y = f_{xz}(w)$ to $\mathcal{B}ob$. We obtain that y, the piece of evidence that $\mathcal{A}lice$ gives to $\mathcal{B}ob$ in order to commit on a pair of classical bit (x, z), is a random element uniformly distributed in $\{0,1\}^n$. Here are the properties that we need to consider.

1. The functions f_{xz} are efficiently computable and publicly known.
2. Given $y = f_{xz}(w)$ no information on (x, z) is known by $\mathcal{B}ob$ (thus the protocol is unconditionally hiding).
3. $\mathcal{A}lice$ knows only one $(x, z, w) \in \{0,1\}^2 \times \{0,1\}^n$ such that $f_{xz}(w) = y$. If she manages to find another $(x', z', w') \in \{0,1\}^2 \times \{0,1\}^n$ distinct from (x, z, w) such that $f_{x'z'}(w') = y$ then she can break the computational assumption (because necessarily $(x, z) \neq (x', z')$).

We shall see that the above conditions for classical commitment, in particular the first two conditions, implies that the proposed method cannot ensure $\mathcal{B}ob$ that most of the remaining q-bits are in pure states. We have described a particular classical bit commitment scheme, but Mayers's attack works with any other classical bit commitment scheme. In the next two subsections we describe $\mathcal{A}lice$'s attacks during the commit phase, then in the third subsection we explain how $\mathcal{A}lice$ can change her mind after the commit phase.

3.2 Alice's Preparation

If $\mathcal{A}lice$ wants to cheat the proposed protocol, as we will see, she has only to send entangled states rather than a mixture of BB84 states. In Kent's protocol, the use of a classical bit commitment scheme is intended to rule out the EPR attack. However, other entanglements can do the job. Let us consider the state $|\gamma(\theta)\rangle$ defined upon [3] $\theta \in \{+, \times\}$ as [4]

$$|\gamma(\theta)\rangle = \frac{1}{\sqrt{2^{n+1}}} \sum_{w \in \{0,1\}^n} |w\rangle |f_{\theta 0}(w)\rangle |0\rangle_\theta |0\rangle_\theta + |w\rangle |f_{\theta 1}(w)\rangle |1\rangle_\theta |1\rangle_\theta. \tag{1}$$

Mayers's theorem also specifies that there should be a superposition over x (or equivalently over θ). The idea is that every random choice, including the choice of the bases, must be done at the quantum level. However, this part of the superposition would collapse immediately because $\mathcal{A}lice$ must announce $x_i \oplus b$, for the classical bit b she has chosen (this is what is specified by the attack). So for simplicity we ignore this part of the superposition.

The state (1) can be efficiently constructed from condition 1 about the classical commitment scheme. The state $|\gamma(\theta)\rangle$ is made out of four registers which we denote from left to right as r_w, r_f, r_z^A and r_z^B. Now suppose $\mathcal{A}lice$ sends the register r_z^B to $\mathcal{B}ob$ instead of a random BB84 pure state. We assume the more general case where $\mathcal{B}ob$ does not measure the received quantum states until the pure states verifications take place. This allows a more reliable test than measuring immediately after reception and testing later on. Let H_A be the Hilbert space

[2] The same kind of argument can also be formalized for general one-way functions rather than one-way permutations. However, no classical and unconditionally bidding bit commitment scheme is yet known to be based only on the existence of one-way functions.

[3] In the following we sometime consider $\theta \in \{+, \times\}$ as being the bit x such that $\theta = \{+, \times\}_{[x]}$. Notations $f_{+z}(x)$ means $f_{0z}(x)$ and $f_{\times z}(x)$ means $f_{1z}(x)$.

[4] When a quantum state is written as $|w\rangle$ for $w \in \{0,1\}^n$ we mean $|w_1\rangle_+ \otimes \ldots \otimes |w_n\rangle_+$.

for registers r_w, r_f and r_z^A and let H_B be the Hilbert space for register r_z^B. By construction, $\mathcal{B}ob$ receives a mixture with density matrix:

$$\rho_B = \mathrm{Tr}_{H_A}(|\gamma(\theta)\rangle\langle\gamma(\theta)|) = \rho_\theta. \tag{2}$$

$\mathcal{A}lice$'s preparation consists of N_B systems $s_1, \ldots, s_{N_B}$ in quantum states $|\gamma(\theta_1)\rangle, \ldots, |\gamma(\theta_{N_B})\rangle$ for $\theta_i \in_R \{+.\times\}$. She sends to $\mathcal{B}ob$ the r_z^B registers for all systems $s_1, \ldots, s_{N_B}$.

3.3 How Alice Deals With Classical Communication

Suppose $\mathcal{A}lice$ has sent all N_B registers r_z^B to $\mathcal{B}ob$. Let $\theta_1, \ldots, \theta_{N_B}$ be the N_B bases picked in $\{+, \times\}$ in order to prepare the states $|\gamma(\theta_1)\rangle, \ldots, |\gamma(\theta_N)\rangle$. (From $\mathcal{A}lice$'s point of view these bases, i.e. the x_i, are not random anymore.) To execute the classical commitment, $\mathcal{A}lice$ must send the classical information $f_{xz}(w)$. Thus far, the values of z and w are not fixed: they are still in superposition. As explained in the previous section, $\mathcal{A}lice$ does not have to obtain w nor z classically to compute $f_{xz}(w)$. For committing, $\mathcal{A}lice$ simply measures in rectilinear basis all registers r_f. She then announces to $\mathcal{B}ob$, for each $i \in \{1, \ldots, N\}$, the result y_i. That is the way Mayers's attack works. Each system s_i is now in state

$$|\gamma'(\theta_i)\rangle = \frac{1}{\sqrt{2}}\left(|w\rangle|y_i\rangle|0\rangle_{\theta_i}|0\rangle_{\theta_i} + |w'\rangle|y_i\rangle|1\rangle_{\theta_i}|1\rangle_{\theta_i}\right)$$

where $w = f^{-1}_{\theta_i 0}(y_i)$ and $w' = f^{-1}_{\theta_i 1}(y_i)$. The above state is guaranteed to occur by property 2 of the classical commitment scheme, and the fact that w is uniquely determined by x (or θ), z and y.

Now suppose $\mathcal{B}ob$ asks $\mathcal{A}lice$ to unveil the commitment for some position $i \in X$. $\mathcal{A}lice$ simply measures registers r_w (in basis $+$) and r_z^A (in basis θ_i) for the system s_i. Let w and z be the outcomes of the measurement. $\mathcal{A}lice$ announces w, z and x_i to $\mathcal{B}ob$. $\mathcal{B}ob$ always verifies that $y_i = f_{x_i z}(w)$. The system ends up in state

$$|\gamma''(\theta_i)\rangle = |w\rangle|y_i\rangle|z\rangle_{\theta_i}|z\rangle_{\theta_i}$$

which leads to a successful verification by $\mathcal{B}ob$. Clearly $\mathcal{A}lice$ can always pass the test without breaking the computational assumption. The main point is that $\mathcal{A}lice$ executes the honest protocol at the quantum level, so any computational bound is useless. It follows that Kent's verification procedure is not a verification that almost all received q-bits are in pure states since equation 2 is obviously not the description of a pure state.

3.4 Breaking the Quantum Scheme

We now show how $\mathcal{A}lice$ can decide freely the bit she wants to unveil. We recall that, at this point, the rule of the game is that all computational assumptions must be dropped. (Otherwise we only have a computationally secure bit commitment, and this can already be done classically.) After the verification procedure only the remaining systems s_i with $i \in Y = \{1, \ldots, N\} \setminus X$ are used. In order to break the quantum protocol it is sufficient to show how $\mathcal{A}lice$ can choose the transmission basis for all photons transmitted to $\mathcal{B}ob$. For all $i \in Y$ the system s_i is in state (we remove the r_f register since it is no more entangled but we remember its observed value y_i):

$$|\gamma'(\theta_i)\rangle = \frac{1}{\sqrt{2}}\left(|w\rangle|0\rangle_{\theta_i}|0\rangle_{\theta_i} + |w'\rangle|1\rangle_{\theta_i}|1\rangle_{\theta_i}\right).$$

To cheat, $\mathcal{A}lice$ must disentangle the register r_w and obtain the state

$$|\gamma''(\theta_i)\rangle = \frac{1}{\sqrt{2}}\left(|w_0\rangle|0\rangle_{\theta_i}|0\rangle_{\theta_i} + |w_0\rangle|1\rangle_{\theta_i}|1\rangle_{\theta_i}\right).$$

where w_0 is some fixed string. If we ignore the disentangled registers r_f and r_w, this state is essentially an EPR pair (modulo a unitary transformation on $\mathcal{A}lice$'s side). So $\mathcal{A}lice$ can cheat as in the original attack defined in [8]. Now, we show how $\mathcal{A}lice$ can disentangle r_w. A simple way to disentangle r_w would be to replace both w and w' by the same output $f_{\theta_i 0}(w) = f_{\theta_i 1}(w') = y_i$. This is a reversible computation executed in the computational basis defined by θ_i for r_z^A and $+\ldots+$ for r_w, so it corresponds to a unitary transformation. This answers the question. However, this answer is somehow misleading because it gives the impression that the attack is as simple as the computation of $f_{\theta_i 0}(w) = f_{\theta_i 1}(w') = y_i$. The problem is that one must still explain how the two distinct inputs w and w' can be replaced by one and the same value y_i. Here we show how this can be done by $\mathcal{A}lice$ if she can inverse the functions $f_{\theta_i z}$. Because she knows y_i and θ_i, she can compute $w = f^{-1}_{\theta_i 0}(y_i)$ associated with $|0\rangle_{\theta_i}$ and $w' = f^{-1}_{\theta_i 1}(y_i)$ associated with $|1\rangle_{\theta_i}$, so she can "erase" the register r_w, that is, she can set this register to $\mathbf{0}$ by a bit-wise addition modulo 2. (She can also set it to any other value she wants, including y_i.) This concludes the description of the attack.

Although $\mathcal{A}lice$ breaks the computational assumption (i.e. inverse the functions f_{xz}) in order to unveil the bit she wishes, this cannot be used as a building block for a secure quantum bit commitment where the computational assumption is no more needed after some time. This is for exactly the same reason than the one allowing to conclude that no quantum bit commitment can be built from a classical computational assumption.

4 Conclusions

The first proof provided for the impossibility of bit commitment (see the Appendix) has completely obliterated the possibility of creating an unconditionally secure bit commitment. However, the attack was only indirectly described. Subsequently, specific attempts to by-pass this general result were proposed[3, 4]. This has shed more light on the nature of the attack which was finally described explicitly in [2]. Our goal here was to provide an analysis of this general attack in the context of a specific example, and to create a wholeness for the different papers published on the subject. The big lesson to learn from all this is that quantum information is always more elusive than its classical counterpart: extra care must be taken when reasoning about quantum cryptographic protocols and analyzing them. We hope that this paper will help to clarify the issue of the impossibility of bit commitment in its full generality.

Appendix

This appendix contains the original proof written by Mayers and sent to few researchers by email on March 14 1996. A modified version of the proof, which also used a reduction, was published in [1]. A direct proof with no reduction was published later in [2].

Abstract

It is currently known that the 1993 BCJL protocol of Brassard, Crépeau, Jozsa and Langlois (BCJL) is insecure. Here we provide the first proof that, not only this protocol, but any quantum bit commitment is either insecure against $\mathcal{A}lice$ or insecure against $\mathcal{B}ob$.

1 Introduction

The fact that the quantum bit commitment protocol of Brassard, Crépeau, Jozsa and Langlois [11] is insecure is known for quite sometime [14]. Lo and Chau have also independently shown that a restricted category of quantum bit commitments is insecure [15]. Now, we provide the first proof that not only these quantum bit commitment protocols, but any other quantum bit commitment protocol is insecure.

The absence of quantum bit commitment is a serious concern because other quantum protocols such as quantum oblivious transfer depend on the security of bit commitment [10, 19, 20, 12]. On the other hand, not all of Quantum Cryptography fall apart because our earlier proof of security for quantum key distribution [21] holds even if secure quantum bit commitment is not possible despite the fact that it is based on an earlier "proof" of security for quantum oblivious transfer that fails in the absence of a secure bit commitment scheme. The reason is that the proof of security for quantum key distribution does not depend on the security of quantum oblivious transfer, but rather on the (correct) proof that quantum oblivious transfer would be secure if implemented on top of a secure bit commitment scheme.

2 Bit Commitment

Any cryptographic task defines the relationship between inputs and outputs respectively entered and received by the task's participants. In bit commitment, $\mathcal{A}lice$ enters a bit b. At a later time, $\mathcal{B}ob$ may request this bit and, whenever he does, he receives this bit, otherwise he learns nothing about b.

In a naive but concrete realization of bit commitment, $\mathcal{A}lice$ puts the bit into a strong-box of which she keeps the key and then gives this strong-box to $\mathcal{B}ob$. At a later time, if $\mathcal{B}ob$ requests the bit, $\mathcal{A}lice$ gives the key to $\mathcal{B}ob$. The main point is that $\mathcal{A}lice$ cannot change her mind about the bit b and $\mathcal{B}ob$ learns nothing about it unless he obtains the key.

3 Quantum Bit Commitment: the attack

For every quantum bit commitment protocol Q, we shall construct a protocol $\widetilde{Q}$, show that the security of Q implies the security of $\widetilde{Q}$ and then show that $\widetilde{Q}$ is insecure.

Let A and B stand for $\mathcal{A}lice$ and $\mathcal{B}ob$ respectively. For any bit commitment protocol Q, the state space H is of the form $H_A \otimes H_B$ where H_A and H_B are state spaces on $\mathcal{A}lice$'s side and $\mathcal{B}ob$'s side respectively. $\mathcal{A}lice$'s and $\mathcal{B}ob$'s generation of classical variables, measurements, unitary transformations, etc in the commit phase of Q can be modeled by two global measurements, one on $\mathcal{A}lice$'s side and the other one on $\mathcal{B}ob$'s side. These two measurements together correspond to an overall measurement on the entire state space $H = H_A \otimes H_B$. This single overall measurement corresponds to the entire commit phase of Q.

Now, we construct $\widetilde{Q}$. For every $P \in \{A, B\}$, the state space on P's side is of the form $\widetilde{H}_P = H_P \otimes H'_P$. The entire state space is $\widetilde{H} = \widetilde{H}_A \otimes \widetilde{H}_B$. The additional parts H'_A and H'_B are used to store the outcome of the overall measurement executed by $\mathcal{A}lice$ and $\mathcal{B}ob$ together, that is, the overall measurement executed by $\mathcal{A}lice$ and $\mathcal{B}ob$ in the commit phase of Q becomes a unitary transformation on $\widetilde{H}$. At the opening phase (or just after the commit phase), $\mathcal{B}ob$ and $\mathcal{A}lice$ obtain the classical variables stored in their respective systems H'_A and H'_B, that is they execute the measurements that they normally execute in Q, and they continue with the opening phase as in Q.

It is not hard to see that the non security of $\widetilde{Q}$ implies the non security of Q. Assume that $P \in \{A, B\}$ can cheat in $\widetilde{Q}$. A dishonest P in Q can do exactly as P in $\widetilde{Q}$. The resulting random situation in Q after the commit phase is the same random situation that holds in $\widetilde{Q}$ after that the other participant $\bar{P}$ has measured his quantum system $H'_{\bar{P}}$. So, if P succeed in $\widetilde{Q}$, P also succeed in Q.

Now, we must show that $\widetilde{Q}$ is insecure. It is a principle that we must assume that every participant knows every detail of the protocol, including the distribution of probability of a random variable generated by another participant. There is no loss of generality in assuming that at the beginning of the protocol, the overall system is in a pure state $|\psi\rangle \in \widetilde{H} = \widetilde{H}_A \otimes \widetilde{H}_B$: the preparation of a mixture could be included as a part of the protocol. The commit phase of the protocol specifies a unitary transformation U_b on the entire system. So at the end of the commit phase, the overall system is in a final state $|\phi_b\rangle = U_b|\psi\rangle$.

It is fair to assume that every thing outside $\widetilde{H}_A$ is under the control of a dishonest $\mathcal{B}ob$. In other words, there are no third system $\widetilde{H}_C$. For $b = 0, 1$, let ρ_b^A and ρ_b^B be the partial traces of $|\phi_b\rangle\langle\phi_b|$ over $\widetilde{H}_B$ and $\widetilde{H}_A$ respectively. The density matrices ρ_0^B and ρ_1^B on $\mathcal{B}ob$'s side must be close one to the other, otherwise $\mathcal{B}ob$ can cheat. We shall do the simpler case $\rho_0^B = \rho_1^B$. The more subtle case where the density matrices are not identical is done in the next section.

Consider the Schmidt decomposition [22, 23] of $|\phi_0\rangle$ and $|\phi_1\rangle$ respectively given by

$$|\phi_0\rangle = \sum_i \sqrt{\lambda_i} |e_i^{(0)}\rangle \otimes |f_i\rangle$$

and

$$|\phi_1\rangle = \sum_i \sqrt{\lambda_i} |e_i^{(1)}\rangle \otimes |f_i\rangle$$

In the above formula, λ_i are eigenvalues of the density matrices ρ^B, ρ_0^A and ρ_1^A. The fact that these density matrices share the same positive eigenvalues with the same multiplicity is part of the Schmidt decomposition theorem [22, 23]. The states $|e_i^{(b)}\rangle$ and $|f_i\rangle$ are respectively eigenstates of ρ_b^A and ρ^B associated with the same eigenvalue λ_i. Clearly, the same unitary transformation that maps $|e_i^{(0)}\rangle$ into $|e_i^{(1)}\rangle$ also maps $|\phi^{(0)}\rangle$ into $|\phi^{(1)}\rangle$. We recall that $\mathcal{A}lice$ knows what are the states $|\phi_0\rangle$ and $|\phi_1\rangle$. Therefore, she can determine the above unitary transformation.

In order to cheat, $\mathcal{A}lice$ creates the state $|\phi_0\rangle$. In other words, $\mathcal{A}lice$ does what she must honestly do when she has $b = 0$ in mind. With the state $|\phi^{(0)}\rangle$ $\mathcal{A}lice$ is able to convince $\mathcal{B}ob$ that she had $b = 0$ in mind: $\mathcal{A}lice$ has only to open the bit as an honest $\mathcal{A}lice$ would in $\widetilde{Q}$ with $b = 0$ in mind. If $\mathcal{A}lice$ want to change her mind, she only has to maps $|\phi^{(0)}\rangle$ into $|\phi^{(1)}\rangle$ before she continue the opening phase as if she had $b = 1$ in mind.

4 The real situation

Now, we consider the real situation where the density matrices ρ_0 and ρ_1 are not identical. If the protocol is to be secure against $\mathcal{B}ob$, the density matrices ρ_0 and ρ_1 must respect some constraint. We express this constraint in terms of measurements on the n photons that return a binary classical outcome $X \in \{0, 1\}$. We recall that $\mathcal{A}lice$ prepares the density matrix ρ_b when she has b in mind, that is, when $B = b$. Without loss of generality, we take the convention that $\mathrm{P}(X = 0|B = 0) \geq \mathrm{P}(X = 0|B = 1)$. We denote X_b the random variable X conditioned by $B = b$ so that $\mathrm{P}(X = x|B = b) = \mathrm{P}(X_b = x)$. The constraint is

$$\left| \frac{1}{2} - PE \right| = \left| \frac{1}{2} - \sum_{b=0}^{1} \mathrm{P}(B = b)\, \mathrm{P}\left(X_b = \bar{b}\right) \right| \leq 2^{-\alpha n}.$$

This constraint says that no matter which measurement $\mathcal{B}ob$ uses to decide between $B = 0$ and $B = 1$, the probability of error is exponentially close to 1/2. It has been shown in [14, 24], building on the work of [22, 25], that this implies the existence of two purifications $|\psi_0\rangle$ and $|\psi_1\rangle$ for ρ_0^B and ρ_1^B respectively such that

$$\langle\psi_0|\psi_1\rangle^2 \geq (1 - 2 \times 2^{-\alpha n}).$$

We have that $|\psi_0\rangle$ and $|\psi_1\rangle$ are almost the same state.

In order to cheat, $\mathcal{A}lice$ prepares the state $|\psi_0\rangle$. If she want to unveil $b = 0$, using the same argument as in the simpler case, she maps $|\psi_0\rangle$ into $|\phi_0\rangle$ and continue as in the honest $\widetilde{Q}$ when she has $b = 0$ in mind. If she wants to unveil $b = 1$, she executes on $|\psi_0\rangle$ the unitary transformation F that would map $|\psi_1\rangle$ into $|\phi_1\rangle$. She obtains the state $F|\psi_0\rangle$. The inner product between the desired state $|\phi_1\rangle = F|\psi_1\rangle$ and the actual state $F|\psi_0\rangle$, is the same as the inner product $\langle\psi_1|\psi_0\rangle$ which is exponentially close to 1. So, for all practical purpose, $\mathcal{A}lice$ can cheat as in the simpler case by applying this transformation F and then continuing as in the honest Q' when she has $b = 1$ in mind. This concludes the proof that every bit quantum bit commitment is insecure.

Note that as a consequence, Yao's proof of security for Quantum Oblivious Transfer [12] fails because it is built on insecure foundations (through no fault of Yao). Ironically, as we stated in the Introduction, the proof of security for Quantum Key Distribution shown in [21] stands despite the fact that it draws on Yao's work because it does not depend on the security of Bit Commitment.

References

[1] Mayers, D., "Unconditionally secure quantum bit commitment is impossible", *Fourth Workshop on Physics and Computation — PhysComp '96*, Boston, November 1996.

[2] Mayers, D., "Unconditionally secure quantum bit commitment is impossible", *Physical Review Letters*, vol 78, pp. 3414 – 3417 (1997). Note that this paper has the same title as [1] even though it uses a different approach.

[3] Brassard, G. and C. Crépeau, "Cryptology column — 25 years of quantum cryptography", SIGACT News, vol. 27, no. 3, pp. 13 – 24 (1996).

[4] Crépeau, C., "What is going on with quantum bit commitment?", *Proceedings of Pragocrypt '96: 1st International Conference on the Theory and Applications of Cryptology*, Prague, October 1996.

[5] Lo, H.–K. and H. F. Chau, "Why quantum bit commitment and ideal quantum coin tossing are impossible." Los Alamos preprint archive `quant-ph/9711065`, November 1997.

[6] Kent, A., "Quantum Bit Commitment from a Computation Bound", Los Alamos preprint archive `quant-ph/9711069`, November 1997.

[7] Kent, A., "Permanently secure quantum bit commitment protocol from a temporary computation bound", Los Alamos preprint archive `quant-ph/9712002`, December 1997.

[8] Bennett, C. H. and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing", *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, Bangalore, India, December 1984, pp. 175 – 179.

[9] Bennett, C. H., F. Bessette, G. Brassard, L. Salvail and J. Smolin, "Experimental quantum cryptography", *Journal of Cryptology*, Vol. 5, no. 1, 1992, pp. 3 – 28.

[10] Bennett, C. H., G. Brassard, C. Crépeau and M.–H. Skubiszewska, "Practical quantum oblivious transfer", *Advances in Cryptology — Proceedings of Crypto '91*, August 1991, Springer – Verlag, pp. 351 – 366.

[11] Brassard, G., C. Crépeau, R. Jozsa and D. Langlois, "A quantum bit commitment scheme provably unbreakable by both parties", *Proceedings of 34th Annual IEEE Symposium on the Foundations of Computer Science*, November 1993, pp. 362 – 371.

[12] Yao, A. C.–C., "Security of quantum protocols against coherent measurements", *Proceedings of 26th Annual ACM Symposium on the Theory of Computing*, 1995, pp. 67 – 75.

[13] Crépeau, C. and L. Salvail, "Quantum oblivious mutual identification", *Advances in Cryptology — Proceedings of Eurocrypt '95*, May 1995, Springer – Verlag, pp. 133 – 146.

[14] Mayers, D., "The trouble with quantum bit commitment", LANL Report No. quant-ph/9603015 (to be published). The author first discussed the result in Montréal at a workshop on quantum information theory held in October 1995.

[15] Lo, H.–K. and H. F. Chau, "Is quantum bit commitment really possible?", *Physical Review Letters*, vol 78, pp. 3410 – 3413 (1997).

[16] Bennett, C. H., C. Fuchs, T. Mor (personal communication).

[17] Brassard, G., D. Chaum and C. Crépeau, "Minimum Disclosure Proofs of Knowledge", Journal of Computer and System Sciences, vol. 37(2), 1988, pp. 247 – 268.

[18] Naor, Moni, R. Ostrovsky, R. Venkatesan and M. Yung, "Perfect Zero-Knowledge Arguments for NP Can Be Based on General Complexity Assumptions", *Advances in Cryptology — Proceedings of Crypto '92*, August 1992, Springer – Verlag, pp. 196 – 214.

[19] Crépeau, C., "Quantum oblivious transfer", *Journal of Modern Optics*, Vol. 41, no. 12, December 1994, pp. 2445 – 2454.

[20] Crépeau, C., J. van de Graaf, and A. Tapp, "Committed Oblivious Transfer and Private Multi-Party Computation", in *Advances in Cryptology: Proceedings of Crypto '95* (Springer – Verlag, Berlin, 1995), Vol. 963, pp. 110 – 123.

[21] Mayers ,D., On the security of the quantum oblivious transfer and key distribution protocols, *Advances in Cryptology — Proceedings of Crypto '95*, August 1995, Springer – Verlag, pp. 124 – 135.

[22] Hughston, L. P., R. Jozsa, and W. K. Wootters, "A complete classification of quantum ensembles having a given density matrix", *Physics Letters A*, vol. 183, pp. 14 – 18, 1993.

[23] Schmidt, E., Math. Ann. 63 (1906) 433.

[24] Mayers, D., "La sécurité des protocoles de la cryptographie quantique", PhD Dissertation, Université de Montréal, 1996. (Supervised by Gilles Brassard.)

[25] Jozsa, R., "Fidelity for mixed quantum states", *Journal of Modern Optics*, vol. 41(12), pp. 2315 – 2323, 1994.