Practical Quantum Oblivious Transfer

Charles H. Bennett
IBM Research [*]

Gilles Brassard [†]
Université de Montréal [‡]

Claude Crépeau [†]
École Normale Supérieure [§]

Marie-Hélène Skubiszewska
Université Paris–Sud [¶]

April 2, 1992

Abstract

We describe a protocol for *quantum oblivious transfer*, utilizing faint pulses of polarized light, by which one of two mutually distrustful parties ("Alice") transmits two one-bit messages in such a way that the other party ("Bob") can choose which message he gets but cannot obtain information about both messages (he will learn his chosen bit's value with exponentially small error probability and may gain at most exponentially little information about the value of the other bit), and Alice will be entirely ignorant of which bit he received. Neither party can cheat (ie deviate from the protocol while appearing to follow it) in such a way as to obtain more information than what is given by the description of the protocol. Our protocol is easy to modify in order to implement the All-or-Nothing Disclosure of one out of two string messages, and it can be used to implement bit commitment and oblivious circuit evaluation without complexity-theoretic assumptions, in a way that remains secure even against cheaters that have unlimited computing power. Moreover, this protocol is practical in that it can be realized with available opto-electronic apparatus while being immune to any technologically feasible attack for the foreseeable future.

[*] IBM T. J. Watson Research Laboratory, Yorktown Heights, New York, NY 10598, USA. e-mail: bennetc@watson.ibm.com.

[†] Supported in part by Canada's NSERC.

[‡] Département IRO, Université de Montréal, C.P. 6128, succursale "A", Montréal (Québec), Canada H3C 3J7. e-mail: brassard@iro.umontreal.ca.

[§] Département de Mathématiques et d'Informatique, École Normale Supérieure, 45 rue d'Ulm, 75230 Paris CEDEX 05, France. e-mail: crepeau@dmi.ens.fr.

[¶] Laboratoire de Recherche en Informatique, Université Paris–Sud, Bâtiment 490, 91405 Orsay, France. e-mail: skubi@dec.prl.com.

1 Introduction and history

Quantum cryptography was initiated by Stephen Wiesner more than two decades ago [30]. Over the years, a large number of theoretical applications of quantum physics to cryptography have been discovered: unforgeable bank notes and multiplexing channel [30], unforgeable subway tokens [5], self-winding one-time pad [4], key distribution [3], oblivious transfer [13], coin flipping [3, 8], and bit commitment [8]. Recently, much excitement was created [29, 18, 14, 26, 15, 28, etc.] when the success of a first experimental prototype was reported for the quantum key distribution protocol [1]. Until now, not only was this prototype the first physical realization of a quantum cryptographic protocol, but key distribution was the only quantum protocol ever proposed that could in fact be implemented reasonably with available technology. Even then, the prototype is not entirely convincing because it achieves secure key exchange over the distance of 32 centimeters!

In this paper, we extend the applicability of quantum cryptography by describing a new protocol for oblivious transfer that is practical in the sense that it can be realized with available opto-electronic apparatus while being immune to any technologically feasible attack for the foreseeable future, regardless of the computing power available to would-be cheaters. Techniques similar to those explained here can also be used to overcome the lack of tolerance to errors apparently inherent to the quantum bit commitment protocol of [8]. In this paper, we only concentrate on the new quantum oblivious transfer protocol, and leave it for the reader to figure out how these techniques apply to the bit commitment protocol. A major advantage of these protocols over the already-feasible key distribution is that bit commitment and oblivious transfer make perfect sense over a short distance.

Previous quantum protocols have been proposed for both these tasks, but either they leaked too much information [30], or they could not have been implemented in practice because they required one party to generate pure single-photon light pulses [3] or because they could not tolerate errors due to detector noise [13, 8].

Before we proceed, let us recall the purpose of Oblivious Transfer (OT). In Rabin's original OT [27], Alice sends a one-bit message to Bob, which he receives with probability 50%, while receiving nothing otherwise. Bob finds out whether or not he received Alice's bit, but Alice remains totally ignorant about this. Neither Alice nor Bob can influence the probability 50% of success. The related notion of 1-out-of-2 Oblivious Transfer ($\binom{2}{1}$–OT) was subsequently [1] invented by Even, Goldreich and Lempel [16]. In this scenario, Alice and Bob, play the following game. Alice starts with two one-bit messages of her choosing. The purpose of the protocol is for Alice to transmit the messages to Bob in such a way that he can choose to receive either one of them (learning its value with exponentially small error probability) but cannot obtain significant partial information on both [2], while Alice remains entirely ignorant of which of the two messages he received. It is shown in [11] that $\binom{2}{1}$–OT and Rabin's OT are equivalent in the sense that either one can be

[1] In fact, what Wiesner called "multiplexing channel" as early as the late 1960's [30] is essentially what we now call 1-out-of-2 oblivious transfer (of messages rather than single bits), but his protocol leaked partial information on both messages and could be subverted by a receiver who lied about the quantum efficiency of his detectors. Thus, it can be said that the original inventor of oblivious transfer is Wiesner and that the current paper, which fixes the shortcomings of Wiesner's protocol, is making quantum cryptography go full circle.

[2] More precisely, if b_0, b_1 are Alice's bits and Δ is the data Bob received through the protocol, then at least one of $\mathbf{H}(b_0|\Delta, b_1)$ or $\mathbf{H}(b_1|\Delta, b_0)$ should be exponentially close to 1.

implemented from a primitive that implements the other. Therefore, at least from a theoretical point of view, it does not matter which of these two protocols we achieve. The Quantum OT protocol described in this paper implements directly $\binom{2}{1}$–OT, which is preferable from a practical point of view.

Although OT might seem to be a bizarre idea at first, it is now well-known [22, 12] that it is a very useful primitive for building up interesting protocols, such as two-party oblivious circuit evaluation (by which Alice owns a secret x, Bob owns a secret y, and both of them compute the value of $f(x, y)$ for an agreed upon function f, in such a way that Alice learns nothing about y and Bob learns nothing about x, except for what can be inferred from one's private input and the public value of $f(x, y)$).

Our new quantum OT protocol is described in Section 2, after a brief review of the main features of quantum physics. Section 3 reviews two fundamental mathematical tools that are useful in order to implement quantum OT in practice and prove its security. Section 4 describes the only possible cheating strategies under the technologically reasonable assumptions that light pulses cannot be stored for a significant length of time. Moreover, Section 4 proves, under this assumption, that our quantum OT protocol cannot be cheated by either party. Finally, Section 5 addresses more sophisticated attacks, which are completely infeasible at present or with any foreseeable technology: pulse storing and coherent measurements. It is shown how to overcome the first of these attacks, but nothing is known about the unconditional security of our protocol against the second attack (which is even more unreasonable than the first, technologically speaking). Nevertheless, even this second attack (or in fact *any* attack consistent with quantum physics) can be thwarted from a computational point of view under the assumption that one-way functions exist.

Let us emphasize that all known classical (ie *non*-quantum) protocols for OT allow at least one among Alice or Bob to cheat without risk of detection if she or he can break an unproved cryptographic assumption of some sort. Moreover, classical OT protocols necessarily offer the opportunity for one party to attempt cheating *off-line*, which means that these protocols fail even if the cryptographic assumption can only be broken at the cost of spending days of computing time on a supercomputer. More importantly, they can fail *retroactively* if the appropriate algorithmic breakthrough is discovered years after the protocol has taken place, as long as the cheating party has kept a transcript of the execution of the protocol. In contrast, the basic quantum OT protocol fails against pulse storing *only* if the attack is carried out *on-line*, while the protocol is taking place. In particular, better technology in the future would not compromise the security of OTs carried out today. Similarly, the computational version of our scheme (assuming the existence of one-way permutations), which is secure against arbitrary technology but not arbitrary computing power, must be cheated on-line if it is to be cheated at all.

2 Method

This section describes a quantum oblivious transfer protocol implementable under realistic physical conditions and assumptions similar to those used in the quantum key distribution protocol of [2]. In particular we assume that the quantum transmission consists of series of very dim pulses of coherent or incoherent polarized light rather than individual photons (which are harder to generate), that the receiver attempts to detect the pulses by noisy, imperfectly quantum-efficient detectors such as photomultiplier tubes, and, as stated in the introduction, that the pulses cannot be stored for a significant length of time, so the receiver must measure each pulse before the next one arrives or else lose the opportunity of measuring it at all.

The quantum transmission used in the protocol uses light pulses of four canonical polarizations: horizontal, vertical, $45°$– diagonal, and $135°$– diagonal, henceforth denoted H, V, P, and Q respectively. As is well known, rectilinear (H and V) photons can be reliably distinguished by one type of measurement, while diagonal (P and Q) photons can be reliably distinguished by another type of measurement; but the uncertainty principle of quantum physics decrees that a random outcome results, and all information is lost, if one attempts to measure the rectilinear polarization of a diagonal photon, or vice versa. More generally, if a ϕ–polarized photon is subjected to a polarization measurement along axis θ, it behaves like a θ–polarized photon with probability $\cos^2(\phi-\theta)$ and like a $(\theta+90°)$–polarized photon with the complementary probability $\sin^2(\phi-\theta)$. Such a measurement can be performed by using a θ–oriented piece of birefringent material such as calcite to split the incoming light beam into two beams (polarized at θ and $\theta+90°$), then directing these beams into two sensitive photon detectors such as photomultiplier tubes. A pair of polarization states, such as H and V, or P and Q, that can be reliably distinguished by some measurement is called a *basis*; we will use polarization states H and V to represent the bits 0 and 1 respectively in the rectilinear basis, and P and Q to represent the same bits in the diagonal basis.

At first it would seem that Rabin's OT could be achieved quite simply by having Alice send Bob a single photon encoding the bit to be obliviously transferred in one of the two canonical bases (rectilinear or diagonal), chosen randomly by Alice. Bob would then randomly choose a basis in which to measure the photon, and finally Alice would tell him the correct basis. At that point Bob would have a half chance of having received Alice's bit in the correct basis, and a half chance of knowing he had spoiled it, but Alice would not know which occurred. This simple protocol is inadequate because its probability of success would be seriously affected by inefficiency or noise in Bob's detectors, and because it would allow Bob to get too much partial information about Alice's bit all the time by measuring in a basis intermediate between rectilinear and diagonal, say $\theta = 22\frac{1}{2}^{\circ}$.

A protocol for achieving $\binom{2}{1}$–OT based on the above idea was proposed by Crépeau and Kilian [13], but it was impractical because it failed dramatically in a realistic setting in which transmission errors may occur and dim light pulses are used rather than single photons. The more complicated protocol below is free from these disadvantages. The first step is necessary to adjust the protocol to the physical limitations of Bob's detection apparatus, but it may be skimmed at first reading, being somewhat peripheral to the main idea of the protocol. (The dark count rate d is a detector's probability of registering a count during a time slot when no photons are incident on it, and the quantum efficiency q is the excess probability, above d, of registering a count when one photon is incident on the detector; a typical photomultiplier tube might have $d = 10^{-5}$ and $q = 25\%$.) Let b_0 and b_1 be Alice's bits and let c be Bob's choice (ie Bob wishes to obtain b_c).

1. Bob tells Alice the quantum efficiency q and dark count rate d of his detectors. If these values are satisfactory (see below), Alice next tells Bob the intensity μ of light pulses she will be using, the fraction a of these pulses she will expect him to detect successfully, and the bit error rate ε she will be willing to correct in his data to compensate for his dark counts and other noise sources. She also decides on a security parameter N used below, which she communicates to Bob. Alice and Bob agree on a linear binary error-correcting code capable of correcting with very high probability N–bit words transmitted with expected error rate ε (see Section 3.1).

   More precisely, a would normally be set to $1 - \mathrm{e}^{-(\mu q+2d)} \approx \mu q$, the Poisson probability of detecting 1 or more photons (or dark counts) in a pulse of intensity μ, but might be set lower to allow for attenuation in the optical path between Alice and Bob. Similarly ε would normally be set to $d/a \approx d/\mu q$, the expected error rate from dark counts in Bob's two detectors, but might be set higher to compensate for other noise sources. Alice's choice of μ is guided by the need to simultaneously set $a \approx \mu q$ high enough and $\varepsilon \approx d/\mu q$ low enough that a cheating Bob, whose detectors were in fact far less noisy and more efficient than he claimed, would not gain a significant advantage from the brighter pulses and more voluminous check information he had thus induced Alice to send. In Section 4.3 it is shown that safe oblivious transfer can be achieved when $\mathbf{H}(2\varepsilon) < \frac{1}{2} - (1 - \mathrm{e}^{-\mu} - \mu\mathrm{e}^{-\mu})/2a$, where $\mathbf{H}$ is the *entropy* function[3]. If this condition cannot be met, Alice aborts the protocol.

   Finally, Alice and Bob engage in a test run in which Alice sends pulses of intensity μ in a prearranged sequence of polarizations, and Bob, reading each pulse in the correct basis, verifies that he can indeed detect the pulses with probability greater than a and error rate less than ε.

2. Alice sends Bob a random sequence of $2N/a$ faint pulses of the four canonical polarizations.

3. Bob randomly decides for each pulse whether to measure it in the rectilinear or diagonal basis, and records the basis and measurement result in a table whenever (with probability approximately a) a pulse is detected. Therefore Bob should successfully receive roughly $2N$ pulses. If he receives a few more, he ignores the excess; if he receives a few less, he completes his table by making a few random guesses, so that it has exactly $2N$ entries. Bob then reports to Alice the arrival times of all $2N$ pulses he committed himself to have received, but not the bases he used or his measurement results.

4. Alice tells Bob the bases she used to send each of the pulses he received.

5. Bob partitions his pulses into two sets of N pulses each: a "good" set consisting (as much as possible) of pulses he received in the correct basis, and a "bad" set consisting (as much as possible) of pulses he received in the wrong basis. He tells Alice the addresses of the two sets, but he does not tell her which is the good set and which is the bad set. At this point, Bob shares with Alice a word (ie an N–bit string) corresponding to his good set of measurements (with an expected error rate

[3] The entropy function is defined as $\mathbf{H}(p) = p \lg \frac{1}{p} + (1-p) \lg \frac{1}{1-p}$.

not greater than ε); he shares nothing (or nearly nothing since he may have received slightly more than N bits in the correct basis) with her with respect to his bad set of measurements provided that he faithfully followed the protocol. Alice does not know which word she shares with Bob. (It may well be that Bob did not quite receive N good pulses because of statistical fluctuations in the number of pulses received — which could be less than $2N$ — and in the proportion of pulses that he measured in the correct basis — which could be slightly under $\frac{1}{2}$. However, when N is large enough, the errors that this might create in his good set are negligible compared to the expected errors due to noise.)

6. Using the error-correcting code chosen at step 1, Alice computes the syndromes of the words corresponding to each set, and she sends them to Bob over an error-free channel. Given this data, Bob should be able to recover the original word corresponding to his good set but not that corresponding to his bad set.

   Furthermore, Alice computes a random subset parity for each set, and tells Bob the addresses defining these random subsets, but not the resulting parities. At this point, Bob knows one of these parities exactly, while knowing nothing (or nearly nothing) about the other parity, and he knows which parity he knows. Of course, Alice knows both parities, but she does not know which one Bob knows. Let x_0 and x_1 denote these parity bits, and let $\hat{c}$ denote which one Bob knows.

7. Bob tells Alice whether or not $c = \hat{c}$. (This is the very first time in the protocol that c enters into play.)

8. If $c = \hat{c}$, Alice gives $x_0 \oplus b_0$ and $x_1 \oplus b_1$ to Bob (in this prescribed order), otherwise she gives him $x_0 \oplus b_1$ and $x_1 \oplus b_0$. From this, Bob extracts b_c.

Theorem: Let Δ be the data Bob obtains from the protocol. At least one of $\mathbf{H}(b_0|\Delta, b_1)$ or $\mathbf{H}(b_1|\Delta, b_0)$ is exponentially close (in N) to 1. Regardless of what happens, Alice learns nothing.

Proof: The rest of this paper constitutes the proof of this theorem. The main idea is that Alice uses an error-correcting code to give Bob enough side information to correct the errors in the good set but not the bad set, then hashes each set down to a single bit in such a way that Bob's residual information on the bit corresponding to his bad set is negligibly small.

Note: Because privacy amplification [7] can be used to distill more than one bit, it is easy to modify the protocol so that b_0 and b_1 are k–bit messages rather than single bits, in effect implementing directly the two-message version of ANDOS, the all-or-nothing-disclosure-of-secrets of [9].

3 Review of useful tools

Two fundamental tools will be needed in order to allow the honest Bob to receive bit b_c while preventing a cheating Bob from learning something about both bits: concatenated codes and privacy amplification.

3.1 Concatenated codes

One major problem in making our protocol work in practice is that we need to furnish Bob with information by which he can correct the small error rate of the good pulses (due to dark counts and other unavoidable noise), and do so with reasonable decoding effort and exponentially small (in N) residual error probability; while at the same time preventing him (except with exponentially small probability of success) from correcting all the errors in a set containing a significant proportion of pulses received in wrong bases, even with unlimited decoding effort. A concatenated code [17] combining a Reed-Solomon (RS) code [23] and a random linear binary code of exponentially smaller size is an appropriate choice for this purpose.

Such codes offer exponentially small residual error probability, while allowing information to be transmitted through a noisy binary symmetric channel efficiently at a rate $R(\varepsilon) \leq 1 - \mathbf{H}(2\varepsilon)$ Their decoding may be accomplished efficiently by Berlekamp's algorithm for the RS-code and by a brute-force search for the random linear code (the brute-force search takes exponential time in the size of the random linear code, but this is efficient since this code is chosen to be exponentially smaller than the RS-code).

Recall that to each binary linear error-correcting code is associated a *parity check matrix* H so that a word b is a codeword if and only if $Hb^\top$ is the zero vector. For an arbitrary word b, the value of $Hb^\top$ is called the *syndrome* of b. Our use of error-correcting codes is somewhat nonstandard. Instead of sending a codeword into the (noisy) quantum channel, Alice sends a random word. To allow efficient decoding by Bob, she also sends him the corresponding syndrome over a noiseless channel. It is easy to see that this does not alter Bob's decoding effort, and it has the advantage of facilitating the use of privacy amplification (see below).

The fact that these codes can also prevent Bob (except with exponentially small probability of success) from correcting the errors in a set containing a significant proportion of pulses received in wrong bases, even with unlimited decoding effort, is far more complicated to demonstrate. We sketch in Section 4.2 that whatever set of (canonical or noncanonical) bases Bob uses at step 3 to get his data and whatever partition he chooses at step 5, the additional information provided to him by Alice does not enable him to correct the errors in both sets, or even gain partial information about more than one of Alice's bits (except with exponentially small probability).

3.2 Privacy amplification

Privacy amplification is a tool developed in [7] for distilling a short very secret string from a longer partly-secret one. Here, we need only a rather simple special case of this technique. Let x denote a string of length N about which Bob knows only k parity bits [4], where $k < N$. A special case of Theorem 10 in [7, p. 224] says that if a random subset of the bits of x is chosen, the probability that Bob has any information about its parity is less than $2^{-(N-k-1)}/\ln 2$.

[4] A parity bit about x is the exclusive-or of an arbitrary subset of the bits of x. In particular, physical bits and check bits generated by linear error-correction codes, such as the syndrome of x, are parity bits.

In particular, consider the case in which Bob already knows t bits of x and consider a security parameter s. If no more than $N-t-s-1$ additional parity bits are given to Bob as the syndrome of x with respect to a linear code, the expected amount of information he has on the parity of a random subset of x is less than $2^{-s}/\ln 2$. (This is a conservative estimate since it is likely that the check bits would not be entirely independent from the bits previously known.) Therefore, if Bob knows a proportion $\gamma < 1 - \mathbf{H}(2\varepsilon)$ of the bits of x before receiving the syndrome, and if enough check bits are provided to correct a proportion ε of errors, then the probability that Bob knows anything about the parity of a random subset of the bits of x remains arbitrarily small provided that x is sufficiently long.

4 Various cheats and how to overcome them

Let us first notice that there is very little that Alice can do in order to cheat the quantum OT protocol of Section 2. Obviously, she can cheat at step 8 by telling Bob the complement of what she should. However, this does not count as genuine cheating since in this case an OT will have been carried out, except that the bit transferred will not have been what it should have (nothing can prevent this type of cheating unless Alice has to commit to her bits before the start of the protocol — an entirely different problem known as *verifiable* oblivious transfer [12]).

What *would* count as a genuine success in cheating for Alice would be if she could determine (or at least get an indication about) which of her bits was of interest to Bob (ie the value of c). But notice that Bob does not say anything that involves c until step 7. Moreover, $\hat{c}$ is purely random and information-theoretically hidden from Alice because she cannot tell which of Bob's sets was the good set [5]. Therefore, telling Alice at step 7 whether or not $c = \hat{c}$ does not reveal any information about c either. Thus, it is information-theoretically impossible for Alice to cheat, regardless of her computing power and available technology, provided that Bob faithfully follows his protocol.

Nevertheless, there *is* one thing that Alice can attempt in the hope that Bob will goof: she can use garbage for one of the two syndromes she sends at step 6. The point is that Bob would have no way of detecting such behaviour if she sends garbage in relation to his bad set (which he does not even try to correct). Therefore, *if Bob complains*, she learns that this must be because she chose to send garbage for the good set. In itself, this cheat would not pay off because Bob would catch Alice in the act of cheating before she had any chance to learn something: Bob's choice c is not used in the protocol until step 7, after Alice is asked to send her syndromes. However, if Bob does *not* complain, Alice might infer that she picked the bad set — since otherwise, she may think, he *would* have complained! This is more serious because in this case the protocol will continue and Alice will learn Bob's choice c, and moreover Bob will not even be aware of this leakage. Of course, each time Alice gambles on this, she runs a 50% chance of being caught, but it may be worthwhile since it could be that even one undetected success is enough to tell Alice a great deal. There is an easy way out for Bob: if he discovers that Alice has cheated, he stoically shuts his mouth and continues as if nothing had happened. (Once

[5] This is why we had to use *non*interactive reconciliation, such as that provided by error-correcting codes, rather than the *interactive* reconciliation protocols of [6, 2].

aware of Alice's dishonesty, he takes whatever actions are necessary to counter her plans, but he must do so discreetly.) If Alice knows that this will be Bob's behaviour, she knows that she cannot hope to learn anything from cheating, and thus she may not even attempt it. Potential harm caused by this kind of cheating behaviour from Alice can also be prevented mathematically rather than psychologically. The protocol used in [13] to reduce $\binom{2}{1}$–OT to so-called α–$\binom{2}{1}$–OT can be used here to ensure that Alice cannot gain information on Bob's choice except with exponentially small probability, and that she is almost certain to be caught in the act (before gaining any information) if she even tries.

In sharp contrast, several cheating strategies are available for Bob to attempt creating two good sets at step 5, or at least two sets so that he learns something about both of Alice's bits. We shall now demonstrate that, regardless of Bob's strategy, there is at least one set that would result in Bob learning at most an exponentially small bias on the corresponding bit of Alice. Without loss of generality, we shall concentrate on *symmetric* strategies, ie cheating strategies that favour neither of the sets formed by Bob. Indeed, any asymmetric strategy would reduce Bob's advantage about one of Alice's bits, and would therefore be less good if Cheating Bob's goal is to learn something about both bits.

4.1 The standard attack

Let us first consider the easy case in which Bob does *not* cheat at step 3. In such a case, Bob's only symmetric strategy would be to select about $N/2$ good bits (and thus $N/2$ bad bits) in each set. As a result, Bob knows only about half of the bits in each set. As long as the number of check bits sent by Alice at step 6 for each set is less than $N/2$, it follows that Bob knows less than N parity bits about each set. Hence, privacy amplification applies to conclude that Bob's expected information on the parities of both random subsets chosen by Alice at step 8 are vanishingly small. Therefore, this attack is futile whenever $\mathbf{H}(2\varepsilon) < \frac{1}{2}$, ie $\varepsilon < 5.501\%$, as we have seen in Section 3.2.

To be technically exact, one should consider the case in which Bob is more lucky than average at step 3 and gets more than N good bits. The number L of good bits follows a Binomial$(2N, 1/2)$. Therefore, the standard deviation of $L/2$ is $\sqrt{N/8}$. This implies that the probability that $L/2$ exceeds $N/2 + 5\sqrt{N/8}$ is about one in two million. Moreover, one should also take account of the privacy amplification parameter s (cf Section 3.2). Setting $s = 21$ makes Bob's probability of knowing the parity of a random subset less than about one in two million as well. Therefore, if the code's syndromes are of length less than $N/2 - 5\sqrt{N/8} - 22$ bits, the probability that Bob succeeds at cheating is no better than one in a million. In practice, this means that ε should be somewhat smaller than 5.501% for Alice to accept to play the game, but that the threshold probability tends to 5.501% as N tends to infinity. In our analysis of the other, more sophisticated, attacks, we shall be somewhat sloppy and determine ε as if Bob did not get more information than average. A more careful analysis will be provided in the final paper.

4.2 The Breidbart attack

An obvious way in which Bob can cheat is by measuring Alice's pulses in bases other than rectilinear or diagonal. The most extreme such strategy would be for him to measure each pulse in the so-called *Breidbart basis* [5], which is angle $22\frac{1}{2}^\circ$, precisely half-way between the canonical bases. When he does this, Bob obtains each of Alice's bits with probability $\beta = \cos^2 22\frac{1}{2}^\circ = (2+\sqrt{2})/4 \approx 85.3553\%$. Note that knowing a bit with probability β yields only $1 - \mathbf{H}(\beta) \approx 0.399$ bit of information in the sense of Shannon, whereas a legitimate measurement in a canonical basis yields an expected 0.5 bit of Shannon information. Nevertheless, it could happen that Breidbart measurements are more useful in the presence of additional check bit information and/or more resistant to privacy amplification.

It turns out that this is not so: no measurement can do better than the legitimate measurements in canonical bases. We now sketch the proof of this claim. Due to space limitation we cannot give a complete proof, which will appear in the journal version of this paper. Rather, we restrict our attention to the situation in which Bob performs only Breidbart measurements [6]. We analyse the volume of check bits that Alice can give to Bob without compromising the secrecy of her two bits. First we make a few legitimate simplifications of the situation we want to analyze. The following scenario summarizes the situation.

- We assume that the quantum channel is error free (this only makes Bob more powerful).
- Alice sends Bob a bit string $b = b_1, b_2, ..., b_N$.
- Bob receives it as $b' = b'_1, b'_2, ..., b'_N$ through a binary symmetric channel which transmit bits correctly with probability β (from the Breidbart measurements).
- Alice reveals the syndrome $S_b = Hb^\top$ to Bob, where H is the $K \times N$ parity check matrix of the linear code considered (K is the syndrome length).
- Alice picks a random subset I of $\{1, 2, ..., N\}$ and announces it to Bob.
- Bob wants to approximate $z = \bigoplus_{i \in I} b_i$.

Let the actual number of errors in Bob's data be D (out of N bits). To simplify the analysis, assume that the exact value of D is revealed to Bob by God. We are about to prove Bob's inability to cheat even when provided with this additional information, which of course implies the same in the real world (since he could elect not to use the information even if provided). As long as $K \leq 3N/5$, we now show that, except with exponentially small probability, Bob will have an exponential number of equally likely candidates for Alice's original string b. Therefore, privacy amplification applies to conclude that his information on z is vanishingly small. In contrast, any value of K larger than $N/2$ would have allowed Bob to guess z with good probability if his information had been obtained

[6] *A priori*, it could be that the best strategy for Bob is a mixed strategy in which he measures some pulses in the Breidbart basis, some in canonical bases, and perhaps some others in yet other bases.

by use of measurements in canonical bases (in which case Bob would know half the bits of b exactly and would have no information on the other half).

For any positive $\delta < 1-\beta$, except with exponentially small probability (as a function of N), for all large enough N,

$$N/2 < N - D < (\beta+\delta)N\,.$$

Moreover, the number of words at Hamming distance D from b' is $\binom{N}{D}$, which is thus lower-bounded by

$$\binom{N}{D} > \binom{N}{(\beta+\delta)N}\,.$$

Using the approximation [23]

$$\frac{2^{\mathbf{H}(\lambda)N}}{\sqrt{8N\lambda(1-\lambda)}} \le \binom{N}{\lambda N} \le \frac{2^{\mathbf{H}(\lambda)N}}{\sqrt{2\pi N\lambda(1-\lambda)}}\,,$$

we get the lower bound

$$\binom{N}{D} > \frac{2^{\mathbf{H}(\beta+\delta)N}}{\sqrt{N}}$$

because $8\lambda(1-\lambda) \le 1$ precisely when $\lambda \ge \beta$ (or $\lambda \le 1-\beta$).

We want to obtain a lower bound on the number of words at distance D from b' that have b's syndrome since those are the only candidates for b in the eyes of Bob. Unfortunately, we cannot simply divide the above lower bound on $\binom{N}{D}$ by 2^K, which is the number of syndromes, because there is no reason *a priori* to believe that all the syndromes are equally represented among the words at distance D from b'. Let M stand for $\binom{N}{D}/2^K$. Let $N_D(x,y)$ be the number of words with syndrome y at distance D from a fixed word w with syndrome x (this function is well defined because its value is independent of the specific choice of w; moreover, $N_D(x,y) = N_D(\vec{0}, x \oplus y) = N_D(y,x)$). Provided that $K \le 3N/5$, we now show that $N_D(S_{b'}, S_b)$ is exponentially large except with exponentially small probability, where the probability is taken over all choices of b' at distance D from b. More precisely, we now show that $N_D(S_{b'}, S_b) \ge 2^{-r}M$ with probability at least $1-2^{-r}$ for any security parameter $r > 0$.

Starting from word b, each syndrome s occurs $N_D(S_b, s)$ times among the words at distance D from b. Therefore syndrome s has probability $N_D(S_b,s)/\binom{N}{D} = N_D(s,S_b)/\binom{N}{D}$ of being selected, ie of being that of the actual b'. Thus, any syndrome s for which $N_D(s,S_b) < 2^{-r}M$ (which would be bad because it would mean less uncertainty for Bob) has probability of occurrence less than $(2^{-r}M)/\binom{N}{D} = \frac{1}{2^K}2^{-r}$. Even if all but one syndrome were in that category, their collective probability would still be less than 2^{-r}. This establishes the claim that $N_D(S_{b'}, S_b) \ge 2^{-r}M$, except with probability less than 2^{-r}.

Putting this together with our lower bound on $\binom{N}{D}$, we conclude that, except with probability less than 2^{-r} and provided that $N/2 < N-D < (\beta+\delta)N$,

$$N_D(S_{b'}, S_b) \ge 2^{-r}M = \frac{\binom{N}{D}}{2^K}\,2^{-r} > \frac{2^{\mathbf{H}(\beta+\delta)N-K-r}}{\sqrt{N}}\,.$$

Let ρ and ν be two arbitrary positive constants such that $\rho + \nu < \mathbf{H}(\beta + \delta) - \frac{K}{N}$, which is always possible provided that $K < 3N/5$ and δ is small enough. Setting $r = \rho N$ yields the final result that $N_D(S_{b'}, S_b) > 2^{\nu N}/\sqrt{N}$, which is exponentially large in N, except with probability less than $2^{-\rho N} + \text{Prob}[N - D \geq (\beta + \delta)N \text{ or } D \geq N/2]$, which is exponentially small in N.

At this point one can invoke the privacy amplification theorem [7] of [7] and say that the residual information about $z = \bigoplus_{i \in I} b_i$ is exponentially small in N because Bob has exponentially many candidates from which to choose b. This completes the argument from which we conclude that canonical measurements would have been more useful to Bob since they would allow Cheating Bob to recover both of Alice's messages easily if she were willing to supply as many as $K = 3N/5$ check bits in her syndromes (see Section 4.1).

The analysis for all possible canonical and noncanonical measurements is somewhat similar to what we just presented but much more complicated. From now on we assume that Bob makes his measurements in the canonical bases, because he would gain less information otherwise.

— for your eyes only —

An adversary's ideal strategy

In this section we analyse the amount of redundancy that Alice can give to Bob without compromising the secrecy of her two bits. As a measure for how much Bob can approximate b_0 and b_1 we consider his ability to predict $b_0 \oplus b_1$. If he could predict both bits with a non-negligeable advantage he could bias this XOR as well. We use this measure because it makes the analysis simpler. Assume for this analysis that the channel is error free (this makes Bob more powerful). The following scenario summarize the problem we would like to analyse:

- Alice sends Bob a bit string $b = b_1, b_2, ..., b_{2N}$ using bases $\phi_1, \phi_2, ..., \phi_{2N}$.
- Bob receives it as $b' = b'_1, b'_2, ..., b'_{2N}$ with bases $\psi_1, \psi_2, ..., \psi_{2N}$ of his choice.
- Alice creates a random codeword $c = c_1, c_2, ..., c_{2N}$ and reveals $x = b \oplus c$ to Bob, together with the bases $\phi_1, \phi_2, ..., \phi_{2N}$.
- Bob computes a string $c' = b' \oplus x$.
- Alice picks a random subset I of $\{1, 2, ..., 2N\}$ and announces it to Bob.

How well can Bob approximate the random variable $C = \bigoplus_{i \in I} c_i$? We give an upper bound on Bob's ability to predict C valid for all choices of angles $\psi_1, \psi_2, ..., \psi_{2N}$ he could make at the second step as a function of the rate R of the codewords used by Alice.

Let us first define two magic constants that will be used in the analysis later. Let $\beta = \cos^2(22.5°) \approx .853553$ and let $\gamma = H^{-1}(0.5) \approx .889972$. Denote $©_p^b$ the random

[7] To be technically exact, one needs Lemma 9 rather than Theorem 10 from [7] in this case.

variable that will take the binary value b with probability p and $1-b$ with probability $1-p$.

Bob will make his guess for the value of C based on his knowledge of the string c' and of the angles he used. We formalize our argument by replacing Bob's data c' by even more reliable data $c'' = c''_1, c''_2, ..., c''_{2N}$ and show that even with that better data he still cannot predict C.

The distribution of Bob's actual data is given by

$$c'_i = ©^{c_i}_{\cos^2(\phi_i - \psi_i)}$$

. As a first approach to the above replacement idea, consider the following distribution

$$c''_i = \begin{cases} c_i & \text{if } \cos^2(\phi_i - \psi_i) > \beta \\ ©^{c_i}_{\beta} & \text{if } \cos^2(\phi_i - \psi_i) \leq \beta \end{cases}$$

Notice that for both these distributions, whatever ψ_i is $\Pr\left[\cos^2(\phi_i - \psi_i) > \beta\right] = \frac{1}{2}$ because ϕ_i is equally likely to be $0°$ or $45°$, unless $\psi_i = 22.5°$ which gives $\Pr\left[\cos^2(\phi_i - 22.5°) \leq \beta\right] = 1$. Notice also that whenever $|\phi_i - \psi_i| \geq 22.5°$ we have that $\cos^2(\phi_i - \psi_i) \leq \beta$ which means that the random variable c''_i as always at least as corrolated to c_i as c'_i. Now we estimate Bob's ability to predict C when the string c'' is given to him.

On average Bob gets half the bits of c exactly and knows that fact and the other half of the bits are given to him through a symmetric binary channel with error probablity $1-\beta$. If he uses some $\psi_i = 22.5°$ he will decrease his expected number of bits received exactly, which is worse. Given that Bob has received the string c'' we would like to count the number of codewords that this string could come from. For this we consider the number of words that agree with βN out of N specific positions of c'' (remember that the other N positions are known precisely). The number of words that agree with βN out of N positions is

$$\binom{N}{\beta N} \approx \frac{2^{H(\beta)N}}{\sqrt{N}}.$$

If we compare this to the number of syndromes of the code 2^{2N-2RN} we find that as long as $R > 1 - \frac{H(\beta)}{2} \approx 0.699561$ there will be an exponential number (in N) of codewords that are equally possible as origin for c''. By a theorem of [BBR] it implies that Bob's knowledge of C is exponentially small in N. **(of cousre it is simplified here but you understand...)**

Define the function

$$\ddot{x} = \cos^2(45° - \arccos(\sqrt{x}))$$

which return the value $\cos^2(45° - \theta)$ given $\cos^2(\theta)$ as input. Let $\delta \approx 0.922$ be a new magic constant such that

$$H(\delta) = \frac{H(\ddot{\delta})}{2}.$$

Now consider a more precise distribution for c'':

$$c_i'' = \begin{cases} c_i & \text{if } \cos^2(\phi_i - \psi_i) > \delta \\ ©^{c_i}_{\delta} & \text{if } \delta \geq \cos^2(\phi_i - \psi_i) \geq \ddot{\delta} \\ ©^{c_i}_{\ddot{\delta}} & \text{if } \cos^2(\phi_i - \psi_i) < \ddot{\delta} \end{cases}$$

Notice now that if ψ_i is such that $\delta \geq \cos^2(\psi_i) \geq \ddot{\delta}$ then $\delta \geq \cos^2(\phi_i - \psi_i) \geq \ddot{\delta}$ for both values of ϕ_i, and for all other (outer) ψ_i we get

$$\Pr\left[\cos^2(\phi_i - \psi_i) > \delta\right] = \Pr\left[\cos^2(\phi_i - \psi_i) < \ddot{\delta}\right] = \tfrac{1}{2}.$$

This definiton of c_i'' is also more corrolated to c than c' because for each possibilty we replace each probability with an upper bound.

Now suppose that Bob decides a times out of $2N$ to be in the case $\delta \geq \cos^2(\psi_i) \geq \ddot{\delta}$ and b times out of $2N$ to be in the opposite case. On average Bob gets $\frac{b}{2}$ bits of c exactly and knows that fact and the other $\frac{b}{2}$ bits are given to him through a symmetric binary channel with error probablity $1 - \ddot{\delta}$. The other a bits are given to him through a symmetric binary channel with error probablity $1 - \delta$. Given that Bob has received the string c'' we would like to count the number of codewords that this string could come from. For this we consider the number of words that agree with δa out of a specific positions of c'' and that agree with $\ddot{\delta}\frac{b}{2}$ out of $\frac{b}{2}$ extra positions (remember that the other $\frac{b}{2}$ positions are known precisely). The number of words that agree with with δa out of a and that agree with $\ddot{\delta}\frac{b}{2}$ out of $\frac{b}{2}$ positions is

$$\binom{a}{\delta a}\binom{\frac{b}{2}}{\ddot{\delta}\frac{b}{2}} \approx \frac{2^{H(\delta)a}}{\sqrt{a}} \frac{2^{\frac{H(\ddot{\delta})}{2}b}}{\sqrt{\frac{b}{2}}} \approx \frac{2^{H(\delta)2N}}{\sqrt{\frac{ab}{2}}}.$$

due to the choice of our magic cosntant δ.

If we compare this to the number of syndromes of the code 2^{2N-2RN} we find that as long as $R > 1 - H(\delta) \approx 0.607$ there will be an exponential number (in N) of codewords that are equally possible as origin for c''. Again the theorem of [BBR] implies that Bob's knowledge of C is exponentially small in N.

This argument can be generalized. Let's do it one more time with a larger number of cases Bob will have to choose from. Let $\zeta \approx 0.987$, $\eta \approx 0.953$ and $\lambda \approx 0.89$ be three new magic constants such that

$$H(\lambda) = \frac{H(\ddot{\lambda}) + H(\eta)}{2} = \frac{H(\ddot{\eta}) + H(\zeta)}{2} = \frac{H(\ddot{\zeta})}{2}.$$

Now consider a even more precise distribution for c'':

$$c_i'' = \begin{cases} c_i & \text{if } \cos^2(\phi_i - \psi_i) > \zeta \\ ©^{c_i}_{\zeta} & \text{if } \zeta \geq \cos^2(\phi_i - \psi_i) > \eta \\ ©^{c_i}_{\eta} & \text{if } \eta \geq \cos^2(\phi_i - \psi_i) > \lambda \\ ©^{c_i}_{\lambda} & \text{if } \lambda \geq \cos^2(\phi_i - \psi_i) \geq \ddot{\lambda} \\ ©^{c_i}_{\ddot{\lambda}} & \text{if } \ddot{\lambda} > \cos^2(\phi_i - \psi_i) \geq \ddot{\eta} \\ ©^{c_i}_{\ddot{\eta}} & \text{if } \ddot{\eta} > \cos^2(\phi_i - \psi_i) \geq \ddot{\zeta} \\ ©^{c_i}_{\ddot{\zeta}} & \text{if } \cos^2(\phi_i - \psi_i) < \ddot{\zeta} \end{cases}$$

Now suppose that Bob decides a times out of $2N$ to be in the case $\lambda \geq \cos^2(\psi_i) \geq \ddot{\lambda}$, b times out of $2N$ to be in the case $\eta \geq \cos^2(\psi_i) > \lambda$ or $\ddot{\lambda} > \cos^2(\psi_i) \geq \ddot{\eta}$, c times out of $2N$ to be in the case $\zeta \geq \cos^2(\psi_i) > \eta$ or $\ddot{\eta} > \cos^2(\psi_i) \geq \ddot{\zeta}$ and d times out of $2N$ to be in the case $\cos^2(\psi_i) > \zeta$ or $\cos^2(\psi_i) < \ddot{\zeta}$. On average Bob gets $\frac{d}{2}$ bits of c exactly and knows that fact and the other $\frac{d}{2}$ bits are given to him through a symmetric binary channel with error probablity $1-\ddot{\zeta}$, $\frac{c}{2}$ bits through a symmetric binary channel with error probablity $1-\zeta$,$\frac{c}{2}$ bits through a symmetric binary channel with error probablity $1-\ddot{\eta}$,$\frac{b}{2}$ bits through a symmetric binary channel with error probablity $1-\eta$,$\frac{b}{2}$ bits through a symmetric binary channel with error probablity $1-\ddot{\lambda}$, and a bits are given to him through a symmetric binary channel with error probablity $1-\lambda$. Given that Bob has received the string c'' we would like to count the number of codewords that this string could come from. For this we consider the number of words that agree with λa out of a specific positions of c'', that agree with $\ddot{\lambda}\frac{b}{2}$ out of $\frac{b}{2}$ positions,that agree with $\eta\frac{b}{2}$ out of $\frac{b}{2}$ positions,that agree with $\ddot{\eta}\frac{c}{2}$ out of $\frac{c}{2}$ positions,that agree with $\zeta\frac{c}{2}$ out of $\frac{c}{2}$ positions,and that agree with $\ddot{\zeta}\frac{d}{2}$ out of $\frac{d}{2}$ positions (remember that the other $\frac{d}{2}$ positions are known precisely). The number of words with this property is

$$\binom{a}{\lambda a}\binom{\frac{b}{2}}{\ddot{\lambda}\frac{b}{2}}\binom{\frac{b}{2}}{\eta\frac{b}{2}}\binom{\frac{c}{2}}{\ddot{\eta}\frac{c}{2}}\binom{\frac{c}{2}}{\zeta\frac{c}{2}}\binom{\frac{d}{2}}{\ddot{\zeta}\frac{d}{2}} \approx \frac{2^{H(\lambda)a}}{\sqrt{a}}\frac{2^{\frac{H(\ddot{\lambda})+H(\eta)}{2}b}}{\frac{b}{2}}\frac{2^{\frac{H(\ddot{\eta})+H(\zeta)}{2}c}}{\frac{c}{2}}\frac{2^{\frac{H(\ddot{\zeta})}{2}d}}{\sqrt{\frac{d}{2}}} \approx \frac{2^{H(\lambda)2N}}{\frac{bc}{4}\sqrt{\frac{ad}{2}}}.$$

due to the choice of our magic cosntants.

If we compare this to the number of syndromes of the code 2^{2N-2RN} we find that as long as $R > 1 - H(\lambda) \approx 0.517$ there will be an exponential number (in N) of codewords that are equally possible as origin for c''. Again the theorem of [BBR] implies that Bob's knowledge of C is exponentially small in N.

Seemingly the inner interval is converging toward $[\ddot{\gamma}, \gamma]$ which is quite natural since γ was chosen so that $H(\gamma) = \frac{1}{2}$ and that we have a trivial lower bound at $\frac{1}{2}$. This bound is due to the fact that if Bob is honest he will get half the bits and if the redundency is more than half he would be able to recover the whole string. ∎

— Back to the Future —

4.3 Beamsplitting

Another way by which Bob can cheat involves step 3 again. The idea is to capitalize on the fact that the pulses sent by Alice at step 2 are not pure single-photon states. Recall that Alice's pulses are sent with an expected μ photon per pulse, where μ is significantly smaller than 1. More precisely, a perfectly efficient photo-counter would count for each pulse a number of photons that follows a Poisson distribution with mean μ. In particular, there is a probability $\xi = 1 - \mathrm{e}^{-\mu} - \mu\mathrm{e}^{-\mu} \approx \mu^2/2$ that a pulse would give rise to a multiple count. We shall assume conservatively that whenever a multiple count is obtained, Bob learns Alice's bit with certainty [8].

It is now important to remember that Honest Bob's detectors have less than perfect efficiency. Recall that Bob's counting efficiency, denoted by a, was determined at step 1 and that the number of pulses sent by Alice at step 2 is $M = T/a$, where $T = 2N$ is the number of pulses that Bob must receive successfully. But now consider the case of Cheating Bob, whose photodetectors are in fact perfect. Such a Bob can obtain Alice's bit with certainty for the entire set of $M\xi = T\xi/a$ multiple-count pulses, and he would report success on those at step 3. Assuming that $\xi < a$ —otherwise, Alice would have aborted the protocol at step 1— Bob still has to report success on an additional $(1-\xi/a)T$ pulses, which he chooses randomly among the single-count pulses, which he read (according to the honest protocol) in random canonical bases.

As a result, Bob knows $T\xi/a$ bits from beamsplitting and about half of the remaining $(1-\xi/a)T$ bits from "honest" behaviour, for a total of γT bits, where $\gamma = \frac{1}{2} + \xi/2a$. This is $(1+\xi/a)$ times more bits than the $T/2$ that he would have expected to know had he not taken advantage of multiple-count pulses. In this case, the symmetric cheating strategy consists of splitting the $T = 2N$ bits in two sets of size N so that he knows a proportion γ of the bits in each set. From here, the analysis is similar to that of the standard attack (Section 4.1), except that Bob knows a larger fraction of the bits of whichever set is chosen by Alice. Therefore, this cheat will be thwarted provided that $\mathbf{H}(2\varepsilon) < 1 - \gamma = \frac{1}{2} - \xi/2a$.

This completes the formal demonstration of the main theorem (Section 2), under the reasonable assumption that Bob must measure each pulse before the next one arrives or else lose the opportunity of measuring it at all: the protocol is safe, even against cheaters having access to unlimited computing power, because step 1 makes sure that $\mathbf{H}(2\varepsilon) < \frac{1}{2} - \xi/2a$. The closer to this value is ε, the more pulses will have to be received successfully by Bob at step 3 in order to take account of expected statistical deviations, as explained in Section 4.1. (It is according to this consideration that Alice chooses the value of N at step 1.)

As a numerical example, consider the case in which the efficiency of Honest Bob's photodetectors is $q = 25\%$, and assume that Alice sends her pulses at intensity $\mu = 0.05$. In this case, ignoring dark counts and attenuation in the optical channel, Bob's expected counting efficiency would be $a = 1 - \mathrm{e}^{-\mu q} \approx 1.242\%$, $\xi \approx 0.1209\%$, $\gamma \approx 54.87\%$, and therefore expected error rates ε up to about 4.725% on the legitimate use of the quantum channel can be tolerated. If errors are due only to dark counts, this implies that one expected dark count every 2000 time slots can be tolerated, which is entirely reasonable with current technology.

[8] In principle, though not with present technology, he could do this by analyzing the photon number state of the original pulse without spoiling its polarization, then separating all two-photon pulses into two single photons and measuring one in each canonical basis. After hearing the correct basis from Alice at step 4, he would know which measurement was relevant and thus learn Alice's bit with certainty. In practice, he could learn Alice's bit with probability 75% for double-count pulses by a much simpler apparatus in which a half-silvered mirror is used to split the beam into two parts, one measured rectilinearly and one diagonally. If two counts were obtained in such an apparatus, Bob would be able, after hearing the correct basis from Alice, to determine her bit accurately except when (with probability 25%) both counts had occurred in the wrong-basis half of the apparatus, in which case he would know that he failed to learn anything about Alice's bit.

5 More sophisticated attacks

In principle, the quantum OT protocol described in Section 2 could be subjected to more sophisticated attacks, which are possible in principle although infeasible at present or in the foreseeable future. The first of these attacks, *pulse storing*, can be overcome at the cost of making the protocol more complicated, although it would remain possible to implement it with current technology. The second attack, *coherent measurements*, may be impossible to counter, but it is even more unrealistic than the first one.

5.1 Pulse storing

Instead of measuring the pulses at step 3, Bob could merely pretend to do so, while in fact storing all the pulses he pretended to detect in a lossless delay line. Then, after Alice has announced the sending bases at step 4, he could measure them in the correct bases — which he now knows — using a perfectly efficient detector. He would then be able to present Alice with two "good" sets of bits, and thus obtain both b_0 and b_1. In order to mount such an attack, it is clear that Bob needs to be able to keep the pulses' polarization for an arbitrary long time (because Alice might suspect Bob of attempting this attack and thus wait for a while before step 4) and that he must have perfect or near-perfect photo-detection. However, even *this* would not be sufficient. Bob's additional difficulty is that he must tell Alice as early as step 3 which pulses he claims to have successfully measured. But recall that the pulses are so dim that even a perfectly efficient apparatus would detect only about a fraction μ of them. No technology is available or foreseeable for determining whether the pulse would be detected (formally, measuring the number-state of the pulse) without in fact attempting to detect it, which would spoil it!

Even if we grant Bob the technology necessary to perform this attack, there is a conceptually easy fix to the OT protocol. First of all, Alice would send $3N/a$ pulses at step 2, allowing roughly $3N$ of them to be successful if Bob is honest. Then, before step 4, Bob would use a bit commitment scheme to commit to each of the bases used for his successful measurements as well as to the bits thus obtained. Still before step 4, Alice would select a random subset of N reported successes, and ask Bob to open his commitments for those. This allows Alice to check that Bob's commitments are correct (subject to error rate ε) when his committed basis is correct and that his commitments are uncorrelated to the correct bits when his basis is incorrect. Not only does this prove to Alice that Bob measured the pulses before step 4, but also that he did not measure them in noncanonical bases (such as the Breidbart basis).

But of course, one may ask which commitment scheme should be used? Obviously, we would lose most of the benefit from quantum cryptography if we used a scheme that is merely computationally secure. Fortunately, *quantum* bit commitment schemes exist [3, 8]. Even though the schemes presented in [3, 8] are technologically unreasonable, as mentioned in the introduction, the techniques used in the current paper can be used also to modify the scheme of [8] in order to render it feasible with current technology.

5.2 Coherent measurements

So far, we have limited Bob to measuring pulses one at a time, and combining the classical results of these measurements with information subsequently obtained from Alice. The formalism of quantum mechanics allows a more general kind of measurement, which is even more infeasible than pulse storing. Such a measurement would treat the entire sequence of M pulses sent during step 2 as a single 2^M–state quantum system, cause it to interact coherently with an intermediate quantum system of comparable complexity, maintain the phase coherence of the intermediate system for an arbitrarily long time, then finally measure the intermediate system in a way depending on the information provided by Alice at step 4.

In the light of the previous section, avoiding this attack appears easy. Indeed the fix we just showed for pulse storing will also apply to this kind of attack. Unfortunately, the bit commitment scheme of [8] is also susceptible to coherent measurements (although in the case of that scheme the receiver will be Alice, which means that *she* will be the one who can potentially cheat). Alternatively, we could use the bit commitment scheme implicit in [3], but *it* is susceptible to an attack related to the Einstein–Podolsky–Rosen paradox (in addition to requiring the use of single-photon pulses, which are hard to generate in practice). As a consequence, it is not known whether our protocols can be made unconditionally secure against all possible attacks consistent with quantum physics.

Nevertheless, an interesting protocol results if we are satisfied with computational security. Indeed, it is well-known that computationally secure bit commitments are possible under the assumption that one-way functions exist [20, 19, 24]. Therefore, quantum physics provides for an OT protocol that is computationally secure against unrestricted technology (including the ability to perform coherent measurements) under the sole assumption that one-way functions exist. This is interesting because Impagliazzo and Rudich have proved that one-way functions are not sufficient to implement OT in the classical (ie non-quantum) model [21]. Moreover, under the assumption that one-way permutations [25] or one-way group actions [10] exist, it is possible to accomplish a quantum OT protocol that will leak no additional information to either party unless the computational assumption is broken *on-line*, while the protocol is taking place. In contrast, all classical OT protocols are susceptible to off-line cheating: at least one party has complete information (in the sense of Shannon) on the other party's secret.

Acknowledgements

We are greatly indebted to Ivan Damgård for his many valuable comments. We thank also Silvio Micali for pointing out that computational complexity based quantum cryptography is interesting since it allows to build oblivious transfer around one-way functions.

References

[1] Bennett, C. H., F. Bessette, G. Brassard, L. Salvail and J. Smolin, "Experimental quantum cryptography", *Advances in Cryptology — Eurocrypt '90 Proceedings*, April 1990, Springer–Verlag, pp. 253 – 265.

[2] Bennett, C. H., F. Bessette, G. Brassard, L. Salvail and J. Smolin, "Experimental quantum cryptography", *Journal of Cryptology*, Vol. 5, no. 1, 1992, to appear.

[3] 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.

[4] Bennett, C. H., G. Brassard and S. Breidbart, "Quantum cryptography II: How to re-use a one-time pad safely even if $\mathcal{P} = \mathcal{N}P$", unpublished manuscript available from the authors, November 1982.

[5] Bennett, C. H., G. Brassard, S. Breidbart and S. Wiesner, "Quantum cryptography, or unforgeable subway tokens", *Advances in Cryptology: Proceedings of Crypto '82*, August 1982, Plenum Press, pp. 267 – 275.

[6] Bennett, C. H., G. Brassard and J.–M. Robert, "How to reduce your enemy's information", *Advances in Cryptology — Crypto '85 Proceedings*, August 1985, Springer–Verlag, pp. 468 – 476.

[7] Bennett, C. H., G. Brassard and J.–M. Robert, "Privacy amplification by public discussion", *SIAM Journal on Computing*, Vol. 17, no. 2, April 1988, pp. 210 – 229.

[8] Brassard, G. and C. Crépeau, "Quantum bit commitment and coin tossing protocols", *Advances in Cryptology — Crypto '90 Proceedings*, August 1990, Springer-Verlag, to appear.

[9] Brassard, G., C. Crépeau and J.–M. Robert, "Information theoretic reductions among disclosure problems", *Proceedings of 27th IEEE Symposium on the Foundations of Computer Science*, October 1986, pp. 168 – 173.

[10] Brassard, G. and M. Yung, "One-way group actions", *Advances in Cryptology — Crypto '90 Proceedings*, August 1990, Springer-Verlag, to appear.

[11] Crépeau, C., "Equivalence between two flavours of oblivious transfers (abstract)", *Advances in Cryptology: Proceedings of Crypto '87*, August 1987, Springer-Verlag, pp. 350 – 354.

[12] Crépeau, C., "Verifiable disclosure of secrets and application", *Advances in Cryptology: Proceedings of Eurocrypt '89*, April 1989, Springer-Verlag, pp. 181 – 191.

[13] Crépeau, C. and J. Kilian, "Achieving oblivious transfer using weakened security assumptions", *Proceedings of 29th IEEE Symposium on the Foundations of Computer Science*, October 1988, pp. 42 – 52.

[14] Deutsch, D., "Quantum communication thwarts eavesdroppers", *New Scientist*, 9 December 1989, pp. 25 – 26.

[15] Ekert, A., "La mécanique quantique au secours des agents secrets", *La recherche*, No. 233, June 1991, pp. 790 –791.

[16] Even, S., O. Goldreich and A. Lempel, "A randomized protocol for signing contracts", *Advances in Cryptology: Proceedings of Crypto '82*, August 1982, Plenum Press, pp. 205 – 210.

[17] Forney, G. D., *Concatenated Codes*, The M.I.T. Press, 1966.

[18] Gottlieb, A., "Conjugal secrets — The untappable quantum telephone", *The Economist*, Vol. 311, no. 7599, 22 April 1989, p. 81.

[19] Håstad, J., "Pseudo-random generation under uniform assumptions", *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing*, May 1990, pp. 395 – 440.

[20] Impagliazzo, R., L. A. Levin and M. Luby, "Pseudo-random generation from one-way functions", *Proceedings of the 21st Annual ACM Symposium on Theory of Computing*, May 1989, pp. 12 – 24.

[21] Impagliazzo, R. and S. Rudich, "Limits on the provable consequences of one-way permutations", *Proceedings of the 21st Annual ACM Symposium on Theory of Computing*, May 1989, pp. 44 – 61.

[22] Kilian, J., "Founding cryptography on oblivious transfer", *Proceedings of the 20th Annual ACM Symposium on Theory of Computing*, May 1988, pp. 20 – 31.

[23] MacWilliams, F. J. and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland, 1977.

[24] Naor, M., "Bit commitment using pseudo-randomness", *Advances in Cryptology — Crypto '89 Proceedings*, August 1989, Springer–Verlag, pp. 128 – 136. To appear in *Journal of Cryptology*, Vol. 4, no. 2, 1991.

[25] Naor, M., R. Ostrovsky, R. Venkatesan and M. Yung, "Perfect zero-knowledge arguments for **NP** can be based on general complexity assumptions", Manuscript available from the authors, 1991.

[26] Peterson, I., "Bits of uncertainty: Quantum security", *Science News*, Vol. 137, 2 June 1990, pp. 342 – 343.

[27] Rabin, M. O., "How to exchange secrets by oblivious transfer", Technical Memo TR–81, Aiken Computation Laboratory, Harvard University, 1981.

[28] Stewart, I., "Schrödinger's catflap", News and Views, *Nature*, Vol. 353, 3 October 1991, pp. 384 – 385.

[29] Wallich, P., "Quantum cryptography", *Scientific American*, Vol. 260, no. 5, May 1989, pp. 28 – 30.

[30] Wiesner, S., "Conjugate coding", manuscript written *circa* 1970, unpublished until it appeared in *Sigact News*, Vol. 15, no. 1, 1983, pp. 78 – 88.

[31] cette reference est tres longue pour forcer un changement de page! cette reference est tres longue pour forcer un changement de page! cette reference est tres longue pour forcer un changement de page! cette reference est tres longue pour forcer un changement de page! cette reference est tres longue pour forcer un changement de page! cette reference est tres longue pour forcer un changement de page! cette reference est tres longue pour forcer un changement de page! cette reference est tres longue pour forcer un changement de page!