Contents

Chapter 6

Distributing keys

In this chapter we finally give and analyze a complete protocol for key distribution. We will approach our objective in a series of steps, ending up with the famous BB84 quantum key distribution (QKD) protocol. Throughout, we will assume that Alice and Bob share an *classical authenticated channel (CAC)*. Whenever Alice and Bob make use of this channel below, we will say they send the information over "the CAC".

6.1 Secure key distribution

The main cryptographic challenge that we will consider is the one of key distribution. Here, our protagonists, Alice and Bob want to protect their communication from the prying eyes of an eavesdropper Eve. Alice and Bob are thereby always honest, and Eve is the adversary. Alice and Bob have control over their secure labs that Eve cannot peek into. However, Eve has access to the communication channel connecting Alice and Bob.

Definition 6.1.1 (Key distribution)**.** *A* key distribution *protocol between Alice and Bob aims to achieve the following goals, given a security parameter $\varepsilon \geq 0$:*

- *$\varepsilon-$ correctness: Alice and Bob both agree on an m-bit key $K \in \{0,1\}^m$, except with some failure probability at most ε. That is, both Alice and Bob have K_A, K_B respectively, and* $\text{Prob}(K_A \neq K_B) \leq \varepsilon$.
- *$\varepsilon-$ security: Any outsider Eve is almost ignorant about the key, i.e. $\rho_{KE}^{\text{real}} \approx_\varepsilon \rho_{KE}^{\text{ideal}}$ where $\rho_{KE}^{\text{ideal}} = \frac{\mathbb{I}_K}{2^m} \otimes \rho_E$.*

To achieve such a key distribution protocol, we will consider the following communication channels which Alice and Bob may have access to:

1. A classical channel: Alice and Bob can send classical bits in either direction over this channel. Eve has complete access to this channel. In particular, she can read all messages, modify them, and even impersonate Alice (or Bob).

2. A classical *authenticated* channel (CAC): A classical communication channel with one extra guarantee: Alice and Bob know that the message originated unaltered from Alice or Bob respectively. This means that while the channel is not secret because Eve can still read all the messages that travel across, she cannot impersonate Alice or Bob or alter messages traveling over the channel.

3. A classical *secret* channel: A classical communication channel in which Eve cannot learn any information (see below!) about the messages traveling across. Yet, while she cannot hope to gain any information about the message, Eve could impersonate Alice or Bob.

4. A classical *secret and authenticated* channel: A classical communication channel combining both guarantees above.

5. A *quantum communication* channel: A channel where Alice may send quantum information (in particular, in the form of qubits) to Bob, where Eve has full access to all the quantum communication.

For simplicity we will start our discussion by assuming that Alice and Bob are already connected by a CAC - we will see later how such a CAC can be built. That is, we will for the moment only worry about establishing a key which is hidden from Eve!

6.2 Information reconciliation

Let's describe a communication scenario and introduce some convenient notation. Alice and Bob hold two strings that we denote X_A and X_B. X_B, the string of Bob, equals X_A plus a string of errors that we denote by S. Alice and Bob are connected by a CAC, and therefore an information reconciliation protocol consists simply in the exchange of messages over this channel in order for Bob to recover X_A. We denote by C the string consisting of all the messages exchanged over the classical channel. Finally, Bob, with the help of X_B and C will output $\hat{X}_A$, that is, an estimate of the string of Alice X_A.

Definition 6.2.1 (Information Reconciliation)**.** *Let X_A, X_B be distributed according to the joint probability distribution $P_{X_A X_B}$. An* information reconciliation *protocol for X_A, X_B is ε-correct and leaks $|C|$ bits if :*

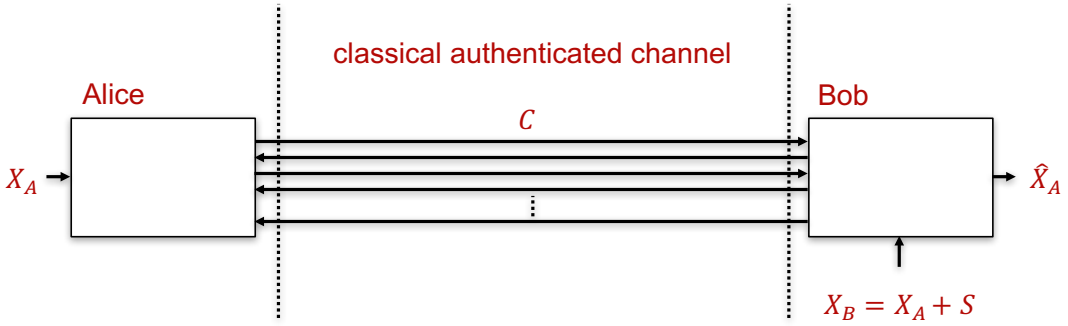

Figure 6.1: Scheme of a generic information reconciliation protocol.

- $\text{Prob}(X_A \neq X_B) \leq \varepsilon$.
- *The length of the messages exchanged on the public channel is* $|C|$.

The goals of the information reconciliation step are two-fold. First and most obvious, Alice and Bob want to ensure that after reconciliation the strings X_A and $\hat{X}_A$ are ε-correct, that is, that the probability that they are different is at most ε. Second, note that all classical communication between Alice and Bob is public, which means that Eve can also gain information from the error correction information they send across! Recalling the chain rule for the min-entropy, we have

$$H_{\min}(X|EC) \geq H_{\min}(X|E) - |C| \,, \tag{6.1}$$

where we used C to denote the error-correcting information sent across the classically authenticated channel. We thus have the min-entropy of Eve *with* the error-correction information C can shrink by at most the number of bits $|C|$ of error-correction information that Alice and Bob send.

Again, it is very easy to achieve any of both goals independently. Why is this? Imagine that a reconciliation protocol consists of Alice sending her whole string to Bob over the classical channel. This is a great protocol if we only care about correctness, the strings will definitely be correct, but the leakage is maximal and after reconciliation we would not have any min-entropy left to do privacy amplification.

On the other hand, imagine a reconciliation protocol that consists in Alice and Bob doing nothing, then, for leakage purposes, the protocol is perfect, the leakage is zero, but the strings might not be correct for the desired ε.

We can classify information reconciliation protocols depending on their usage of the classically authenticated channel. The most general protocol might consist in the exchange of messages in both directions, that is from Alice to Bob and

from Bob to Alice. We call such a protocol a two-way or an interactive protocol. However, much more simpler, and in many circumstances, it is already sufficient that the whole reconciliation consists of a single message from Alice to Bob, that is, the communication happens one-way. We refer to such protocols as one-way reconciliation protocols, where Alice encodes her string X_A into a single (significantly shorter) message that we denote by C_A and sending it through the classical channel. Then Bob uses C_A to eliminate errors from X_B, effectively recovering X_A.

6.3 Syndrome coding

In the following we will explore one concrete one-way reconciliation scheme. The scheme that we will describe is based on linear codes. Let us review the definition and main elements of linear codes:

Definition 6.3.1 (Linear code)**.** *Let $\mathbb{F}_q$ be the finite field of size q, a (n, k) q-ary linear code C is a linear subspace of $\mathbb{F}_q^n$ of dimension k.*

The individual elements (which are q-ary strings of length n) contained in the subspace defining a linear code are called *codewords*, and a (n, k) q-ary code has q^k different codewords. We will only be concerned with binary codes, so in the following we let $q = 2$.

Since a code is just a subspace, if we want to construct an n length binary code, any procedure that characterizes a subspace of $\mathbb{F}_2^n$ will allow us to construct a code. One convenient characterization is by using a $m \times n$-dimensional *parity check matrix* H. The code is defined as the set of vectors v such that $H \cdot v^T = 0$. The dimension of the code induced by H is $k = n - \text{rank}(H)$.

The map $s_H : \mathbb{F}_2^n \to \mathbb{F}_2^m$ given by $v \mapsto H \cdot v^T$ is called the *syndrome*. It turns out that the syndrome is very useful beyond inducing the code. In particular, an example of a reconciliation procedure is when Alice and Bob agree on a particular parity matrix H. Let's consider this in more detail. First, let us discuss the encoding step performed by Alice. The reconciliation scheme is based on linear codes, given a parity check matrix H and Alice's vector X_A, the encoding of X_A is its syndrome. That is, the message that Alice sends to Bob for information reconciliation is the syndrome of X_A.

Example 6.3.1. *Let $v = (001)$ and $H = \begin{pmatrix} 110 \\ 011 \end{pmatrix}$. Then $s_H(v) = H \cdot v^T = (01)^T$. What this means, is that if v were the string of Alice, the message that Alice would send to Bob for reconciliation would be $s_H(v)$ which as we calculated is: $(01)^T$.* ∎

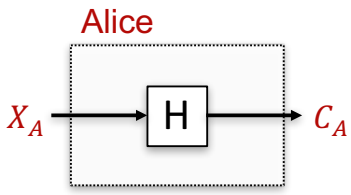

Figure 6.2: The encoder in syndrome coding based reconciliation.

Let us now move to Bob, who has a decoder, which is essentially a procedure that helps him to estimate the error string S (where recall $X_B = X_A + S$), so that he may recover X_A. The decoder is a little bit more complicated than the encoder, since it takes both C_A and X_B as inputs. The first thing that happens in the decoder is that it computes the syndrome of X_B that we call C_B. Then, Bob computes $C_S = C_B + C_A$, where recall that C_A is the syndrome of X_A. We call this resulting string C_S because it is indeed the syndrome of the error string, i.e. $C_S = s_H(S)$. Then, C_S is sent into a module that estimates the error string S and outputs the estimate that we call $\hat{S}$. Finally $\hat{S}$ is added to X_B and this will be the decoded string that Bob will receive.

Exercise 6.3.1. *Show that C_S is indeed the syndrome of the error string S, or in other words, that $s_H(S) = s_H(X_A) + s_H(X_B)$.* ∎

Exercise 6.3.2. *Show that if the error estimate is correct $\hat{X}_A$ will equal X_A.* ∎

A simple but relevant property of the scheme is that as soon as the length of the string that Alice sends Bob is smaller than the length of the string of Alice not all errors can be corrected. In order to see this recall that $\hat{X}_A$ is $X_B + \hat{S}$, but $\hat{S}$ is a function of C_S, the syndrome of S. Hence, even if the estimator function outputs a different value for each syndrome we have 2^m different outputs while there are 2^n different error strings, in other words, unless m equals n, it is not possible to correct all errors. However, if different errors occur with different probabilities we might be satisfied if we correct the most likely errors.

Example 6.3.2. *Let us go back to Example 6.3.1. Let us now describe the decoder, for this we need to make explicit the estimation function. There are four different syndromes. We will assign as error estimate for each syndrome the following strings:*

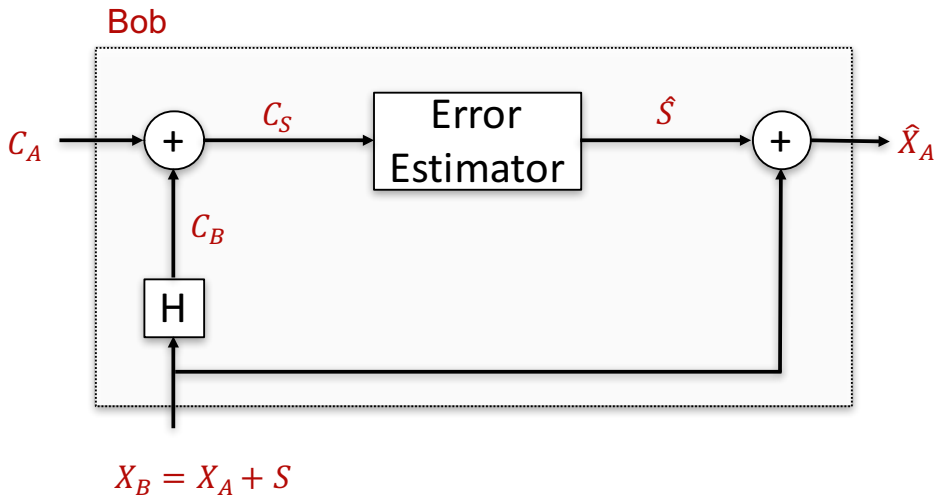

Figure 6.3: The decoder in syndrome coding based reconciliation.

Syndrome	*Error Estimate*
00	*000*
01	*001*
10	*100*
11	*010*

Can you guess why we chose this particular map? The idea is that if there is zero or one errors, the estimator will output the correct error estimate. However, this also means that any other error string will be wrongly estimated. ∎

6.4 Limits of reconciliation

We have seen a concrete scheme for reconciliation. What are the fundamental limits of reconciliation? In order to answer this question we need some structure. Let us assume that the strings of Alice and Bob, that we denote for precision X_A^n and X_B^n, are of length n, where each of the symbols is drawn independently from the same joint distribution $P_{X_A X_B}$ (where X_A, X_B are binary random variables). Then, any information reconciliation protocol that leaks $|C|$ bits satisfies:

$$|C| \geq n \cdot H(X_A|X_B) \tag{6.2}$$

Moreover, the inequality can be achieved when $n \to \infty$ [**slepian1973noiseless**].

In a realistic scenario n is finite and the information reconciliation protocol needs to be computationally efficient. Instead of dealing with the implementation details of an information reconciliation protocol, it is sometimes convenient to approximate the leakage value of a realistic protocol by $\xi \cdot nH(X_A|X_B)$, where $\xi > 1$ is the reconciliation efficiency. The constant ξ is often chosen $\xi \approx 1.2$. However, this approximation should be used with care since ξ will be a function of the length, the noise model and the correctness considered [**tomamichel2014fundamental**].

The errors between Alice's and Bob's string can also generally be modelled by a BSC. That is, we can see their strings as the input and output of a BSC: whenever Alice inputs a bit x, Bob receives x with probability p, but with probability $1-p$ the bit is flipped to $x+1$[1]. In the case of a BSC, Eq. (6.2) simplifies to $|C| \geq nh(p)$ where $h(p) = -p\log(p) - (1-p)\log(1-p)$ is the binary entropy function.

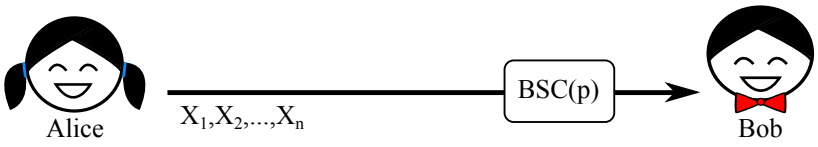

Figure 6.4: The errors between Alice's and Bob's string are generally modelled by a BSC.

6.5 BB84 Quantum key distribution

In last week's lectures we discussed a special classical channel, where Eve is guaranteed to have some amount of noise in her attempts at intercepting bits transmitted by Alice to Bob. In such a classical protocol, if one takes away the guarantee about Eve but instead allow her to arbitrarily intercept messages on the special channel, then it is clear that there is no more security: Eve can learn all the bits of the string x. When considering a quantum channel, the classical protocol would amount to sending a string x encoded in a single fixed basis. For example, we might send x in the standard basis as $|x\rangle\langle x| = |x_1\rangle\langle x_1| \otimes \ldots \otimes |x_n\rangle\langle x_n|$. Eve, knowing the basis, can measure the transmitted quantum state to recover $x_1 \cdots x_n$ without error, copying each bit without causing any disturbance to the state. Of course, we might also encode x in a different basis, for example the Hadamard basis, as $H^{\otimes n}|x\rangle\langle x|H^{\otimes n}$.

[1] In the context of quantum key distribution (QKD), this error probability $1-p$ is also often called the quantum bit-error rate (QBER).

Yet, the fact remains, if Eve knows the basis, then she can copy the bits without being detected!

Exercise 6.5.1. *Consider a bit b encoded in the Hadamard basis $H|b\rangle\langle b|H$. Give a measurement that recovers b (knowing it was encoded in the Hadamard basis!). Compute the post-measurement states for each possible outcome. What do you conclude?* ∎

However, recall that by the no-cloning theorem presented in the Week 0 lecture notes, it is impossible to copy arbitrary qubits, i.e., qubits that could live anywhere on the Bloch sphere. This is precisely the case when Eve does not know the encoding in advance. We are thus motivated to let Alice not just choose bits x_j at random, but for each bit she will also randomly choose a basis θ_j. This gives rise to the BB84 encoding:

Definition 6.5.1 (BB84 states/encoding)**.** *The BB84 states are $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$. This set of states corresponds to encoding a classical bit $x_j \in \{0,1\}$ in a randomly chosen basis $\theta_j \in \{0,1\}$ where $\theta_j = 0$ labels the standard basis, and $\theta_j = 1$ the Hadamard basis.*

Note that the standard basis and the Hadamard basis are the eigenbases of the Pauli-Z and Pauli-X matrices respectively.

We first assume that Alice and Bob are connected via an classical authenticated channel (CAC), which they will use during the protocol. Later in the notes we will investigate how Alice and Bob could construct such a channel.

The BB84 protocol can be described as follows:

Protocol 1 (BB84 QKD (no noise))**.** *Outputs $k \in \{0,1\}^\ell$ to both Alice and Bob. Alice and Bob execute the following:*

1. *For a small real-valued parameter $\eta \ll 1$, and a large integer $n \gg 1$, Alice chooses a string $x = x_1, \ldots, x_N \in \{0,1\}^N$ uniformly at random where $N = (4+\eta)n$. She also chooses a basis string $\theta = \theta_1, \ldots, \theta_N$ uniformly at random. She sends to Bob each bit x_j by encoding it in a quantum state according to the basis θ_j: $H^{\theta_j}|x_j\rangle$.*

2. *Bob chooses a basis string $\tilde{\theta} = \tilde{\theta}_1, \ldots, \tilde{\theta}_N$ uniformly at random. He measures qubit j in the basis $\tilde{\theta}_j$ to obtain outcome $\tilde{x}_j$. This gives him a string $\tilde{x} = \tilde{x}_1, \ldots, \tilde{x}_N$.*

3. *Bob tells Alice over the CAC that he has received and measured all the qubits.*

4. *Alice and Bob tell each other over the CAC their basis strings θ and $\tilde{\theta}$ respectively.*

5. *Alice and Bob discard all rounds j in which they didn't measure in the same basis. Let $S = \{j|\theta_j = \tilde{\theta}_j\}$ denote the indices of the rounds in which they measured in the same basis. Since Alice and Bob chose $\theta, \tilde{\theta}$ at random, for large values of n, they throw away roughly $N/2 \approx 2n$ bits.*

6. *Alice picks a random subset[2] $T \subseteq S$ for testing and tells Bob T over the CAC. That is, Alice and Bob test roughly $|T| \approx N/4 \approx n$ bits.*

7. *Alice and Bob announce x_T and $\tilde{x}_T$ to each other over the CAC, where we denote by x_T the substring of x corresponding to the indices in the test set T. They compute the error rate $\delta = W/|T|$, where $W = |\{j \in T \mid x_j \neq \tilde{x}_j\}|$ is the number of errors when Alice and Bob did measure in the same basis.*

8. *If the error rate $\delta \neq 0$, then Alice and Bob abort the protocol. Otherwise, they proceed to denote $x_{\text{remain}} = x_{S\setminus T}$ and $\tilde{x}_{\text{remain}} = \tilde{x}_{S\setminus T}$ as the remaining bits, i.e., the bits where Alice and Bob measured in the same basis, but which they did not use for testing. The length of x_{remain} and $\tilde{x}_{\text{remain}}$ is approximately n bits.*

9. *Alice and Bob perform privacy amplification: Alice picks a random r, and computes $k = Ext(x_{\text{remain}}, r)$. She sends r to Bob, who computes $k = Ext(\tilde{x}_{\text{remain}}, r)$.*

Note that even though steps 1 and 2 seem to take place one after the other, and you may be tempted to think that Alice and Bob require quantum storage, this is not necessarily the case. Alice can prepare the qubits one-by-one, and Bob can also measure them one-by-one. This is very appealing, since Alice and Bob only need very simple quantum devices - preparing and measuring single qubits is already enough!

Let us first investigate why the protocol is correct. If there are no errors in transmission, then whenever Bob measures in the same basis as the one chosen by Alice ($\theta_j = \tilde{\theta}_j$), then he learns the bit perfectly ($x_j = \tilde{x}_j$). If there is no eavesdropper, they will pass the test. Since $x_{\text{remain}} = \tilde{x}_{\text{remain}}$ and Bob knows r they produce the same output k as before.

But why should this protocol be secure? The intuition is that whenever Eve tries to intercept, and gain some information from the transmitted qubits, she will

[2] A random subset T of S is where each element in S is included in T with probability $1/2$. By this definition, if $|S|$ is large, then $|T| \approx |S|/2$.

invariably disturb the quantum states — and Alice and Bob can detect such disturbance. It can be proven that

$$H_{\min}(X_{\text{remain}}|E) \gtrsim n\left[1 - h(\delta)\right] \ , \tag{6.3}$$

where $h(\delta) = -\delta \log \delta - (1-\delta)\log(1-\delta)$ is the binary entropy function, and recall from Protocol 1 that δ is the error rate between Alice and Bob's sample. We note that analyzing the sampling procedure, i.e., which qubits to test, for small values of N is an intricate problem which requires great care, as analyzed in [**pfister2016sifting**]. Here we will not dive into this but instead consider only the limit of large N.

In the case where errors in transmission occur, the error rate is always $\delta \neq 0$. This affects the correctness of Protocol 1, since Alice and Bob will get $x_{\text{remain}}, \tilde{x}_{\text{remain}}$ respectively, where $x_{\text{remain}} \neq \tilde{x}_{\text{remain}}$. In fact, when $x_{\text{remain}} \neq \tilde{x}_{\text{remain}}$, by property of the extractor Ext, with probability almost equal to 1, $Ext(x_{\text{remain}}, r) \neq Ext(\tilde{x}_{\text{remain}}, r)$. This means that almost for certain, Alice and Bob will end up with different keys!

To overcome this problem, Alice and Bob will have to perform an additional step of *information reconciliation* in the protocol. Thus instead of Protocol 1, they execute the following protocol:

Protocol 2 (BB84 QKD (with noise))**.** *Outputs $k \in \{0,1\}^{\ell}$ to both Alice and Bob. Alice and Bob execute the following:*

1-7. Same as Protocol 1.

8. If the error rate is larger than a certain threshold $\delta > \delta_t$, Alice and Bob abort the protocol. Otherwise, they proceed to denote $x_{\text{remain}} = x_{S\setminus T}$ and $\tilde{x}_{\text{remain}} = \tilde{x}_{S\setminus T}$ as the remaining bits, i.e., the bits where Alice and Bob measured in the same basis, but which they did not use for testing.

9. Alice and Bob perform information reconciliation: Alice sends some error correcting information C across the classical authenticated channel to Bob, and Bob corrects the errors in his string $\tilde{x}_{\text{remain}}$, so that he can obtain x_{remain} from the process as well.

10. Alice and Bob perform privacy amplification: Alice picks a random r, and computes $k = Ext(x_{\text{remain}}, r)$. She sends r to Bob, who computes $k = Ext(\tilde{x}_{\text{remain}}, r)$.

In Protocol 2, Alice and Bob allow for errors under the assumption that all the errors can be caused by a malicious Eve. They bound the amount of min-entropy Eve has about X_{remain} by (i) first invoking Eq. (6.3), and (ii) taking into account

the amount of error correction information C sent across the channel from Alice to Bob. The size of k (given by the number of bits l) then depends on both δ and $|C|$.

Later on we will use the guessing game from the previous chapter in order to prove that the BB84 protocol presented aboveis secure for certain positive values of l.

6.6 Security of BB'84

To prove security of the BB'84 protocol we make two small modifications to the protocol. Although it will at first appear like these modifications give more power to the eavesdropper, they will facilitate the analysis.

6.6.1 A purified protocol

The first modification is rather benign. Consider the following two experiments. In the first experiment, Alice chooses $x, \theta \in \{0,1\}$ uniformly at random and returns $|x\rangle_\theta = H^\theta |x\rangle$, an encoding of the bit x in the basis specified by θ (the standard basis if $\theta = 0$ and the Hadamard basis if $\theta = 1$). In the second experiment, Alice first prepares an EPR pair $|\mathrm{EPR}\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$. She then chooses a $\theta \in \{0,1\}$ uniformly at random and measures the first qubit in the basis $\{|0\rangle_\theta, |1\rangle_\theta\}$, obtaining outcome $x \in \{0,1\}$. She returns the second qubit.

We claim that the two experiment are absolutely equivalent. There are two things to verify. First, while in the first experiment Alice makes a choice of x uniformly at random, in the second experiment x is determined as the outcome of a measurement on the EPR pair. But we know that, since the reduced density matrix of the EPR pair on the first qubit is the totally mixed state, any basis measurement on that qubit will return each of the two possible outcomes with probability $1/2$. So the distribution of x is identical in the two experiments.

Second, we should check that when Alice obtains outcome x by measuring the first qubit of the EPR pair in the basis θ, the qubit she returns, i.e. the second qubit of the EPR pair, is indeed projected onto the state $|x\rangle_\theta$. Again this is a property of the EPR state that is valid for any choice of basis measurement on the first qubit, so we are good — the two experiments are indeed equivalent.

Let us then consider an equivalent formulation of the BB'84 protocol in which, instead of directly preparing BB'84 states, Alice first prepares EPR pairs, keeps the first qubit of each pair to herself, and sends the second qubit to Bob. At a later stage she measures her qubit in a basis $\theta_j \in \{0,1\}$ chosen uniformly at random, and records the outcome x_j.

Thanks to the observation we made above this new formulation of the protocol is completely equivalent to the standard one. Even though it may look more complicated, the essential advantage of the new formulation is that it allows us to delay the moment in the protocol at which Alice needs to make her choice of basis. Although the difference is only conceptual, we can think of this delay as giving less power to Eve: we will now be able to easily argue that certain actions of the eavesdropper, taken early on in the protocol, could not have depended on Alice's basis choice, since the choice has not yet have been made at the time.

Here is the modified protocol in detail. For simplicity we again consider the case where there is no noise. It is called the "purified" BB'84:

Protocol 3 (Purified BB'84 (no noise).)**.** *Outputs $k \in \{0,1\}^\ell$ to both Alice and Bob.*

1. *For a small real-valued parameter $\eta \ll 1$, and a large integer $n \gg 1$, let $N = (4+\eta)n$. Alice prepares N EPR pairs $|\mathrm{EPR}\rangle_{AB}$, and sends the second qubit of each pair to Bob. She chooses a uniformly random basis string $\theta = (\theta_1, \ldots, \theta_N) \in \{0,1\}^N$ and measures each of her qubits in the bases θ to obtain a string $x = x_1, \ldots, x_N$.*

2. *Bob chooses a uniformly random basis string $\tilde{\theta} = (\tilde{\theta}_1, \ldots, \tilde{\theta}_N) \in \{0,1\}^N$. He measures the j-th qubit he received from Alice in the basis $\tilde{\theta}_j$ to obtain outcome $\tilde{x}_j$.*

3. *Bob tells Alice over the CAC that he received and measured all the qubits.*

4. *Alice and Bob exchange their basis strings θ and $\tilde{\theta}$ over the CAC.*

5. *Alice and Bob throw away the data from all rounds $j \in \{1, \ldots, N\}$ in which they didn't measure in the same basis. Let $S = \{j | \theta_j = \tilde{\theta}_j\}$ denote the indices in which they measured in the same basis.*

6. *Alice picks a random subset $T \subseteq S$ of size $T \approx |S|/2$ for testing and tells Bob T over the CAC.*

7. *Alice and Bob announce x_T and $\tilde{x}_T$ to each other over the CAC. They compute the error rate $\delta = W/|T|$, where $W = |\{j \in T \mid x_j \neq \tilde{x}_j\}|$ is the number of disagreements found in T. If δ is too large, they abort the protocol.*

8. *Let $R = S \backslash T$. Alice and Bob perform information reconciliation and privacy amplification on x_R.*

The idea of considering a purified variant of the BB'84 protocol can be traced back to a different proposal for quantum key distribution put forward by Ekert in 1991 [**ekert1991quantum**]. Ekert's main insight was that if Alice and Bob were able to test for the presence of entanglement between their qubits, then (intuitively) by the monogamy of entanglement they would be able to certify that their systems are uncorrelated with Eve's. We will explore Ekert's protocol (and prove the intuition correct!) next week when we analyze quantum key distribution in the so-called device-independent setting.

Remark 6.6.1. *Even though the purified protocol requires Alice to prepare EPR pairs, the formulation will only be used for the purposes of analysis. As we already discussed, from the point of view of any eavesdropper which protocol Alice and Bob actually implement makes no difference at all, so it is perfectly fine to prove security of the purified protocol but use the original BB'84 protocol in practice. This is convenient because it is much easier to prepare single-qubit BB'84 states than to distribute EPR pairs across long distances.*

6.6.2 More power to the eavesdropper

The second modification we make to the BB'84 is less benign, and will appear to give much more power to the eavesdropper. But once again it will be convenient for the analysis. Moreover, if we can prove security against stronger eavesdroppers without too much extra effort, why not do it?

The motivation for this second modification is that it is very hard to model the kinds of attacks Eve might apply to the quantum communication channel between Alice and Bob. For example, she might partially entangle herself with the qubits sent by Alice, creating a joint state ρ_{ABE} on which we have little control.

Exercise 6.6.1. *Consider the case of a single EPR pair ($n = 1$), and suppose that Eve applies a CNOT on her qubit $|0\rangle_E$, controlled on the qubit B that Alice sends to Bob (Eve then forwards the qubit over to Bob). Compute the resulting joint state ρ_{ABE}. Compute the probability that Alice and Bob choose the same basis $\theta = \tilde{\theta}$ and obtain $x = \tilde{x}$. Is this a good attack?* ∎

Because it is hard to model general intercepting attacks of the form described in the exercise, we will modify the protocol by allowing Eve to prepare an arbitrary pure state ρ_{ABE}, where the A and B systems are each made of N qubits, then give A to Alice, B to Bob, and keep E to herself. Alice and Bob will each measure their respective qubits using random choices of bases as in the protocol, and proceed from there on. By giving more power to Eve (she prepares the states, instead of Alice) we're preventing ourselves from thinking too hard about having a model for the attacks: in the new setup, Eve can prepare any state she likes!

This may sound crazy: if we let the eavesdropper prepare any state, then why doesn't she choose, say, $\rho_{ABE} = |000\rangle^{\otimes N}_{ABE}$? Observe that such a state would pass the "matching outputs" test when $\theta_j = \tilde{\theta}_j = 0$ (standard basis), but it would completely fail whenever $\theta_j = \tilde{\theta}_j = 1$ (Hadamard basis). So even though we're claiming Eve could prepare any state she likes, not all states will be accepted by Alice and Bob in the "matching outputs" test they perform in Step 7. How powerful is this test? Can it be used to certify that the state handed over by Eve indeed has the correct form, of being (close to) a tensor product of N EPR pairs? This may sound surprising, as the test only involves local measurements: can local measurements really detect entanglement? The answer is yes. Let's see how it works.

6.6.3 Locally implementing an entangled measurement

Suppose we modified the purified BB'84 protocol by adding an initial step as follows:

0. Upon receiving their N respective qubits from Eve, Alice and Bob jointly measure each pair of qubits using the two-outcome POVM $\{|\mathrm{EPR}\rangle\langle\mathrm{EPR}|_{AB}, \mathbb{I}_{AB} - |\mathrm{EPR}\rangle\langle\mathrm{EPR}|_{AB}\}$, where $|\mathrm{EPR}\rangle_{AB}$ denotes the EPR pair on Alice Bob's joint system. If the number of pairs of qubits that were not found to equal $|\mathrm{EPR}\rangle_{AB}$ is larger than δn they abort Protocol 3. Otherwise, they proceed as usual.

With this modification the protocol is clearly secure, and the tests performed in step 7 have become superfluous. Indeed, after the completion of step 0, Alice and Bob already have the guarantee that at least $(1-\delta)n$ of their shared pairs of qubits are perfect EPR pairs (since they are projected in the post-measurement state $|\mathrm{EPR}\rangle$). In particular, any bit of the raw key obtained from measurements on these states is perfectly uniform and uncorrelated with Eve (remember that correlations generated from pure states are always perfectly monogamous).

The problem with step 0 is that it requires Alice and Bob to perform a joint entangled measurement, which they cannot implement locally. Or can they?

Exercise 6.6.2. *Suppose given a tripartite state ρ_{ABE}, where A and B are each systems of a single qubit. Show that the probability that a measurement of systems A and B in the standard basis results in matching outcomes is exactly $\mathrm{Tr}(\Pi_1 \rho_{AB})$, where*

$$\Pi_1 = |\mathrm{EPR}\rangle\langle\mathrm{EPR}| + |\Psi_{01}\rangle\langle\Psi_{01}|, \qquad \textit{and} \qquad |\Psi_{01}\rangle = \frac{1}{\sqrt{2}}|00\rangle - \frac{1}{\sqrt{2}}|11\rangle\,. \tag{6.4}$$

Similarly, show that if the measurement is performed in the Hadamard basis then the probability of obtaining matching outcomes is $\mathrm{Tr}(\Pi_2\rho_{AB})$*, with*

$$\Pi_2 = |\mathrm{EPR}\rangle\langle\mathrm{EPR}| + |\Psi_{10}\rangle\langle\Psi_{10}|, \qquad \textit{and} \qquad |\Psi_{10}\rangle = \frac{1}{\sqrt{2}}|01\rangle + \frac{1}{\sqrt{2}}|10\rangle\,. \tag{6.5}$$

■

The exercise suggests that the "matching outcomes" test that Alice and Bob implement in step 7 of the original Protocol 3 is essentially equivalent to its replacement step 0 introduced above. Therefore, the security of Protocol 3 with step 0 implemented should directly imply the security of the protocol without step 0, but with step 7 instead.

To express the relation between the two steps, consider the reduced density matrix $\rho_{A_jB_j}$ of the state prepared by Eve on any two qubits for Alice and Bob, and let p_j be the probability of succeeding in the "matching outcomes" test, averaged over the choice of a uniformly random (but identical for both A_j and B_j) basis in which to perform the test. Then by expressing $\rho_{A_jB_j}$ in the Bell basis as

$$\rho_{A_jB_j} = q_{00}|\mathrm{EPR}\rangle\langle\mathrm{EPR}| + q_{01}|\Psi_{01}\rangle\langle\Psi_{01}| + q_{10}|\Psi_{10}\rangle\langle\Psi_{10}| + q_{11}|\Psi_{11}\rangle\langle\Psi_{11}|, \tag{6.6}$$

using the expressions for Π_1 and Π_2 obtained in Exercise 6.6.2 you can check that the condition

$$q_{00} = \langle\Psi_{00}|\,\rho_{A_jB_j}\,|\Psi_{00}\rangle \geq 2p_j - 1 \tag{6.7}$$

is satisfied. In particular, if the probability of success in the test is close to 1, say $p_j = 1-\delta$ for some small δ, then the overlap q_{00} is correspondingly large, at least $1-2\delta$.

If the test performed in step 7 of Protocol 3 was really equivalent to our hypothetical step 0 projecting all pairs of qubits on EPR pairs, then we would be done with our proof of security. However, although the intuition is valid the argument is not quite complete: the two tests are not *exactly* equivalent, and making the argument precise is going to require more work.

A first distinction is that, in Protocol 3, step 7 is performed on the rounds T selected for testing, whereas it is the rounds in $R = S\backslash T$ that are used for the raw key (the outputs used for the raw key are never tested for equality, as this would leak them to Eve!). Another difficulty is that the results of the tests performed in different rounds are not independent from each other: although Alice and Bob make independent measurements, the state ρ_{ABE} prepared by Eve does not necessarily have a tensor product form.

The second distinction raises a thorny difficulty, to which we'll return in more detail next week. For now, let's concentrate on the first objection: how do we infer conditions on the qubits in rounds $j \in S\backslash T$ from results of tests performed on the qubits in rounds $j \in T$?

6.6.4 A concentration inequality

Let us summarize the situation. Suppose for simplicity that the number $|S|$ of rounds in which Alice and Bob make the same basis choice is exactly $|S| = 2n$, and that T has size $|T| = |S|/2 = n$. For each $j \in S$, introduce an indicator random variable $Z_j \in \{0, 1\}$ such that $Z_j = 1$ indicates failure in the matching outcomes test: $Z_j = 0$ if and only if $x_j = \tilde{x}_j$. With this notation the condition verified by Alice and Bob at step 7 of Protocol 3 can be written as $\sum_{j\in T} Z_j \leq \delta|T|$. In order to analyze security of their key, however, they would like to bound $\sum_{j\in S\backslash T} Z_j$. How can we do this?

The key idea is to use the fact that T is chosen as a random subset. Intuitively the average number of failures in T should be about the same as the average in the whole of S: indeed, which rounds are included in T or not is chosen at random by Alice, independently from whether the outcomes in those rounds happened to match or not.

The main tool required to make this intuition precise is called a concentration bound. There are many such bounds. The most widely used are usually referred to as the "Chernoff bound" or "Hoeffding's inequality", which is a generalized version of the Chernoff bound. If you have never heard of them, go look them up! The following is a variant of the Chernoff bound that turns out to be perfectly tuned for our scenario:

Theorem 6.6.2 (Lemma 7 in [**tomamichel2015rigorous**])**.** *Let $m = n + k$ and consider binary random variables $X_1, \ldots, X_m$. (The X_i may be arbitrarily correlated.) Let T be a uniformly random subset of $\{1, \ldots, m\}$ of size k. Then for any $\delta, \nu > 0$,*

$$\Pr\left(\sum_{j\in T} X_j \leq \delta k \quad \wedge \quad \sum_{j\in\{1,\ldots,m\}\backslash T} X_j \geq (\delta + \nu)n \right) \leq e^{-2\nu^2 \frac{nk^2}{(n+k)(k+1)}}. \tag{6.8}$$

To see what the theorem says in our setting, set $m = |S| = 2n$ and $k = n$. Let's also choose $\nu = \delta$ for convenience. Plugging in these parameters we get the bound

$$\Pr\left(\sum_{j\in T} Z_j \leq \delta n \quad \wedge \quad \sum_{j\in S\backslash T} Z_j \geq 2\delta n \right) \leq e^{-\delta^2 \frac{n^2}{n+1}}, \tag{6.9}$$

which is valid for any choice of $\delta > 0$. This bound implies that the probability that the test performed in step 7 passes, but the outcomes obtained in the non-tested rounds $R = S\backslash T$ do not match in a fraction larger than 2δ of these rounds, is tiny — exponentially small in n! Writing ABORT to denote the event that Alice and Bob abort in Step 7 of Protocol 3, we can use Bayes' rule to rewrite the bound above as

$$\Pr\left(\sum_{j\in S\backslash T} Z_j \geq 2\delta n \,\middle|\, \neg\,\mathrm{ABORT}\right) \leq \frac{e^{-\delta^2 n}}{\Pr(\neg\,\mathrm{ABORT})}. \tag{6.10}$$

Writing the bound in this way points to an important subtlety in how the security of quantum key distribution is defined. As you can see, the bound is only good if $\Pr(\neg\,\mathrm{ABORT})$ is not too small; if this probability was extremely tiny, then the right-hand side of Eq. (6.10) would suffer a corresponding blow-up. The probability that the protocol does not abort is not something that we can control or test, and it is natural that this probability has to be taken into account when defining security: we should always allow the protocol to have a very small probability of not aborting, in which case no claim can be made on the security.

Unfortunately we are still not done, due to the issue of dependencies between different tests, which may arise due to the eavesdropper preparing a state that is not in tensor product form.

If the state ρ_{ABE} *is* a tensor product across different rounds, i.e. it is of the form $\rho_{ABE} = \otimes_{j=1}^{n} \rho_{A_j B_j E_j}$ then we can complete the proof. Using that the choice of basis $\theta_j, \tilde{\theta}_j$ for $j \in R$ is uniformly random, we can conclude from the bound in Eq. (6.10) that a large fraction of $j \in R$ are such that the state $\rho_{A_j B_j}$ *would* pass the matching outcomes test, in *both* bases, with high probability (this is because, if it were not the case, there would be a sufficiently high chance that we make a choice of basis with respect to which the state fails the test, leading to a contradiction with Eq. (6.10)). From the analysis in Section 6.6.3 we can deduce that $\rho_{A_j B_j}$ has a correspondingly large overlap with an EPR pair, and thus that the outcomes obtained by Alice and Bob when measuring in the same basis are highly correlated with one another, but (due to monogamy) have very weak correlation with Eve's system. Working out the parameters will give us a bound on the min-entropy of X_j in each round $j \in R$, which can be added up over all rounds by using the independence of different rounds.

If the state ρ_{ABE} is not a tensor product, unfortunately the analysis becomes more difficult; for instance the min-entropy does not add up easily across rounds. We will give a detailed analysis under the independence assumption next week, and we will also discuss the non-independent case in greater detail.