Unconditionally secure key-agreement : two case studies.

Pascal Junod, pascal.junod@switzerland.org

Contents

List of Figures

to Mimi

1 Acknowledgments

First of all, I would thank Prof. Ueli Maurer for the freedom he gave me for the subject of this "semester thesis". To have the possibility of doing a little work in his research area was very exciting and stimulating.

Then, I would thank Dr. Stefan Wolf, my supervisor, for his help for letting me understand the topic, for the helpful exchange of ideas, and for motivating and stimulating me to try, to try and to try much more.

Last but not least, I would like to give special thanks to Reto and to Bartosz.

2 Unconditionally secure secret-key agreement

2.1 Motivation

One of the fundamental problems in cryptography is to exchange a message between Alice, the sender, and Bob, the receiver such that an eavesdropper, Eve, cannot have any information about this message, or in other words, that Eve has no other possibility as guessing the message, even with an infinite amount of computer power. Cryptosystems which allow such a level of security, the perfect secrecy, are said to be *information theoretically secure*.

At this point, we can notice that the majority of the cryptosystems which are used today *don't* allow such a level of security; their are based on the assumption that the computer power of Eve is finite, that it would take too much time (months, years, or even more) for the most powerful machines to break these algorithms. Their are said to be *computationally secure*.

Shannon showed that communication can only take place in perfect secrecy if the key used for the encryption has at least so much entropy as the message itself. Therefore, for a long time, it was widely believed that perfect secrecy is not practical to use.

Recently it was pointed out by researchers that Shannon's assumptions for his theorem are unnecesseraly unrealistic : in its model, Eve has a perfect knowledge of the cipher text. It was shown that the noise of physical communication channels, which is a natural property, can be used to generate an (almost) perfect secure key; furthermore, it was shown that it is beneficial for Alice and Bob to exchange additionally messages over an other channel, even if it is totally insecure. A goal of information theory being to find good codes which allow to communicate over a noisy channel, it is very surprising to know that this noise can be used for cryptographic purposes.

Another interesting direction for the purpose of generating perfect secure keys is the quantum cryptography, which is based on Heisenberg's uncertainty principle. By using this fact, which states that there are pairs of incompatible properties in the quantum world, i.e., that measuring a property (such that the polarization of a photon) randomizes necessarily the other, Benett et al [1] introduced a basic protocol for secret-key agreement. We must note that quantum cryptography is expensive to implement, which is not the case in using the noise of communication channels.

A direction of the recent research in the area of noisy channels is to characterize the situations where it is possible to agree on a perfect secret key. The study of special scenarii goes in this way, the ultimate goal being to find

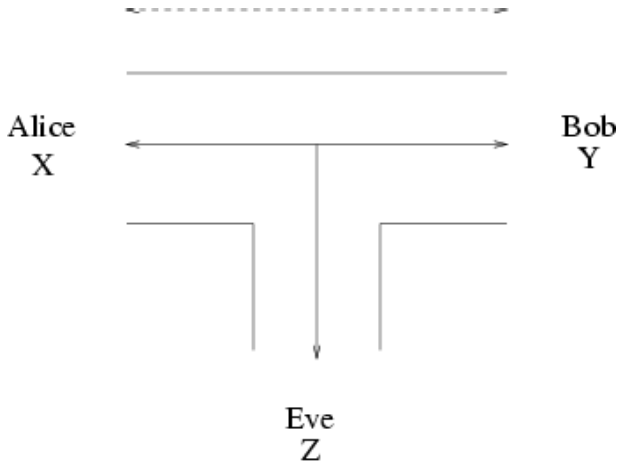

Figure 1: The general scenario

when secret-key agreement is possible for the general case, which is characterized by a probability distribution P_{XYZ} without special properties.

2.2 The secret-key rate

As pointed out in the previous section, the goal is to characterize mathematically the situations where perfect secret-key agreement is possible.

Consider the general scenario (see Figure (1)), which was first described in [5] : Alice, Bob and Eve have access to repeated, independent realizations of random variables X, Y and Z, respectively, with joint probability distribution P_{XYZ}. A special scenario could be the following : P_{XY}, P_{XZ} and P_{YZ} all describe independent binary symmetric channels, for example. Assume that Eve has no information about X and Y other than through her knowledge of Z. Furthermore, assume that Alice and Bob can communicate over an insecure, but *authenticated* channel. Finally, Alice and Bob know the distribution P_{XYZ}.

A protocol for this general scenario can be described as follows : at each step, either Alice or Bob sends a message to Bob or vice-versa. These messages depends on the random variables X and Y and on the messages exchanged in the previous steps. We can denote without loss of generality the messages sent by Alice with $C_1, C_3, C_5, ...$ and those by Bob with $C_2, C_4, C_6, ...$ At the end of a t-steps protocol, Alice and Bob each compute a key S and S', respectively, as functions of their own random variable and the messages $C^t = [C_1, C_2, ..., C_t]$ exchanged over the insecure channel. Their goal is to maximize the entropy $H(S)$ of the key under the condition that S and S' are

the same keys with a high probability and that Eve has very little knowledge about one of these two keys. We can summarize these conditions more formally as follows :

$$
\begin{aligned}
H(C_i|C^{i-1}X) &= 0 \qquad \text{for odd } i & (1)\\
H(C_i|C^{i-1}Y) &= 0 \qquad \text{for even } i & (2)\\
H(S|C^tX) &= 0 & (3)\\
H(S'|C^tY) &= 0 & (4)\\
P[S \neq S'] &\leq \epsilon & (5)\\
I(S;C^tZ) &\leq \delta & (6)
\end{aligned}
$$

where ϵ and δ are very small.

The *secret-key Rate* $S(X,Y||Z)$ is defined as the maximal rate at which Alice and Bob can generate a secret key by public discussion :

Definition 1 (Secret-Key rate)
The secret-key rate of X and Y with respect to Z, denoted $S(X,Y||Z)$, is the maximum rate at which Alice (X) and Bob (Y) can agree on a secret key S while keeping the rate at which Eve (Z) obtains information arbitrarily small, or more formally, it is the maximal rate R such that $\forall\epsilon > 0$ there exists a protocol for sufficiently large N satisfying (1)-(5) with X and Y respectively replaced by X^N and Y^N satisfying

$$\frac{1}{N}I(S;C^tZ^N) \leq \epsilon \tag{7}$$

and achieving

$$\frac{1}{N}H(S) \geq R - \epsilon \tag{8}$$

2.3 The scenario EC2

To characterize when information theoretically secure secret-key agreement is possible is equivalent to search conditions on P_{XYZ} for which $S(X,Y||Z) \neq 0$. It seems to be very hard to solve this problem for general joint probability distributions. Therefore, different special scenarii are to be investigated (see [4], e.g.).

One of them is the so-called *scenario EC2*, which is presented in Figure (2). Bob's information about Alice's random variable is biased by a binary symmetric channel with a probability ϵ. Eve receives X and Y through two independent erasure channels which have an erasure probability of $1 - r_X$ and of $1 - r_Y$, respectively. Or in a few words, Bob always receives Alice's bit, but this bit is perhaps false, while Eve don't receive always Alice's and

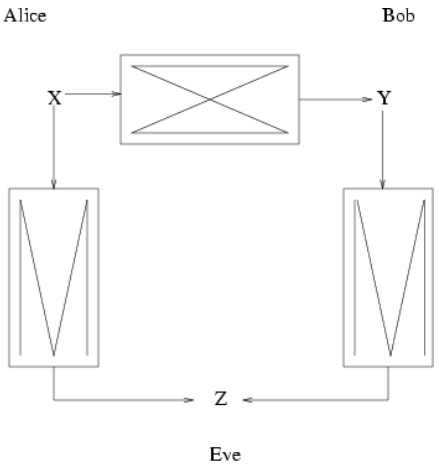

Figure 2: The scenario EC2

Bob's bits, but for her, a received bit is a true bit. At first sight, it can be surprising that perfect secure secret-key agreement is possible in this situation; the possibility of exchanging messages over the insecure channel allows it.

2.3.1 Protocol RC

By using protocol RC ("Repeat-Code"), which is described below, one can show that secret key agreement is possible.

Protocol RC

Let N be fixed. Alice chooses a random bit C, and she sends

$$[C \oplus X_1, C \oplus X_2, ..., C \oplus X_N]$$

over the public channel. Bob computes

$$[(C \oplus X_1) \oplus Y_1, ..., (C \oplus X_N) \oplus Y_N]$$

and accepts exactly if and only if this is equal to either $[0, ..., 0]$ or $[1, ..., 1]$.

In this situation, Alice and Bob make use of a repeat code of length N with only the two codewords $[0, ..., 0]$ and $[1, ..., 1]$. The trick here is that they reduce their error probability by using codewords of length N and also that they improve their situation compared with the opponent's by accepting the bit only in case of highly reliable communications.

First one can note that the adversary Eve takes advantage from a greater N, too. Secondly it's clear that Eve's optimal strategy in this case for guessing C is to compute the block $[(C \oplus X_1) \oplus Z_1, ..., (C \oplus X_N) \oplus Z_N]$ and to take a majority decision about the bits in this block.

The analysis of scenario EC2 with protocol RC was done in [4]. In the following, we use this notation : let α be the probability that $X \neq Y$, and let r_X and r_Y be the probabilities that Eve *reads* respectively X and Y. We assume that Alice's information is better protected than Bob's : $r_Y \geq r_X$. A first upper bound on r_X comes from the following theorem, which was first proved in [6]. A simplified version of the proof is in [8].

Theorem 1
For every distribution P_{XYZ},

$$S(X;Y \parallel Z) \geq \max\{I(X;Y) - I(X;Z), I(Y;X) - I(Y;Z)\} \qquad (9)$$

The following lemma comes from [8]. We give a slightly more detailed proof:

Lemma 1
In scenario EC2, $S(X;Y \parallel Z)$ is strictly positive if

$$r_X \leq 1 - \frac{h(\alpha)}{1 - r_Y + r_Y h(\alpha)}$$

Proof :

We have $r_X \geq r_Y$, so the term $I(X;Y) - I(X;Z)$ in (9) applies. We have :

$$\begin{aligned} I(X;Y) &> I(X;Z) \\ H(X) - H(X|Y) &> H(X) - H(X|Z) \\ -H(X|Y) &> -H(X|Z) \\ H(X|Y) &< H(X|Z) \end{aligned}$$

We now have to compute these two values.

$$\begin{aligned} H(X|Y) &= h(\alpha) \\ H(X|Z) &= 1 \cdot P_Z[\Delta\Delta] + h(\alpha) P_Z[\Delta 0] + h(\alpha) P_Z[\Delta 1] \end{aligned}$$

where

$$P_Z[\Delta\Delta] = (1 - r_X)(1 - r_Y)$$

and

$$\begin{aligned} P_Z[\Delta 0] = P_Z[\Delta 1] &= 0.5(1-r_X)r_Y\alpha + 0.5(1-r_X)r_Y(1-\alpha) \\ &= 0.5(1-r_X)r_Y. \end{aligned}$$

This gives us

$$H(X|Z) = (1-r_X)(1-r_Y) + h(\alpha)(1-r_X)r_Y$$

We have as straightforward calculation :

$$\begin{aligned} h(\alpha) &< (1-r_X)(1-r_Y) + h(\alpha)(1-r_X)r_Y \\ &= 1 - r_X - r_Y + r_X r_Y + h(\alpha)r_Y - h(\alpha)r_X r_Y \\ &= 1 - r_X\left(1 + r_Y h(\alpha) - r_Y\right) - r_Y + h(\alpha)r_Y. \\ r_X\left(1 - r_Y + h(\alpha)r_Y\right) &< 1 - h(\alpha) - r_Y + h(\alpha)r_Y \\ &= 1 - h(\alpha) + r_Y\left(h(\alpha) - 1\right). \\ r_X\left(1 + r_Y\left(h(\alpha) - 1\right)\right) &< \left(h(\alpha) - 1\right)\left(r_Y - 1\right). \\ r_X &< \frac{\left(h(\alpha) - 1\right)\left(r_Y - 1\right)}{1 + r_Y\left(h(\alpha) - 1\right)} \\ &= \frac{r_Y - 1}{\frac{1}{h(\alpha)-1} + r_Y} \\ &= \frac{r_Y - 1}{\frac{1 + r_Y(h(\alpha)-1)}{h(\alpha)-1}} \\ &= \frac{\left(r_Y - 1\right)\left(h(\alpha) - 1\right)}{1 + r_Y\left(h(\alpha) - 1\right)} \\ &= \frac{r_Y h(\alpha) - h(\alpha) - r_Y + 1}{1 + r_Y h(\alpha) - r_Y} \\ &= \frac{1 + r_Y h(\alpha) - r_Y}{1 + r_Y h(\alpha) - r_Y} - \frac{h(\alpha)}{1 + r_Y h(\alpha) - r_Y} \\ &= 1 - \frac{h(\alpha)}{1 + r_Y\left(h(\alpha) - 1\right)}. \end{aligned}$$

which concludes the proof.

□

2.3.2 Protocol RCE

At this point, one can note that it can be surprising to know that protocol RC allows secret-key agreement, because this code doesn't seem to be very

appropriate in a situation where the adversary Eve has a perfect knowledge of X or Y with some positive probability. Revealing one bit of the repeat-code block means revealing the entire block. A protocol using blocks which contain a certain number of incorrect bits (less than the half) is better in this situation; on the other side, the effect that Alice's and Bob's bits become more reliable is weaker with this protocol. Protocol RCE ("Repeat-Code with Errors") was proposed in [8, 4] :

Protocol RCE

Let N be fixed. Bob randomly chooses a bit C and a random N-bit block $[C_1, ..., C_N]$ such that tN of the bits are equal to C and $(1-t)C$ are equal to its complement $C' := C \oplus 1$. $t > 1/2$ is a parameter of the code, and tN is an integer. Bob compute

$$[C_1 \oplus Y_1, ..., C_N \oplus Y_N]$$

and sends this block over the public channel. Alice computes

$$[(C_1 \oplus Y_1) \oplus X_1, ..., (C_N \oplus Y_N) \oplus X_N]$$

and accepts if and only if this equals $[0, ..., 0]$ or $[1, ..., 1]$.

We can note that protocol RCE corresponds to Protocol RC for $t = 1$. RCE is, as protocol RC, efficient in terms of computation but wasteful with respect to the achievable rate of generated secret-key.

2.3.3 The analysis of protocol RCE

An analysis of Scenario EC2 with protocol RCE is made in [8, 4]. First, we can compute the conditional probability β_N that *Alice receives the bit sent by Bob incorrectly, given that she accepts.* We have :

$$\beta_N = \frac{\alpha^{tN}(1-\alpha)^{(1-t)N}}{(1-\alpha)^{tN}\alpha^{(1-t)N} + \alpha^{tN}(1-\alpha)^{(1-t)N}} \tag{10}$$

Let be

$$K = K(t) := \frac{1}{4t-2} \tag{11}$$

The equation (10) can be approximated as follows :

Lemma 2

$$\frac{\alpha^{tN}(1-\alpha)^{(1-t)N}}{(1-\alpha)^{tN}\alpha^{(1-t)N} + \alpha^{tN}(1-\alpha)^{(1-t)N}} \leq \left(\frac{\alpha}{1-\alpha}\right)^{N/(2K)} \tag{12}$$

Proof :

$$\begin{aligned}
\frac{\alpha^{tN}(1-\alpha)^{(1-t)N}}{(1-\alpha)^{tN}\alpha^{(1-t)N}+\alpha^{tN}(1-\alpha)^{(1-t)N}} &= \frac{\alpha^{tN}(1-\alpha)^{(1-t)N}}{\alpha^{tN}\left((1-\alpha)^{tN}\alpha^{(1-2t)N}+(1-\alpha)^{(1-t)N}\right)} \\
&= \frac{(1-\alpha)^{(1-t)N}}{(1-\alpha)^{tN}\alpha^{(1-2t)N}+(1-\alpha)^{(1-t)N}} \\
&= \frac{(1-\alpha)^{(1-t)N}}{(1-\alpha)^{(1-t)N}\left(1+\alpha^{(1-2t)N}(1-\alpha)^{(2t-1)N}\right)} \\
&= \frac{1}{1+\alpha^{-(2t-1)N}(1-\alpha)^{(2t-1)N}} \\
&= \frac{1}{1+(\frac{1-\alpha}{\alpha})^{(2t-1)N}} \\
&\leq \frac{1}{\left(\frac{1-\alpha}{\alpha}\right)^{(2t-1)N}} \\
&= \left(\frac{\alpha}{1-\alpha}\right)^{(2t-1)N}
\end{aligned}$$

According to (11), we have finally

$$\frac{\alpha^{tN}(1-\alpha)^{(1-t)N}}{(1-\alpha)^{tN}\alpha^{(1-t)N}+\alpha^{tN}(1-\alpha)^{(1-t)N}} \leq \left(\frac{\alpha}{1-\alpha}\right)^{N/(2K)} \tag{13}$$

which concludes the proof.

□

It seems difficult to give the optimal strategy for Eve in the case of protocol RCE in the scenario EC2. Such a strategy will clearly minimize her error probability. We summarize the different possibilites for Eve to make an error in an event tree (see Figure (3)). First of all, Alice can accept the bit (`AliceB`) or not (`notAliceB`). In the latter case, which is clearly the most frequent one, the protocol go on. In the former case, Alice accepts a wrong bit (`AliceW`) or not (`notAliceW`). The latter situation brings us to `A`. Now, Eve's channel with Alice can be informative (`ChannelA`), with a more or less high probability (in this case, one character only need to be read by Eve), or not informative (`notChannelA`). In the former case (situation `B`), if Eve knows one bit of the block of Alice, then her error probability γ_N is equal to 0. Finally, when Eve receives sN of the tN correct bits of Bob's block and exactly the same number of incorrects bit, and that she learns nothing about Alice's block (she receives only erasure symbols), then we are in the situation `C` if we take the

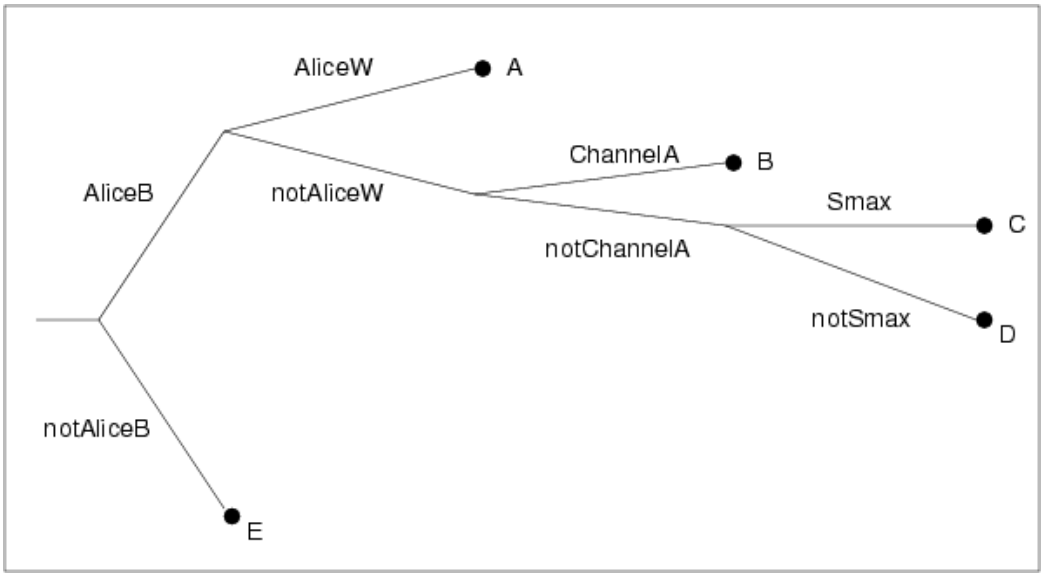

Figure 3: An event tree for Eve's error probability

s that maximizes Eve's error probability, and in the situation `D` otherwise. We have seen that in situation `B`, Eve's error probability is clearly defined. In case of `C`, the optimal strategy for Eve is to flip a fair coin, because she has no information about Bob's block, half of the bits being wrong, the other half being right. This gives us a lower bound for Eve's error probability :

$$\gamma_N \geq \frac{1}{2} \cdot \max_{0 \leq s \leq (1-t)} \left\{ \binom{tN}{sN} (r_Y)^{sN} (1-r_Y)^{(t-s)N} \right. \\ \cdot \binom{(1-t)N}{sN} (r_Y)^{sN} (1-r_Y)^{(1-t-s)N} \\ \left. \cdot (1-r_X)^N \right\} \tag{14}$$

Otherwise, situation `D` is uninteresting because of the asymptotic behavior of the binomial coefficients, and situation `A` can slightly increase Eve's error probability, if her channel with Bob is uninformative. The most interesting potential improvement is the case where the number of false bits which Eve receives from Bob is not equal to the number of true bits. Unfortunately, it is difficult to find the right weights of the different cases in the expression.

To give an analysis of $S(X, Y||Z)$ in this case, one can use Lemma (3) (see [4, 8]).

Lemma 3
Let X, Y and Z be arbitrary random variables, and let C be a bit, randomly chosen by Alice. Assume that for all N, Alice can generate a message M from X^N and C (and possibly some random bits) such that with some probability $p_{\alpha,N} > 0$, Bob (who knows M and Y^N) publicly accepts and can compute a bit C' such that $P[C \neq C'] \leq b^N$ for some $b \geq 0$. If in addition, given that Bob accepts, for every strategy for guessing C when given M and Z^N, the average error probability γ_N of Eve is at least c^N for some $c > b$ and for sufficiently large N, then $S(X, Y||Z) > 0$.

As approximation of (14), the following is used in [4, 8]:

Lemma 4
The lower bound (14) implies that

$$\gamma_N^{2K/N} \geq 1 - \frac{1}{4K} - \frac{1}{16(1-r_Y)K} - 2Kr_X \tag{15}$$

if $r_Y \leq 1 - t$ holds, and if N is sufficiently large.

To find a good condition on r_X, the idea is to find the best possible choice for $K := K(t)$ with respect to the fixed parameters α and r_Y. Furthermore, the condition on r_Y in lemma (4) must hold. This optimal choice of K leads to an upper bound on r_X, such that if r_X is smaller than this bound, then protocol RCE works for secret-key agreement. By using lemma (3), which states that it's sufficient for secret-key agreement by public discussion if Eve's error probability about the bit sent by Bob is asymptotically greater than Alice's error probability for $N \longrightarrow \infty$, it is possible to prove the following bound:

Theorem 2
In scenario EC2, protocol RCE allows for secret-key agreement, and $S(X, Y||Z)$ is hence positive, if

$$r_X < \frac{2(1 - \frac{\alpha}{1-\alpha})^2(1 - r_Y)}{5 - 4r_Y} \tag{16}$$

(when $1 - \alpha/(1-\alpha) \leq 5/4 - r_Y$), or if

$$r_X < (1 - r_Y)\left(1 - \frac{\alpha}{1-\alpha} - \frac{1 - r_Y}{2} - \frac{1}{8}\right) \tag{17}$$

(when $1 - \alpha/(1-\alpha) > 5/4 - r_Y$).

2.4 Towards a new lower bound

We give now our own attempts to analyse the protocol RCE in the scenario EC2. The basis is a different estimation of expression (14). We need first a mean to approximate the binomial coefficients, which are not very convenient to compute with. The following result comes from the well-known Stirling formula :

Lemma 5
For a constant C and a sufficiently large N, we have :

$$\binom{aN}{bN} \geq \frac{C}{\sqrt{N}} \left(\frac{a^a}{b^b (a-b)^{a-b}} \right)^N .$$

Proof :

Stirling's formula states that

$$\sqrt{2\pi} N^{N+1/2} e^{-N} e^{(12N+1)^{-1}} < N! < \sqrt{2\pi} N^{N+1/2} e^{-N} e^{(12N)^{-1}}$$

Therefore, we have:

$$\begin{aligned}
\binom{aN}{bN} &= \frac{(aN)!}{(bN)!\,((a-b)N)!} \\
&\geq \frac{\sqrt{2\pi}\,(aN)^{aN+1/2}\, e^{-aN} e^{(12aN+1)^{-1}}}{\sqrt{2\pi}\,(bN)^{bN+1/2}\, e^{-bN} e^{(12bN)^{-1}} \sqrt{2\pi}\,((a-b)N)^{(a-b)N+1/2}\, e^{-(a-b)N} e^{(12(a-b)N)^{-1}}} \\
&= \frac{1}{\sqrt{2\pi}} \cdot \frac{e^{(12aN+1)^{-1}}}{e^{(12(a-b)N)^{-1}} e^{(12bN)^{-1}}} \cdot \frac{(aN)^{aN+1/2}}{(bN)^{bN+1/2} ((a-b)N)^{(a-b)N+1/2}} \\
&= \frac{1}{\sqrt{2\pi N}} \cdot \frac{e^{(12aN+1)^{-1}}}{e^{(12(a-b)N)^{-1}} e^{(12bN)^{-1}}} \cdot \frac{\sqrt{a}}{\sqrt{b}\sqrt{a-b}} \cdot \left(\frac{a^a}{b^b (a-b)^{a-b}} \right)^N \\
&= \frac{1}{\sqrt{2\pi N}} \cdot e^{\frac{-12abN + 12b^2 N + 2a^2 N + a}{12(12aN+1)(b-1)bN}} \cdot \frac{\sqrt{a}}{\sqrt{b}\sqrt{a-b}} \left(\frac{a^a}{b^b (a-b)^{a-b}} \right)^N
\end{aligned}$$

Let C:=$\frac{\sqrt{a}}{\sqrt{b}\sqrt{a-b}\sqrt{2\pi}}$. The second part of the product tends to 1 as $N \to \infty$. Thus, we conclude that

$$\binom{aN}{bN} \geq \frac{C}{\sqrt{N}} \left(\frac{a^a}{b^b (a-b)^{a-b}} \right)^N .$$

□

Lemma 6
The lower bound (14) implies that

$$\begin{aligned}
\gamma_N^{2K/N} \geq & (1 - 2K r_X) \cdot \left(1 - \frac{1}{2K(1-r_Y)} \right) \\
& \cdot \left(\frac{(4K^2 - 1)(1-r_Y)^{2r_Y}}{(4K^2 (1-r_Y)^2 - 1)^{1-r_Y} (4K)^{2r_Y}} \right)^K
\end{aligned} \tag{18}$$

if $r_Y/2 \leq 1 - t$ holds and if N is sufficiently large.

Proof :

We have to approximate (14). We can first note that $r_Y/2 \leq 1-t$ means that $s := r_Y/2$ is a possible choice; in fact, this is the optimal one. By applying Lemma 5 and by replacing the binomial coefficients by the corresponding expression, we get

$$\begin{aligned}\gamma_N^{1/N} \geq & \frac{t^t}{(\frac{r_Y}{2})^{\frac{r_Y}{2}}(t-\frac{r_Y}{2})^{t-\frac{r_Y}{2}}} \cdot \frac{(1-t)^{1-t}}{(\frac{r_Y}{2})^{\frac{r_Y}{2}}(1-t-\frac{r_Y}{2})^{1-t-\frac{r_Y}{2}}} \\ & \cdot \left(\frac{1}{2}\right)^{1/N} \cdot r_Y^{r_Y}(1-r_Y)^{1-r_Y}(1-r_X) \\ = & \frac{t^t(1-t)^{1-t}}{\left(\frac{t-\frac{r_Y}{2}}{1-r_Y}\right)^{t-\frac{r_Y}{2}} \left(\frac{1-t-\frac{r_Y}{2}}{1-r_Y}\right)^{1-t-\frac{r_Y}{2}}} \cdot 2^{r_Y-\frac{1}{N}} \cdot (1-r_X)\end{aligned}$$

We can now take a first approximation :

$$\lim_{N\to+\infty} 2^{r_Y-\frac{1}{N}} = 2^{r_Y} \tag{19}$$

Furthermore, as r_Y ranges over $[0,1]$, the right part of equation (19) ranges over $[1,2]$. So we can eliminate this term. Our expression resumes now to :

$$\gamma_N^{1/N} \geq \frac{t^t(1-t)^{1-t}}{\left(\frac{t-\frac{r_Y}{2}}{1-r_Y}\right)^{t-\frac{r_Y}{2}} \left(\frac{1-t-\frac{r_Y}{2}}{1-r_Y}\right)^{1-t-\frac{r_Y}{2}}} \cdot (1-r_X) \tag{20}$$

Let be $K := \frac{1}{4t-2}$. We now have

$$(1-t)^{1-t} \cdot t^t = (2K+1)^{\frac{2K+1}{4K}} \cdot (2K-1)^{\frac{2K-1}{4K}} \cdot \frac{1}{4K} \tag{21}$$

Introducing (21) in (20), we get

$$\begin{aligned}\gamma_N^{1/N} \geq & \left(\frac{(2K+1)^{2K+1}(2K-1)^{2K-1}}{\left(\frac{2K(1-r_Y)+1}{1-r_Y}\right)^{2K(1-r_Y)+1} \left(\frac{2K(1-r_Y)-1}{1-r_Y}\right)^{2K(1-r_Y)-1}}\right)^{\frac{1}{4K}} \\ & \cdot (1-r_X) \cdot (4K)^{-r_Y}\end{aligned} \tag{22}$$

As last algebraic modification before doing approximations, we raise (22) to the power $2K$ and we do a straightforward calculation to get

$$\begin{aligned}\gamma_N^{2K/N} \geq & \left(\frac{4K^2-1}{(4K^2(1-r_Y)^2-1)^{1-r_Y}}\right)^K \\ & \cdot \sqrt{\frac{(2K+1)(2K(1-r_Y)-1)}{(2K-1)(2K(1-r_Y)+1)}} \cdot \left(\frac{(1-r_X)(1-r_Y)^{r_Y}}{(4K)^{r_Y}}\right)^{2K} \\ = & \left(\frac{4K^2-1}{(4K^2(1-r_Y)^2-1)^{1-r_Y}}\right)^K \cdot \sqrt{\frac{2K+1}{2K-1}} \cdot \sqrt{\frac{2K(1-r_Y)-1}{2K(1-r_Y)+1}} \\ & \cdot \left(\frac{1}{4K}\right)^{2Kr_Y} \cdot (1-r_X)^{2K} \cdot (1-r_Y)^{2Kr_Y}\end{aligned} \tag{23}$$

Now we can try to approximate (23) to get a lower bound as simple and as tight as possible. First we can note that $\sqrt{\frac{2K+1}{2K-1}} \geq 1 \quad \forall K \geq \frac{1}{2}$. Hence we can eliminate this term from (23). Secondly we have $(1-r_X)^{2K} \geq (1-2Kr_X)$. And finally,

$$\begin{aligned}\sqrt{\frac{2K(1-r_Y)-1}{2K(1-r_Y)+1}} = & \sqrt{1-\frac{2}{2K(1-r_Y)+1}} \\ \geq & \sqrt{1-\frac{1}{K(1-r_Y)}} \\ \geq & 1-\frac{1}{2K(1-r_Y)}\end{aligned}$$

Putting these observations into (23) gives us

$$\begin{aligned}\gamma_N^{2K/N} \geq & (1-2Kr_X) \cdot \left(1-\frac{1}{2K(1-r_Y)}\right) \\ & \cdot \left(\frac{(4K^2-1)(1-r_Y)^{2r_Y}}{(4K^2(1-r_Y)^2-1)^{1-r_Y}(4K)^{2r_Y}}\right)^K\end{aligned} \tag{24}$$

which concludes the proof.

□

Unfortunately, this bound is not very useful because of its complexity. A simple simulation showed that for $t > 0.7$, this bound is better as (4) in more than 90% of the time. But the problem is that to get a result such as Theorem 2, the expression (24) needs to be simplified and approximated a lot more: we need to compare it to Bob's error probability and to derive it to get the optimal K. Because of the complexity of the expression, we could not manage to find this optimal K analytically, even with the help of Maple. However, it is not worth to invest a lot of more time to get a simpler expression.

3 Study of an information theoretic conjecture

In the following, we use the usual notation for information theoretic measures, $H(X)$ for the Shannon-entropy of a random variable X and $I(X;Y)$ for the mutual information between the two random variables X and Y. If not stated otherwise, $log(.)$ is the notation for the logarithm function in base 2. [2] was used as an information theory reference book.

3.1 Introduction

The goal of the second part of this semester thesis is to analyse the following conjecture which comes from [8]:

Conjecture 1
Let X be a binary *random variable and Y, Z and U be random variables such that $YZ \longrightarrow X \longrightarrow U$ is a Markov chain. Let be $I(X;Y) \leq I(X;Z)$. Then*

$$I(U;Y) \leq I(U;Z) \tag{25}$$

We can first note that this conjecture is not verified if X is not a binary random variable. An intuitive counterexample is the following: take two pages of a journal, each with some information about the historical development of the stock market, for example. We can see these two pages as the random variable YZ. Now, we can "process" these two pages to extract some information about a precise stock. This extracted information can be seen as the random variable X. As last step, we cut with a shear through the information X, and wo do this independantly from the extraction of the information. Let's call the result of such a process U. This is easily and intuitively concevaible. Now, assume that X gives more information about Y as about Z. It is possible that the shear has canceled all the information about Z, so that U gives now more information about Z as about Y. So this conjecture cannot hold for general random variable $UXYZ$.

This section is organized as follows: first we present a simulation written to find a counterexample to this conjecture, for the cases of binary variable Y, Z and U and then for ternary ones. Then we present some theoretic views about this result, and the way we did towards the proof of the conjecture.

3.2 Search for a counterexample

To be almost sure that this conjecture holds, we first search a counterexample with the help of the computer. For this purpose, we wrote two programs in C, `conj2.c` and `conj3.c`, which are given in annexe. The next part explains the design of these programs, which were not able to find such a counterexample for YZU being binary random variable (`conj2.c`) and ternary random variables (`conj3.c`).

3.2.1 Pseudo-random numbers generation

First of all, we need a good pseudo-random generator, which must have the two following properties: first, it has to have a long period, because we have to generate several billions of pseudo-random numbers without taking the risk of generating the same situation twice, and second, it has to be sufficiently fast.

Because we are generating more than 100'000'000 random numbers in our search, a good pseudo-random generator is `pj_ran3()`, as recommended in [7], Chapter 7. This generator has a period $> 2 \cdot 10^{18}$, which is extremely large. The basic idea of this algorithm is the combination of two different sequences with different periods so as to obtain a new sequence whose period is at least the least common multiple of the two periods. This idea comes from [3] and the reference implementation from [7].

3.2.2 The search algorithm

We present here the complete procedure for searching counterexamples. We treat the two cases (binary and ternary YZU) together, because of their similarity.

- A random variable P_{YZ} is randomly generated: for the binary case, random numbers $r_1, r_2 \in_R [0,1]$ are generated, $P_Y := \{r_1, 1-r_1\}$ and $P_Z := \{r_2, 1-r_2\}$ for the binary case, and $r_1, r_2, r_3, r_4 \in_R [0,1]$ are generated, $P_Y := \{r_1, (1-r_1)r_2, (1-r_1)(1-r_2)\}$ and $P_Z := \{r_3, (1-r_3)r_4, (1-r_3)(1-r_4)\}$ for the ternary case. Then, $P_{YZ} := P_Y \cdot P_Z$.

- To do the step $YZ \longrightarrow X$, a stochastic matrix M_1 is randomly generated: $r_i \in_R [0,1]$ are random numbers; $i := 1..4$ for the binary case and $i := 1..9$ for the ternary case

$$M_1 := \begin{pmatrix} r_1 & 1-r_1 \\ \\ ... & ... \\ \\ r_i & 1-r_i \end{pmatrix}$$

  The probability distribution P_X is then computed as $P_X := P_{YZ} \cdot M_1$.

- In an analogous way, the step $X \longrightarrow U$ is computed as follows: $r_{ij} \in_R [0,1]$ are random numbers; the stochastic matrix is the following for the binary case:

$$M_2 := \begin{pmatrix} r_{11} & 1-r_{11} \\ r_{21} & 1-r-21 \end{pmatrix}$$

and for the ternary case :

$$M_2 := \begin{pmatrix} r_{11} & (1-r_{11})r_{12} & (1-r_{11})(1-r_{12}) \\ r_{21} & (1-r_{21})r_{22} & (1-r_{21})(1-r_{22}) \end{pmatrix}$$

Then, $P_U := P_X \cdot M_2$.

- As a last step, the following information theoretic values are computed: $I(X;Y) - I(X;Z)$ and $I(U;Y) - I(U;Z)$, where

$$\begin{aligned} I(X;Y) &= -\sum_X P_X \cdot \log P_X - \sum_X \sum_Y P_{X|Y} \cdot \log P_{X|Y} \\ I(X;Z) &= -\sum_X P_X \cdot \log P_X - \sum_X \sum_Z P_{X|Z} \cdot \log P_{X|Z} \\ I(U;Y) &= -\sum_U P_U \cdot \log P_U - \sum_U \sum_Y P_{U|Y} \cdot \log P_{U|Y} \\ I(U;Z) &= -\sum_U P_U \cdot \log P_U - \sum_U \sum_Z P_{U|Z} \cdot \log P_{U|Z} \end{aligned}$$

  Then, these values are compared.

This process is then iterated $1'000'000'000$ of times for both situations.

This took approximately 16 hours on a Sun workstation with a normal load for each process. As conjectured, *it couldn't be found any counterexample* either for ternary or for binary random variables.

3.3 A theoretic approach

We present now some theoretic considerations about this problem. To see that, we can first note that it is possible to write the random variable YZ as a new one W. We thus have the following situation: $W \longrightarrow X \longrightarrow U$ being a Markov chain and X being a binary random variable.

The key observation is the following: $I(W;U)$ has to be monotone in $I(X;U)$. Or in other words, these two quantites must have the same behaviour, the first being increasing if and only if the second is increasing, or, inversely, they have to be decreasing at the same time. To try to understand more deeply this fact, we use the following notation: let be the following new random variables:

$$P_0 := P_{W|X=0}$$

and

$$P_1 := P_{W|X=1}$$

Furthermore, let be

$$p_u := P_{X|U=u}(0)$$

We can now write the two quantites $I(W;U)$ and $I(X;U)$, using the definition of the mutual information of two random variables, as follows:
$H(W) - H(W|U)$ and $H(X) - H(X|U)$, respectively.

At this point, we can note that $H(W)$ and $H(X)$ are two constant values, and that only $H(W|U)$ and $H(X|U)$ are varying, with U varying. Thus, the monotony of the two first quantites can be reduced to the mutual monotony of the two latter. Using the notation defined formerly, we can write this as follows:

$$\sum P_U(u) \cdot H(p_u \cdot P_0 + (1 - p_u) \cdot P_1)$$

has to be monotone in

$$-\sum P_Y(y) \log P_Y(y)$$

To prove this conjecture, we have to prove that this condition is holding, or that this condition is a natural property of this kind of Markov chain.

We have studied this condition fore some simple and reduced examples, and this experimentally. It seems that there is no major argument against this fact. But we were unable to prove it analytically and formally. We found very difficult to handle with this kind of mathematical expressions, which are defined on a discrete number of cases, namely the different possible cardinalities of the random variables.

As a conclusion for this part, we claim that we are more convinced that

this conjecture holds than at the beginning, but the lack of formal arguments makes that we can cannot be *totally* convinced !

4 Conclusion

In this semester thesis, we have first studied the foundations of perfect secret-key agreement. Then, a special scenario, the RCE (repeat-code with errors) protocol, which has been proposed and studied in [8], has been considered. In [8], the goal was to show that protocol RCE can be better fore some situations in the EC2 scenario; we have tried to do a more complete analysis. Unfortunately, the derived bound is not useful because of its complexity.

In a second part, we have studied a conjecture, which comes from [8], too. We first wrote a simulation, seeking a counterexample to this conjecture. It couldn't be found any. Then, we have tried to understand the theoretical foundations of this conjecture. We have shown a few possibilities to get its complete formal proof.

This semester thesis was my first "rendez-vous" with a research field. It was very interesting to try to underestand the way that researchers take to study the foundations of perfect secret-key agreement. I understood how difficult it can be to proove a lemma, how frustrating it can be to try to write down some ideas. I'm disappointed with the practical results of my attempts, which are close to nothing, but on the other side, I've learned a lot of things, and this not only in the field of cryptology. I think it's the more important !

References

[1] Charles H. Bennett, François Bessette, Gilles Brassard, Louis Salvail, and John Smolin. Experimental quantum cryptography. In I. B. Damgård, editor, *Advances in Cryptology—EUROCRYPT 90*, volume 473 of *Lecture Notes in Computer Science*, pages 253–265. Springer-Verlag, 1991, 21–24 May 1990.

[2] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory.* Wiley Series in Telecommunications. John Wiley & Sons, New York, NY, USA, 1991.

[3] P. L'Ecuyer. Efficient and portable combined random number generators. *Commun. of the ACM*, 31, 6:742–749, 1988.

[4] Ueli Maurer and Stefan Wolf. Unconditionally secure key agreement and the intrinsic conditional information. *IEEE Transactions on Information Theory*, 45(2):499–514, March 1999.

[5] Ueli M. Maurer. Secret key agreement by public discussion from common information. *IEEE Transactions on Information Theory*, 39(3):733–742, May 1993.

[6] Ueli M. Maurer. The strong secret key rate of discrete random triples. In Richard E. Blahut et al., editors, *Communications and Cryptography: Two Sides of One Tapestry.* Kluwer, 1994.

[7] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C.* Cambridge University Press, Cambridge, second edition, 1992.

[8] Stefan Wolf. *Information-theoretically and computationally secure key agreement in cryptography.* Ph.D. dissertation, ETH Zürich, 1999.