Information-theoretical Secret-key agreement and Bound information

Master Thesis / Diplomarbeit

by

Martin Kreißig

Institute of Photonic Sciences

Quantum Optics group

Advisor: Prof. Dr. Antonio Acin

Universitat Politecnica de Catalunya

Escola Tecnica Superior d'Enginyeria de Telecomunicacio de Barcelona

Co-Advisor: Prof. Josep Sole Pareta

Universität Stuttgart

Institut für Kommunikationsnetze und Rechnersysteme

Co-Advisor: Dipl.-Ing. Andreas Gutscher, Prof. Paul J. Kühn

Abstract

One big problem of the communication between two parties is the secrecy. That means how much information a third party can obtain by intercepting the messages transmitted from one honest party to the other one. Therefore cryptography offers a wide range of protocols to ensure security with assumptions on the eavesdropper. So one was looking for an information-theoretical description of the scenario to get unconditional secure communication. In this scenario we are considering two honest parties that want to communicate over an authenticated channel that the eavesdropper is wiretapping.
This scenario introduced the definition of the intrinsic information and the secret-key rate which are a measure of the secrecy in this setting. Later because of strong analogies to quantum mechanics it turned out that this description was lacking a phenomena called bound information which is the disability of a probability distribution to create a secret-key even though it has predicted secrecy.

Nearly ten years of research have shown the existence of bound information for the multipartite case where several parties are communicating but not yet for the bipartite case. Hence the approach of non-distillability seems a very promising one to find this conjecture. Motivated by this the approach we implemented this tool and simulated some distributions that have conjectured bound information. Thereby we improved the tool to reduce its calculation time and to get closer to the aim.

Contents

List of Figures

List of Tables

1 Secret key agreement

For the discussion of information-theoretical key agreement it is necessary to explain the basic concepts of information theory. In the following we explain the measure for the information of one source and the measure for the correlation between two sources. With this we can introduce the scenario and the discussion of secret key agreement in the subsequent sections of this chapter.

1.1 Introduction to information theory

The first description of information theory was introduced by Shannon in 1948 [2] who defined the *information I* of an *event x* as:

$$I(x) = \log \frac{1}{P(x)} \tag{1.1}$$

This measure associates the information of the event with its uncertainty, given by its probability. That means the less often this event occurs, the larger is its information. When we have a source of the discrete alphabet $\mathscr{X}$ with elements $\{x_1, x_2, ..., x_N\}$ we can compute the *average amount of information* given from this source by its expectation value. Thus Shannon introduced the *entropy H* like:

$$H(X) = E[I(X)] = \sum_{i=1}^{N} P(x_i) \cdot I(x_i) = \sum_{i=1}^{N} P(x_i) \cdot \log \frac{1}{P(x_i)} \tag{1.2}$$

with $0 \leq H(X) \leq \log N$ where the maximal value belongs to a uniform distribution.
Now, we can introduce the *conditional entropy* which is the entropy of a source X given the side information from source Y as:

$$\begin{aligned} H(X|Y) &= \sum_{x \in X} \sum_{y \in Y} P(x,y) \log \frac{1}{P(x|y)} \\ &= H(X,Y) - H(Y) \end{aligned} \tag{1.3}$$

where $P(x, y)$ is the joint distribution of X and Y and $P(x|y)$ the conditional distribution of X given Y.
A measure for the correlation between two random variables X and Y distributed according

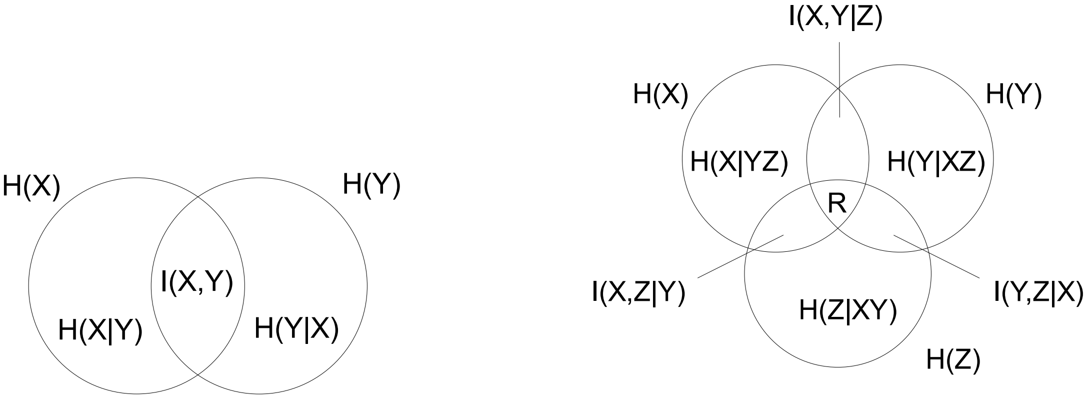
to the probability distribution $P(x,y)$ is given by the *mutual information* $I(X,Y)$.

$$\begin{aligned} I(X,Y) &= \sum_{x\in X}\sum_{y\in Y} P(x,y)\log\frac{P(x,y)}{P_1(y)P_2(y)} \\ &= H(X) - H(X|Y) \end{aligned}$$

When we introduce a third random variable in the scenario the *conditional mutual information* that gives us the amount of correlation between the two random variables X and Y given the information of the third one is given by:

$$\begin{aligned} I(X,Y|Z) &= H(X|Z) - H(X|Y,Z) && (1.4) \\ &= \sum_{z\in Z} P_Z(Z=z)\cdot I(X,Y|Z=z) && (1.5) \end{aligned}$$

To summarize these results we want to give a graphical representation of the relations between the (conditional) entropy and the (conditional) mutual information. Therefore have a look at figure 1.1 (a) where you can see a diagram for the bipartite case of the two alphabets $\mathscr{X}$ and $\mathscr{Y}$. When we expand this scenario to the tripartite case we can derive the relationship shown in figure 1.1 (b), where all arguments in R can be interchanged because of their symmetry.



(a) Bipartite case

(b) Tripartite case with $R(X,Y,Z) := I(X,Y) - I(X,Y|Z)$

Figure 1.1: Graphical representation of the entropy, the conditional and the mutual information

For further details we want to refer to appendix A at this point.

1.2 Motivation for secret-key agreement

It is an old problem that two honest parties, called in what follows Alice and Bob, want to communicate secure messages in a real environment. The known methods like Secure Shell

(SSH) or Pretty Good Privacy (PGP) are based on *computational security*, namely they are secure against an eavesdropper, Eve, who is assumed to have limited computational capabilities. The improvement of computational capabilities, and in particular the advent of quantum computation, sheds light on the medium- or long-term applicability of these methods. Thus, it is desirable to design alternative cryptographic protocols having a stronger form of security, *information-theoretic security*. Contrary to the previous schemes, the security of these protocols can be proven using concepts from Information Theory.

The first formulations in this field were done by Shannon [3] in 1949 who suggested to encrypt a *plaintext* message M with a *secret key* S to the *ciphertext* C. This system is said to be perfect secure if the ciphertext does not reveal any information about the plaintext or in information theoretical writing: $I(M, C) = 0$.
This was based on a proposal by Vernam [4] where the honest parties use a pre-established private key to encrypt and decrypt the binary message, by making a modulo 2 sum with this key. More precisely, the emitter, say Alice, sums her message with the key, which is only known by the honest parties, and sends the resulting string of bits to Bob via the insecure channel. At the receiver side, Bob performs a second modulo 2 sum and obtains the message. This scheme is known as the *one-time pad* because perfect security is possible only when the key is used once. Moreover the key has to be at least as long as the message or in a more precise way: $H(S) \geq H(M)$.

As already stated, we do need a *secret key* to encrypt and decrypt a message to communicate secretly between the honest parties Alice and Bob. Somehow, this is just a reformulation of the initial problem, since now the main question is how to distribute the initial secret key in a secure way. This defines the key-agreement problem.

1.3 The scenario

In the next lines, we introduce the standard key-agreement scenario considered throughout this work. The two honest parties and the eavesdropper receive correlated symbols, denoted by the random variables X, Y and Z, respectively, distributed according to the probability distribution P_{XYZ}. The goal of the honest parties is to map these symbols into a secret key, which will later be used for secure information transmission via one-time pad. In order to do that, Alice and Bob exchange information over an authenticated but insecure communication channel. The channel is authenticated but insecure because Eve can receive the whole communication, but cannot tamper it. A more general scenario is given by a completely insecure channel where Eve can also modify and introduce messages. However, here we will not consider this possibility as there are easy mechanisms to check the integrity of the messages.

Thus we can introduce a model like shown in figure 1.2.

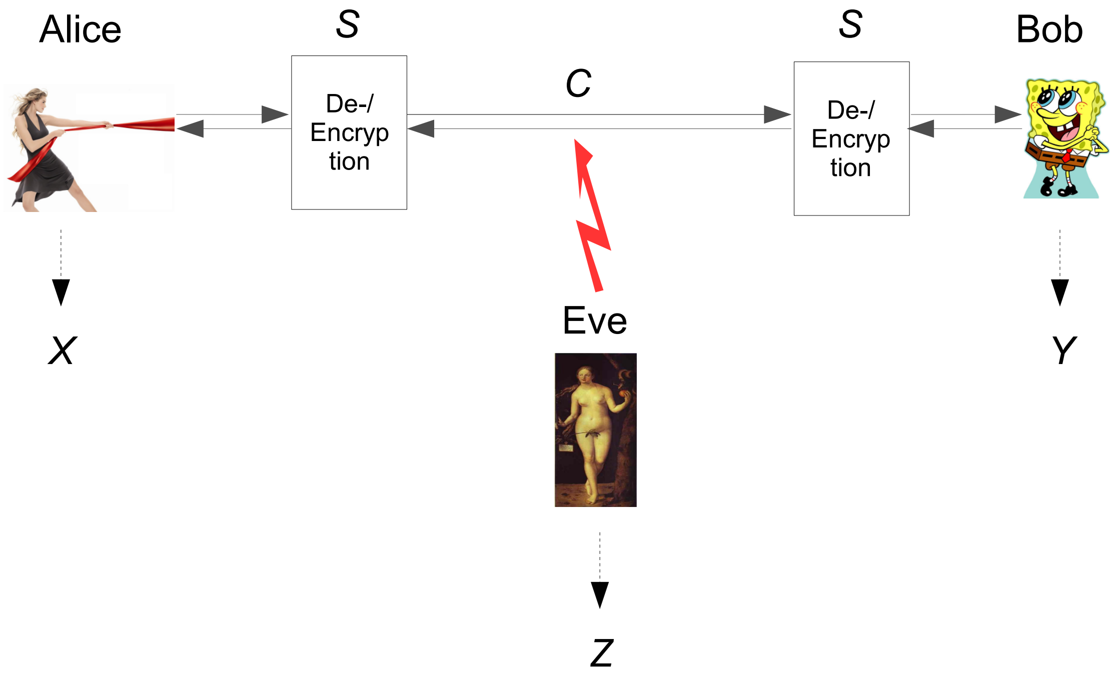


Figure 1.2: Communication model used in this thesis

An example of a probability distribution P_{XYZ} is given in table 1.1. Here Alice and Bob have a binary random variable X and Y, while Eve's variable belongs to an alphabet of size equal to three. Note that when Eve obtains the value 3, she knows for sure that Alice's random variable is equal to one and Bob's to zero. Hence for this specific outcome, which occurs with probability β, Alice and Bob have no secrecy. But if Eve gets $Z = 1, 2$, she does not have perfect knowledge on Alice and Bob's variables. This might open the possibility for secret correlations between Alice and Bob that could be distilled into a secret key.

X Y (Z)	0	1
0	(1) α	(3) β
1	(2) γ	(1) δ (2) ε

Table 1.1: Arbitrary distribution P_{XYZ}. Condition $\alpha + \beta + \gamma + \delta + \varepsilon = 1$.

The one-time pad allows us to reduce the problem of information-theoretical secure communication to the information-theoretical secure key agreement. So the question is if the two honest parties are able to use independent realizations of a given distribution to obtain a secret key.

1.4 Unconditional secret key agreement

As mentioned above, the main goal for Alice and Bob is to transform the initial probability distribution P_{XYZ}, into a secret key, namely a list of perfectly correlated symbols about which Eve has no information. In order to do that, they can apply local operations to their symbols and exchange messages over the insecure channel. These protocols define the set of local operations assisted by public communication, briefly denoted by LOPC. To get unconditional secrecy we have to show in information-theoretical terms that there is no or at least asymptotically no correlation between the adversary and the secret key. Hence we aim at finding a LOPC protocol transforming N realizations of the initial distribution P_{XYZ} into a new distribution arbitrarily close to M secret bits, defined as the tripartite probability distribution $S_{XYZ} = S_{XY}P_Z$, that is $P^{(N)}_{XYZ} \xrightarrow[LOPC]{} S^{(M)}_{XYZ}$, where S is given in table 1.2.

X Y	0	1
0	1/2	0
1	0	1/2

Table 1.2: Distribution S between Alice and Bob for a secret bit

The secret key rate of the initial probability distribution corresponds to the rate for the optimal protocol, that is, the maximum of M/N over all LOPC protocols. This was formulated in a more rigorous way in [5] and [6].

Definition 1 *The* secret key rate *of X and Y with respect to Z, denoted by $S(X, Y||Z)$, is the maximum rate at which Alice and Bob can agree on a secret key S in such a way that the amount of information that Eve obtains about S is arbitrarily small. In other words, it is the maximal R such that for every $\epsilon > 0$ and for all sufficiently large N there exists a protocol, using public discussion over an insecure but authenticated channel, such that Alice and Bob who receive $X^N = [X_1, \cdots, X_N]$ and $Y^N = [Y_1, \cdots, Y_N]$, respectively, compute the same key S with probability at least $1 - \epsilon$ satisfying*

$$\begin{aligned} I(S, CZ^N) &\leq \epsilon & (1.6)\\ H(S) &\geq \log|S| - \epsilon & (1.7)\\ \frac{1}{N}H(S) &\geq R - \epsilon & (1.8)\end{aligned}$$

where C denotes the communication, i.e. the collection of all messages M, sent over the channel and $|S|$ denotes the alphabet of S.[1]

[1] All logarithms throughout this thesis are to the basis 2

That means that: (1.6) the correlation (mutual information) between N copies of Eve's random variable, the communication C and the key S is arbitrary small, (1.7) the entropy of the secret key is arbitrary close to its maximum, namely an uniform distribution and (1.8) the rate is non-zero in the limit of large blocks.

Based on this definition it is hard to find the secret-key rate for a given distribution. Hence it is very useful to establish more easily computable upper and lower bounds for this quantity. It seems quite reasonable that the rate at which Alice and Bob agree on a secret bit cannot be lower than their shared information degraded by the mutual information between Eve and one of them. In other words the secret-key rate must be larger than what Eve knows about one random variable of the honest parties.

$$\max\{I(X,Y) - I(X,Z), I(Y,X) - I(Y,Z)\} \leq S(X,Y||Z) \tag{1.9}$$

This bound is reachable using LOPC protocols where all the communication goes in one direction, say from Alice to Bob [7] if $\max\{I(X,Y) - I(X,Z), I(Y,X) - I(Y,Z)\} = I(X,Y) - I(X,Z)$. Indeed, Alice and Bob can first use error correction to eliminate their errors and agree on a perfectly correlated list of symbols and then apply *privacy amplification* (see also [8]) to the new list to obtain an unconditionally secure key. Privacy amplification is a map of a K-bit string to L bits, where $K > L$, by universal hash functions. It is known however that two-way communication protocols are more powerful than one-way, since it was shown in [5] - and we will prove it in chapter 1.5 - that it is even possible to create a positive secret-key rate in situations where $I(X,Z) > I(X,Y)$ and $I(Y,Z) > I(X,Y)$.

Furthermore it seems quite intuitive that the secret-key rate cannot exceed the mutual information between Alice and Bob because that is the total amount of information they share. It cannot be larger either than the mutual information between the honest parties conditioned on the eavesdropper. Thus, one has:

$$S(X,Y||Z) \leq \min\{I(X,Y), I(X,Y|Z)\} \tag{1.10}$$

But what happens if Eve performs any kind of local operation on her random variable? Then she is able to change the conditional mutual information and hence we get a tighter bound on the secret-key rate. This operation can be described through a channel characterized by the conditional probability $P_{\overline{Z}|Z}$ with Z being the input and $\overline{Z}$ being the output random variable.

Definition 2 *Given a distribution P_{XYZ} the* intrinsic (conditional mutual) information *is defined as*

$$I(X,Y\downarrow Z) := \inf_{P_{\overline{Z}|Z}} \left\{ I(X,Y|\overline{Z}) : P_{XY\overline{Z}} = \sum_{z\epsilon Z} P_{XYZ}\cdot P_{\overline{Z}|Z} \right\} \tag{1.11}$$

This leads to a stronger upper bound on the secret-key rate

$$S(X,Y||Z) \leq I(X,Y \downarrow Z) \leq I(X,Y|Z) \tag{1.12}$$

Another quantity to classify the correlations of Alice and Bob was introduced in [9] that is the rate at which Alice and Bob can generate a distribution by public communication that is at least as good as P_{XYZ}.

Definition 3 *Let P_{XYZ} be the joint distribution of three discrete random variables X, Y and Z. The* information of formation of X and Y given Z, *denoted by $I_{form}(X,Y|Z)$, is the infimum of all numbers $R \geq 0$ with the property that for all $\varepsilon > 0$ there exists N_0 such that for all $N \geq N_0$, there exists a protocol between Alice and Bob with communication C and achieving the following: Alice and Bob, both knowing the same random $\lfloor RN \rfloor$-bit string S, can finally compute X' and Y', respectively, sucht that there exist random variables X^N, Y^N and Z^N jointly distributed according to $(P_{XYZ})^N$ (this is the distribution corresponding to n-fold independent repetition of the random experiment P_{XYZ}) and a channel $P_{\overline{C}|Z^N}$ sucht that*

$$Prob\left[(X',Y',C) = (X^N,Y^N,\overline{C})\right] \geq 1-\varepsilon \tag{1.13}$$

holds.

This shows that the synthesis of our distribution is somehow only depending on P^N_{XY} because the communication C can be simulated by an Eve knowing the corresponding Z^N. From this we can also formalize the fact, that Eve does not gain any information by observing C.
It was also proven that the information of formation is lower bounded from the intrinsic informtion. This means that the intrinsic information bounds the minimum number of secret bits required to create the desired distribution. Hence we have for every distribution P_{XYZ}

$$S(X,Y||Z) \leq I(X,Y \downarrow Z) \leq I_{form}(X,Y|Z) \tag{1.14}$$

We want to remark here that a distribution can be established by LOPC if and only if $I_{form}(X,Y|Z) = 0$ [10]. If $I_{form}(X,Y|Z) > 0$, the distribution requires the use of secret correlations for its generation.

Before concluding this section, we would like to discuss other possible bounds on the secret-key rate. One may for instance consider how the secret-key rate is affected when Eve gets some additional side information U from an oracle. This can be formulated as $Z' = [Z,U]$ and would only affect equation (1.6) in the way that $I(S,CZ'^N) \leq \epsilon$. But this formulation is already included in $I(S,CZ^N) \leq \epsilon$ and hence we can follow:

$$S(X,Y||[Z,U]) \leq S(X,Y||Z) \tag{1.15}$$

Another interesting question is what happens if Alice or Bob perform local maps $P_{\overline{X}|X}$ and $P_{\overline{Y}|Y}$. This can be described as the following: Let X, Y, Z, $\overline{X}$ and $\overline{Y}$ be random variables jointly distributed like $P_{XYZ\overline{XY}} = P_{XYZ} \cdot P_{\overline{X}|X} \cdot P_{\overline{Y}|Y}$. Then we can state due to the fact that the secret-key rate is the maximum rate taken over all possible protocols between Alice and Bob that

$$S(X, Y||Z) \geq S(\overline{X}, \overline{Y}||Z) \tag{1.16}$$

This shows us that Alice and Bob cannot increase their secrecy by applying any kind of local operation which leads us to another interesting quantity the *binarization* of the alphabet of the honest parties which is the reduction of one's alphabet $\mathscr{X}$ or $\mathscr{Y}$, respectively, to a binary one $\mathscr{B}_A$ or $\mathscr{B}_B$. It is shown in [6] and [1] that the restriction of the ranges: $\mathscr{X} \rightarrow \tilde{\mathscr{X}}$ with $\tilde{\mathscr{X}} \leq \mathscr{X}$ and $\mathscr{Y} \rightarrow \tilde{\mathscr{Y}}$ with $\tilde{\mathscr{Y}} \leq \mathscr{Y}$ does not increase the secret-key rate.

Lemma 4 *Let X, Y and Z be random variables with ranges $\mathscr{X}$, $\mathscr{Y}$ and $\mathscr{Z}$ and joint distribution P_{XYZ}. For $\tilde{\mathscr{X}} \subset \mathscr{X}$ and $\tilde{\mathscr{Y}} \subset \mathscr{Y}$, we define a new random experiment with random variables $\tilde{X}$ and $\tilde{Y}$ (with ranges $\tilde{\mathscr{X}}$ and $\tilde{\mathscr{Y}}$, respectively). If Ω is the event that $X \in \tilde{\mathscr{X}}$ and $Y \in \tilde{\mathscr{Y}}$, then the joint distribution of $\tilde{X}$ and $\tilde{Y}$ with Z is defined as follows:*

$$P_{\tilde{X}\tilde{Y}Z}(x, y, z) := \frac{P_{XYZ}(x, y, z)}{P_{XYZ}[\Omega]} \tag{1.17}$$

for all $(x, y, z) \in \tilde{\mathscr{X}} \times \tilde{\mathscr{Y}} \times Z$. Then

$$S(X, Y||Z) \geq P_{XYZ}[\Omega] \cdot S(\tilde{X}, \tilde{Y}||Z) \tag{1.18}$$

This follows from the definition of the secret-key rate which is already the maximal rate for key generation.

In the next section, we introduce the most commonly used key distillation protocol and then discuss how it can be used for secret key distillation in a relevant scenario.

1.5 Protocol: Advantage distillation

Advantage distillation is an LOPC protocol for key agreement that uses two-way communication. It may allow distilling a key even in situations when standard one-way communication techniques fail [5,6]. Although initially presented in the binary case, the protocol works for variables of arbitrary size. It works as follows: Alice locally generates a random variable C of the same size d as X. Then, she take N realizations of X and computes the N values M_i satisfying

$$C = M_i + X_i,$$

where the sum is modulo d. The N variables M_i are then transmitted to Bob over the public channel. Bob receives the bit-string M_i and performs the same sum with his corresponding set of variables Y_i:

$$Y_i + M_i.$$

Bob will accept the codeword only if all these sums give the same result, D, which he keeps as his new symbol.

We will now give an example showing how this protocol can distill a key from a probability distribution where Eve's information on Alice and Bob's variables is larger than the correlations between the honest parties. We will see how it enables mapping the initial probability distribution into a new probability distribution where equation (1.9) is positive. Thus, the honest parties can apply error correction and privacy amplification to the new distribution, obtained after advantage distillation, and obtain an unconditional secure key.

Consider the situation in which the joint distribution is coming from a broadcasted signal with random variable R and $P_R(0) = P_R(1) = 1/2$ as shown in the figure 1.3. This signal arrives at Alice, Bob and Eve with different error probabilities $P_{X|R}(1,0) = P_{X|R}(0,1) = \epsilon_A/2$, $P_{Y|R}(1,0) = P_{Y|R}(0,1) = \epsilon_B/2$ and $P_{Z|R}(1,0) = P_{Z|R}(0,1) = \epsilon_E/2$ Moreover $\delta_A = 1 - \epsilon_A$, $\delta_B = 1 - \epsilon_B$ and $\delta_E = 1 - \epsilon_E$ being the probabilities of a correct transmission.

Without loss of generality we can assume that $\epsilon_A = \epsilon_B = \epsilon$ and hence $\delta_A = \delta_B = \delta$, i.e. Alice's and Bob's channels are identical[2]. Hence all three parties have a different knowledge about the transmitted bits and this is reflected by the probability distribution 1.3. In this scenario, one can see that no one-way communication protocol enables secret-key distillation if Eve's error is smaller than Alice and Bob's.

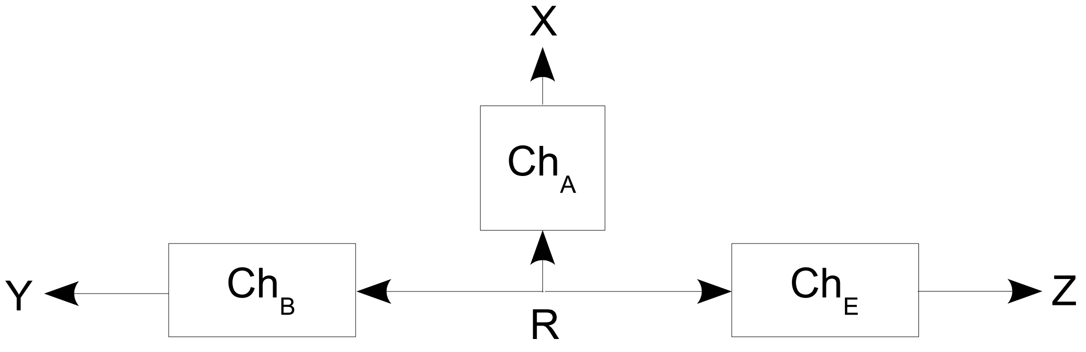


Figure 1.3: Model of the cascaded channels used in the broadcasting scenario

Let's analyze how the initial distribution changes after application of the advantage distil-

[2] If e.g. $\epsilon_A < \epsilon_B$ we can cascade another channel with error probability $(\epsilon_B - \epsilon_A)/(1 - 2\epsilon_A)$ to obtain $\epsilon_A = \epsilon_B$

lation protocol previously described. After this protocol, the probability that Bob accepts a message correctly is given by the *fidelity* $\mathcal{F}$

$$\mathcal{F}^N = (\delta^2 + \epsilon^2)^N \tag{1.19}$$

and in the case of a false accepted message we obtain the *disturbance* $\mathcal{D}$

$$\mathcal{D}^N = \Big(1 - (\epsilon^2 + \delta^2)\Big)^N \tag{1.20}$$

Moreover we can say that Bob accepts a message in general with the probability

$$p_{accept} = \mathcal{F}^N + \mathcal{D}^N \tag{1.21}$$

and thus we can derive Bob's overall error probability, getting:

$$\beta_N = \frac{\mathcal{D}^N}{\mathcal{F}^N + \mathcal{D}^N} \tag{1.22}$$

We are now interested in the conditional probability γ_N that Eve decides the wrong message

X Y (Z)	0	1
0	(0) $\delta_{\mathcal{F}} \cdot \frac{\mathcal{F}}{2}$ (1) $(1-\delta_{\mathcal{F}}) \cdot \frac{\mathcal{F}}{2}$	(0) $\delta_{\mathcal{D}} \cdot \frac{\mathcal{D}}{2}$ (1) $(1-\delta_{\mathcal{D}}) \cdot \frac{\mathcal{D}}{2}$
1	(0) $(1-\delta_{\mathcal{D}}) \cdot \frac{\mathcal{D}}{2}$ (1) $\delta_{\mathcal{D}} \cdot \frac{\mathcal{D}}{2}$	(0) $(1-\delta_{\mathcal{F}}) \cdot \frac{\mathcal{F}}{2}$ (1) $\delta_{\mathcal{F}} \cdot \frac{\mathcal{F}}{2}$

Table 1.3: Resulting distribution after receiving the broadcasted signal

under the condition that Bob accepts the correct one. Therefore we introduce the probability $\delta_{\mathcal{F}} = P(Z = X|X = Y)$ that Eve makes a right decision given that Bob accepts the correct one.

$$\delta_{\mathcal{F}} = \delta^2\delta_E + \epsilon^2\epsilon_E \tag{1.23}$$

Thus we can derive the probability when Eve decides a wrong bit given that Bob accepts the correct one $P(Z \neq X|X = Y)$ as

$$1 - \delta_{\mathcal{F}} = \delta^2\epsilon_E + \epsilon^2\delta_E \tag{1.24}$$

The other case can be described as, given that Bob accepts a wrong bit Eve can either decide for the correct one ($\delta_{\mathcal{D}} = P(Z = X|X \neq Y) = \delta\epsilon\epsilon_E + \epsilon\delta\delta_E$) or the false one ($1 - \delta_{\mathcal{D}} = \delta\epsilon\delta_E + \epsilon\delta\epsilon_E$). This leads us directly to table 1.3 which illustrates the probabilities of a correct decision on Bob's side (fidelity and disturbance) as well as the conditional probabilities of Eve based on the outcome of Bob for each bit broadcasted.

We can now conclude Eve's error probability for the whole message as the following:

$$\begin{aligned}\gamma_N &= \frac{1}{2}\cdot\frac{1}{p_{accept}}\cdot\sum_{i=N/2}^{N}\binom{N}{i}\left(\delta^i_{\mathcal{F}}(1-\delta_{\mathcal{F}})^{N-i}+\delta^i_{\mathcal{D}}(1-\delta_{\mathcal{D}})^{N-i}\right) &(1.25)\\ &\geq \frac{1}{2}\cdot\frac{1}{p_{accept}}\cdot\binom{N}{N/2}\left(\delta^{N/2}_{\mathcal{F}}(1-\delta_{\mathcal{F}})^{N/2}+\delta^{N/2}_{\mathcal{D}}(1-\delta_{\mathcal{D}})^{N/2}\right) &(1.26)\end{aligned}$$

As one can easily see $\delta^{N/2}_{\mathcal{D}}(1-\delta_{\mathcal{D}})^{N/2} = (\delta\epsilon)^N$ which is a very small value that we can also neglect. Moreover when using the Stirling formula (see [6]): $\binom{N}{N/2} \geq \frac{1}{\sqrt{2\pi N}}\cdot 2^N$, we can rewrite (1.26) as:

$$\gamma_N \geq \frac{1}{2\sqrt{2\pi N}}\cdot\frac{1}{p_{accept}}\cdot\left(2\sqrt{\delta_{\mathcal{F}}(1-\delta_{\mathcal{F}})}\right)^N \tag{1.27}$$

For $\epsilon < 1/2$ and with equalities (1.23) and (1.24) we can show that:

$$\begin{aligned}\sqrt{\delta_{\mathcal{F}}(1-\delta_{\mathcal{F}})} &= \sqrt{(1-2\epsilon+\epsilon^2-\epsilon_E+2\epsilon\epsilon_E)(\epsilon^2-2\epsilon\epsilon_E+\epsilon_E)}\\ &\geq \epsilon(1-\epsilon) &(1.28)\end{aligned}$$

which is equal for $\epsilon_E = 0$ and maximal for 1/2. From equation (1.20) we can conclude $\mathcal{D} = 2\cdot(\epsilon-\epsilon^2)$. This leads us to the result:

$$\begin{aligned}\gamma_N &> \frac{1}{2\sqrt{2\pi N}}\cdot\frac{1}{p_{accept}}\cdot\mathcal{D}^N &(1.29)\\ &= \frac{1}{2\sqrt{2\pi N}}\cdot\beta_N &(1.30)\end{aligned}$$

This shows that Bob's error decreases exponentially with the number of copies compared to Eve's.
We proof later that secret key agreement is only possible if

$$\frac{\mathcal{D}}{1-\mathcal{D}} < 2\sqrt{\delta_{\mathcal{F}}(1-\delta_{\mathcal{F}})} \tag{1.31}$$

Now we have to show that with $\beta_N < \gamma_N$ the secret-key rate is positive. Therefore consider the scenario that we want to construct the random variables $\tilde{X}$ and $\tilde{Y}$ from our X^N and Y^N random variables exchanged over the authenticated channel. Then it suffices to show that

$$I(\tilde{X},\tilde{Y}) - I(\tilde{X},\tilde{Z}) = H(\tilde{X}|\tilde{Z}) - H(\tilde{X}|\tilde{Y}) > 0 \tag{1.32}$$

where $\tilde{Z} = [Z^N, V]$ with V being the collection of all messages over the public channel. Then we define $\tilde{X}$ and $\tilde{Y}$ as follows: If Bob accepts $\tilde{X} = C$ and $\tilde{Y} = C'$ and if he rejects publicly

$\tilde{X} = \tilde{Y} =$ "*reject*". Given that Bob accepts Alice's bit and given that $\beta_N = b^N$ we can state that:

$$H(C|C') = h(\beta_N) \leq 2b^N \cdot \log(1/b^N) = 2b^N \cdot N \cdot \log(1/b) < \gamma_N \tag{1.33}$$

for sufficiently large N and $h(p) = -p \log p - (1-p)\log(1-p)$ being the binary entropy. The first inequality in (1.33) is derived from $-p \log p \geq -(1-p)\log(1-p)$ for $p \leq 1/2$ and the second one from the Jensen's inequality .
Eve's side can be described as

$$H(C|\tilde{Z}) = \sum_{\tilde{z}} P_{\tilde{Z}}(\tilde{z}) H(C|\tilde{Z} = \tilde{z}) = E[h(q(\tilde{Z}))] \geq E[q(\tilde{Z})] = \gamma_N \tag{1.34}$$

where $q(\tilde{z})$ denotes the probability of Eve guessing C incorrectly with her optimal strategy, i.e. $q(\tilde{z}) \leq 1/2$ and $h(q(\tilde{z})) > q(\tilde{z})$.
For Bob's public rejection we have

$$H(\tilde{X}|\tilde{Y}) = H(\tilde{X}|\tilde{Z}) = H(\tilde{X}|V) = 0 \tag{1.35}$$

Concluding this we get $I(\tilde{X}, \tilde{Y}) - I(\tilde{X}, \tilde{Z}) > 0$ which proofs the information-theoretical security of this protocol.

Now we proof that condition (1.31) leads to secret key agreement by showing that the contrary condition that means

$$\frac{\mathcal{D}}{1-\mathcal{D}} \geq 2\sqrt{\delta_{\mathcal{F}}(1-\delta_{\mathcal{F}})} \tag{1.36}$$

leads to non-distillability. Therefore we consider our distribution of table 1.3 where Alice and Bob are independent and hence $I(X, Y|Z) = 0$ which is a sufficient condition for non-distillability if

$$P_{XYZ}(0,0|z) \cdot P_{XYZ}(1,1|z) = P_{XYZ}(0,1|z) \cdot P_{XYZ}(1,0|z) \tag{1.37}$$

For $z = 0$ (and $z = 1$ respectively) we obtain from the previous condition that

$$\delta_{\mathcal{F}}(1-\delta_{\mathcal{F}}) \cdot \frac{\mathcal{F}^2}{4} = \delta_{\mathcal{D}}(1-\delta_{\mathcal{D}}) \cdot \frac{\mathcal{D}^2}{4}$$

By the definition of $\delta_{\mathcal{D}}$ we see that $\delta_{\mathcal{D}} = 1 - \delta_{\mathcal{D}}$ and this concludes the equality of (1.36).
It remains to show the inequality of (1.36). Herefore we assume that Eve performs the maps $P_{\tilde{Z}|Z}(0,0) = p$, $P_{\tilde{Z}|Z}(1,0) = 1-p$, $P_{\tilde{Z}|Z}(1,1) = q$ and $P_{\tilde{Z}|Z}(0,1) = 1-q$. It is easy to check that this leads to the conditions $\frac{\mathcal{D}}{1-\mathcal{D}} = \sqrt{\delta_{\mathcal{F}}^2 + (1-\delta_{\mathcal{F}})^2}$ which is included in condition 1.36 and this concludes the proof.

The protocol explained here does not claim to be the most efficient one. It shall only illustrate the mechanisms.

2 Bound information

2.1 Motivation

Given a probability distribution, it is not easy to check whether it is distillable since the computation of the secret-key rate requires a maximization over all possible LOPC protocols. Clearly, the positivity of the intrinsic information is a necessary condition for a probability distribution to be distillable (see equation (1.12)). However, it is an open problem whether this condition turns out to be sufficient. If this was the case, all probability distributions that could not be created by LOPC would be distillable. On the other hand, the existence of non-distillable probability distributions with positive intrinsic information would imply the existence of an irreversible form of secret correlations, as (i) some sort of secrecy is needed for the preparation of the distribution but (ii) this secrecy cannot be distilled into a secret key. This irreversible form of secret correlations is known as *bound information*.

A similar phenomenon was observed in quantum physics, concerning the problem of distilling pure-state entanglement from an entangled quantum state. There, it was shown that the distillation of an entangled quantum state into a maximally entangled state is not always possible. This irreversible form of entanglement is known as *bound entanglement*[1].

The strong analogies between entanglement distillation and secret-key agreement motivated the conjecture that there may also exist probability distributions containing secret correlations, in the sense that its formation by LOPC is impossible, that cannot be used for secret key agreement. These distributions would be the classical cryptographic analog of bound entangled states. The relation between bound entanglement and bound information has been discussed in previous articles and is beyond the scope of this thesis. We refer to references [1] and [11] for details.
Before proceeding, let us precisely define bound information.

Definition 5 *A distribution contains bound information when (i) its formation by LOPC is impossible and (ii) no secret key can be distilled out of it by LOPC.*

[1] Entanglement is the term usually employed to name quantum correlations. We would like to refer to appendix B for an introduction to quantum physics

This definition is general and applies to scenarios with more than two honest parties, as it will be discussed below. However, in this thesis we mainly consider the standard case of two honest parties plus the eavesdropper. Then, the existence of bound information is equivalent to finding a probability distribution such that:

$$S(X,Y\|Z) = 0 \qquad\qquad I_{form}(X,Y|Z) > 0 \tag{2.1}$$

The first condition means that the distribution is useless for key distribution, while the second one implies that the formation of the distribution by LOPC is impossible. This last condition can also be replaced by $I(X,Y\downarrow Z) > 0$, since $I(X,Y\downarrow Z) > 0$ if and only if $I_{form}(X,Y|Z) > 0$.

2.2 Example of conjectured bound information

We give in what follows an example of a probability distribution which is conjectured to have bound information. The distribution was presented in [1] and was constructed from a bound entangled quantum state. The distribution reads (to be normalized):

X Y (Z)	1	2	3
1	(0) 2	(4) 5-α	(3) α
2	(1) α	(0) 2	(5) 5-α
3	(6) 5-α	(2) α	(0) 2

Table 2.1: Distribution P_{XYZ} coming from example 3 in [1]

In the following we will show that distribution P_{XYZ} is distillable, non-distillable and has conjectured bound information for different ranges of α [1].

Initially we calculate the conditional mutual information of this distribution, having

$$\begin{aligned}
I(X,Y|Z) &= P_Z(0)\cdot I(X,Y|Z=0) + \sum_{z=1}^{6} P_Z(z)\cdot I(X,Y|Z=z) \\
&= \frac{\sum_{x,y} P_{XYZ}(x,y,0)}{\sum_{x,y,z} P_{XYZ}(x,y,z)} \cdot \sum_{i=1}^{3} \frac{P_{XYZ}(i,i,0)}{\sum_{x,y} P_{XYZ}(x,y,0)} \log \frac{\sum_{xy} P_{XYZ}(x,y,0)}{P_{XYZ}(i,i,0)} + 0 \\
&= \frac{6}{\sum_{x,y,z} P_{XYZ}(x,y,z)} \log 3 \qquad (2.2) \\
&> 0 \qquad (2.3)
\end{aligned}$$

Our first goal is to see when the distribution has positive intrinsic information. Therefore we

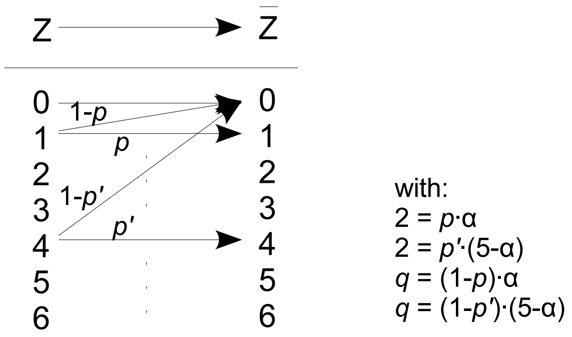


Figure 2.1: Map on Eve's side to reduce the secret key rate

consider the maps by Eve of figure 2.1: she maps all values of her alphabet except zero onto themselves with probability p or p', depending on their probability of occurrence, given in P_{XYZ}, and onto zero with probability $(1-p)$ and $(1-p')$. The conditions in figure 2.1 result from the fact that all $P_{XYZ}(x, y, z = 0) = 2$ in table 2.1.
Thus we end in a distribution that we can split in $P_{XY\overline{Z}1}$ and $P_{XY\overline{Z}2}$ like shown in table 2.2:

$P_{XYZ} =$

X Y ($\overline{Z}$)	1	2	3
1	(0) 2	(0) 2	(0) 2
2	(0) 2	(0) 2	(0) 2
3	(0) 2	(0) 2	(0) 2

+

X Y ($\overline{Z}$)	1	2	3
1	(0) 0	(4) q'	(3) q
2	(1) q	(0) 0	(5) q'
3	(6) q'	(2) q	(0) 0

Table 2.2: $P_{XY\overline{Z}1}$ and $P_{XY\overline{Z}2}$ after the map of Eve

One can see that the intrinsic information of these two distributions $P_{XY\overline{Z}1}$ and $P_{XY\overline{Z}2}$, and thus also for our distribution P_{XYZ}, is zero. So we can state:

$$S(X, Y||Z) \leq I(X, Y \downarrow Z) = I(X, Y|\overline{Z}) = 0 \tag{2.4}$$

This is however only possible for $2 \leq \alpha \leq 3$ because only then all values $P_{XYZ}(x, y, z) \geq 2$ and hence can be mapped to $P_{XYZ}(x, y, z = 0) = 2$. What happens for $0 < \alpha < 2$?

In this case, we consider a generic map by Eve $Z \rightarrow \overline{Z}$, as shown in 2.3. Here she maps every value of her alphabet onto arbitrary symbols labeled by i.

Applying these maps leads us to the distribution $P_{XY\overline{Z}}$ of table 2.4 which shows us the slice of one specific i.

channel probability	restriction
$P_{\overline{Z}\vert Z}(i,0) = a_i$	$\sum_i a_i = 1$
$P_{\overline{Z}\vert Z}(i,1) = b_i$	$\sum_i b_i = 1$
$P_{\overline{Z}\vert Z}(i,2) = c_i$	$\sum_i c_i = 1$
$P_{\overline{Z}\vert Z}(i,3) = \alpha_i$	$\sum_i \alpha_i = 1$
$P_{\overline{Z}\vert Z}(i,4) = \beta_i$	$\sum_i \beta_i = 1$
$P_{\overline{Z}\vert Z}(i,5) = \gamma_i$	$\sum_i \gamma_i = 1$
$P_{\overline{Z}\vert Z}(i,6) = \delta_i$	$\sum_i \delta_i = 1$

Table 2.3: Generalized maps for the intrinsic information with $0 < i < 6$

$\overline{Z} = i$ X Y	1	2	3
1	$2 \cdot a_i$	$(5-\alpha) \cdot \beta_i$	$\alpha \cdot \alpha_i$
2	$\alpha \cdot b_i$	$2 \cdot a_i$	$(5-\alpha) \cdot \gamma_i$
3	$(5-\alpha) \cdot \delta_i$	$\alpha \cdot c_i$	$2 \cdot a_i$

Table 2.4: Distribution $P_{XY\overline{Z}}$ obtained from a general map of P_{XYZ}

The question we will follow now is if distribution $P_{XY\overline{Z}}$ can be transformed in an independent one for the range $0 < \alpha < 2$. Taking into account the condition of independency of two random variables and by linear combination and substituting α, β, γ and δ we end up in a distribution only depending on a_i, b_i and c_i presented in table 2.5.

X Y $(\overline{Z})$	1	2	3
1	$2a_i$	$(5-\alpha)\frac{(2a_i)^2}{\alpha(5-\alpha)b_i}$	$\alpha\frac{(2a_i)^3}{\alpha^3 b_i c_i}$
2	αb_i	$2a_i$	$(5-\alpha)\frac{(2a_i)^2}{\alpha(5-\alpha)c_i}$
3	$(5-\alpha)\frac{\alpha^2 b_i c_i}{(5-\alpha)2a_i}$	αc_i	$2a_i$

Table 2.5: Distribution only depending on a_i, b_i and c_i under the condition that X and Y are independent (to be normalized)

Now we will derive a contradiction in this distribution from which we can follow that it is not possible to make the random variables X and Y independent and concluding $I(X, Y \downarrow Z) > 0$ for $\alpha < 2$.

Therefore we compare the entries $X = 3$ and $Y = 1$ in tables 2.1 and 2.5 and conclude the following inequality:

$$1 \geq \sum_i \left(\frac{2}{\alpha}\right)^3 \frac{a_i^3}{b_i c_i} \tag{2.5}$$

Without loss of generality we can state the $\sum_i \frac{a_i^3}{b_i c_i} = 1$ if $a_i \equiv b_i \equiv c_i$. Then inequality (2.5) is satisfied for $\alpha \geq 2$ which equivalent to the case prooven above. For $a_i \not\equiv b_i \not\equiv c_i$ we can show that $\sum_i \frac{a_i^3}{b_i c_i} > 1$ following an analogy given in [1] and then the lower bound for α is even bigger. That means Eve cannot map our distribution to a uniform one for $\alpha < 2$ and concluding $I(X, Y \downarrow Z) > 0$.

Now we introduce a possible protocol which shows that the secret key rate is positive for $0 \leq \alpha < 1$. Therefore Alice and Bob binarize their sets and only accept the values $X, Y \in \{1, 3\}$ and obtain the following distribution:

X Y (Z)	1	3
1	(0) 2	(3) α
3	(6) 5-α	(0) 2

Table 2.6: Binarization of the distribution P_{XYZ} from 2.1

As we know it is sufficient for the advantage distillation protocol to show that Alice and Bobs' error probability is lower than 1/2. In this case by the usage of many copies ($N \gg 0$) the error probability tends to zero and the honest parties end up in a probability distribution having only non-zero values at $X = Y$. To show this we consider the disturbance $\mathcal{D} = P_{XYZ}(1, 3, z) + P_{XYZ}(3, 1, z)$:

$$\beta_N = \frac{\mathcal{D}^N}{\mathcal{D}^N + \mathcal{F}^N} \xrightarrow{N \gg 0} \begin{cases} 0 \text{ for } \mathcal{D} < 1/2 \\ 1 \text{ for } \mathcal{D} > 1/2 \end{cases} \tag{2.6}$$

If we also manage to distribute $P(X = Y|Z)$ uniformly, the secrecy becomes obvious because that means whenever Eve obtains one value she is not able to distinguish whether Alice and Bob both have a one or a three.

This is achieved if Alice and Bob perform the local maps $P_{\overline{X}|X}(1, 0) = P_{\overline{X}|Y}(0, 1) = \frac{5-2\alpha}{14-2\alpha}$ and leave it otherwise. This leads to the distribution of table 2.7.

One can see the uniform distribution between $(X, Y) = (1, 1)$ and $(3, 3)$ and with $\alpha < 1$ we can show that $P(X = Y) > 1/2$. Then applying the advantage distillation protocol leads to a positive secret-key rate.

X Y (Z)	1	3
1	(0) $2\cdot\frac{9}{14-2\alpha}$ (6) $(5-\alpha)\cdot\frac{9}{14-2\alpha}\cdot\frac{5-2\alpha}{14-2\alpha}$	(0) $2\cdot 2\cdot\frac{5-2\alpha}{14-2\alpha}$ (6) $(5-\alpha)\cdot\left(\frac{5-2\alpha}{14-2\alpha}\right)^2$ (3) α
3	(6) $(5-\alpha)\cdot\left(\frac{9}{14-2\alpha}\right)^2$	(0) $2\cdot\frac{9}{14-2\alpha}$ (6) $(5-\alpha)\cdot\frac{9}{14-2\alpha}\cdot\frac{5-2\alpha}{14-2\alpha}$

Table 2.7: Maps of Alice and Bob to exclude Eve

For $\alpha > 1$ this protocol cannot be applied and there has not been found any other that leads to positive secret-key rate in this range.

The results obtained for this example are summarized and illustrated in figure 2.2 where one can also see the strong relation from the classical and the quantum world.

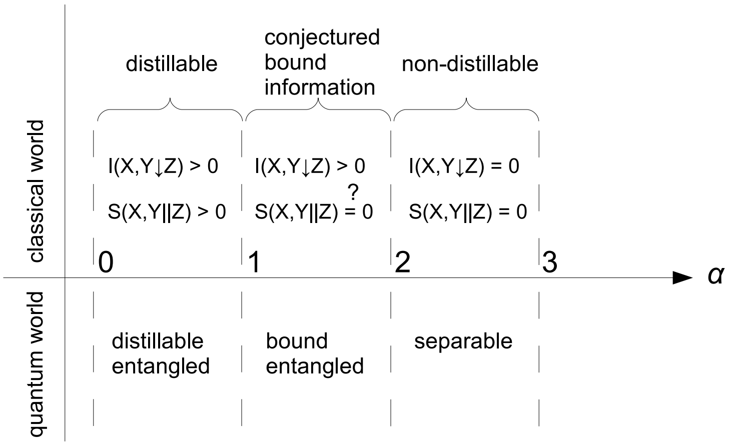


Figure 2.2: Overview over the ranges of separability/entanglement and distillability/non-distillability

2.3 Multipartite bound information and activation

The example examined in chapter 2.2 showed that there exists a range for the distribution where one cannot say certainly that secret key-agreement is possible. The existence of bound information for this range could only be conjectured. Further investigations ([10,12]) showed that the existence of bound information can be proven in the multipartite scenario, where the number of honest parties is larger than two.

The first example of a tripartite distribution containing bound information was presented in [10] and is shown in table 2.8.

A	B	C	E	P_1(A,B,C,E)
0	0	0	0	1/6
0	0	1	1	1/6
0	1	0	2	1/6
1	0	1	3	1/6
1	1	0	4	1/6
1	1	1	0	1/6

Table 2.8: Distribution for a tripartite scenario

In the following we summarize the proof of the existence of bound information for this distribution. This proof exploits the connection between the secrecy properties of the initial distribution and those of the bipartite distributions obtained by splitting the parties into two groups, where all the parties in each group are considered as a single party.

First, assume that two of the $N > 2$ honest parties, denoted by A_i and A_j are able to distill a secret key in the multipartite scenario. Then, the same holds for all bipartite splittings for which these two parties belong to different groups. Indeed, they should simply follow the same protocol as in the multipartite scenario (which, however, is not necessarily optimal for secret-key generation in the corresponding bipartite scenario).

This means that the intrinsic information has to be positive for all bipartite splittings of the parties for which A_i and A_j are separated. On the other hand, assume that the correlations in the multipartite scenario can be generated by LOPC. Then, the same is true for all bipartite distributions obtained from it. Again, for each bipartite splitting, the parties should simply apply the same LOPC protocol for the formation as in the multipartite case. Consequently, the intrinsic information has to be zero for all these bipartite distributions.

Let us now apply all these arguments to the probability distribution P_1 of table 2.8. Consider first the bipartition $AB - C$, where A and B are together. We have that $I(AB; C|E = e) = 0 \forall e \in \{1, ..., 4\}$ so we only have to consider the case of $e = 0$ which is represented in the distribution P_2 given in table 2.9.

Here $P_2(a, b, c, e) = P_1(a, b, c|e = 0) = \frac{P_1(a,b,c,e)}{\sum_{a,b,c} P_1(a,b,c,e=0)}$. The conditional mutual information is calculated as $I(AB; C|E) = \frac{1}{3}$. Now Eve performs the following maps: $1 \rightarrow 0$ and $4 \rightarrow 0$ with a remaining partial distribution P_3:

A	B	C	E	P_2(A,B,C,E)
0	0	0	0	1/2
1	1	1	0	1/2

Table 2.9: Distribution P_2 showing the part with positive conditional mutual information of P_1

A	B	C	$\overline{E}$	P_3(A,B,C,$\overline{E}$)
0	0	0	0	1/4
0	0	1	0	1/4
1	1	0	0	1/4
1	1	1	0	1/4

Table 2.10: Distribution P_3 after Eve's map

Now we obtain for the intrinsic information:

$$I(AB;C \downarrow E) = \frac{1}{3} \cdot \left(2 \cdot \frac{1}{2} \log_2 2 + 2 \cdot \frac{1}{2} \log_2 2 - 4 \cdot \frac{1}{4} \log_2 4\right) = 0 \tag{2.7}$$

Note that the same results hold for $AC - B$ because of symmetry. Therefore, none of the parties is able to distill a key. Indeed if this was the case, the intrinsic information could not be zero for both the $AB - C$ and $AC - B$ splittings.

If we now consider the last bipartition, where B and C are together, the resulting probability distribution becomes distillable. Indeed, it is enough for the single BC party to announce those cases where their symbols coincide. This leads us, with probability 1/3, to distribution P_4 in table 2.11.

A	BC	B=C	E	P_4(A,B,C,E)
0	00	1	0	1/6
0	01	0	1	1/6
0	10	0	2	1/6
1	01	0	3	1/6
1	10	0	4	1/6
1	11	1	0	1/6

Table 2.11: Distribution P_4 obtained for those cases where $B = C$

Now, the resulting probability distribution is precisely equal to a perfect secret bit, as Alice and Bob-Charlie's symbols are perfectly correlated and Eve has no information at all. Therefore, the secret-key rate for the initial $A - BC$ distribution is at least equal to $1/3$. But this is precisely equal to the conditional mutual information, so we have that the mutual information, the intrinsic information and the secret-key rate coincide and are

$$I(A; BC|E) = I(A; BC \downarrow E) = \frac{1}{3}$$

As we have positive intrinsic information for one of the bipartite splittings, the LOPC generation of P_1 by the three honest parties is impossible. However, none of the parties is able to distill this secrecy into a secret key. Hence following the definition, we have found bound information for distribution P_1. Moreover this protocol shows a way to *activate* the secret correlations in the distribution, since this secrecy become distillable when B and C are together.

3 A non-distillability criterion

The examples that have conjectured (chapter 2.2) and provable multipartite bound information (chapter 2.3) have been derived from quantum states having a very similar characterization. However, the existence of bipartite bound information still remains open. As mentioned, the main difficulty comes from the fact that one has to prove that no LOPC protocol can distill a key from a given probability distribution. In the quantum case, this was possible because there exists an easily computable criterion for non-distillability, namely the positivity of partial transposition (see appendix B.3). However, in the classical cryptographic case, we lack such a simple criterion.

The first step in this direction was provided in [13]. There, a possible criterion for detecting the non-distillability of a given probability distribution was proposed. Potentially, it could detect the presence of bound information. Therefore, the purpose of this work is to apply the criterion to some examples of probability distribution with conjectured bound information and see how it performs. The main hope was to prove bound information, but, unfortunately, this has not been the case. Actually, as we discussed later, it is also possible that the criterion is useless for detecting bound information.

In this chapter, we first present the concept of secret-bit fraction in Section 3.1, which plays a key role in all what follows, and then discuss in Section 3.2 the non-distillability criterion proposed in [13]. Later, we will apply the criterion to some candidates of probability distributions having conjectured bipartite bound information.

3.1 General ideas

As mentioned, to derive the criterion we first need to introduce the concept of secret-bit fraction.

Definition 6 *Given a distribution P_{ABE}.The secret bit fraction is given by:*

$$\lambda\left[P_{ABE}\right] = 2\frac{\sum_e \min\{P_{ABE}(0,0,e), P_{ABE}(1,1,e)\}}{\sum_{a,b,e} P_{ABE}(a,b,e)} \tag{3.1}$$

and the maximal extractable secret bit fraction is calculated by:

$$\Lambda\left[P_{ABE}\right] = \sup_{\mathcal{M}_A,\mathcal{N}_B} \lambda\left[\mathcal{M}_A\mathcal{N}_B P_{ABE}\right] \tag{3.2}$$

over all linear maps $\mathcal{M}_A$ and $\mathcal{N}_B$ acting on the spaces $\mathscr{H}_A$ and $\mathscr{H}_B$ respectively in the way: $\mathcal{M}_A : \mathscr{H}_A \rightarrow \mathscr{B}_A$ and $\mathcal{N}_B : \mathscr{H}_B \rightarrow \mathscr{B}_B$ where $\mathscr{B}$ denotes the binary output space of each alphabet.

The secret bit fraction represents the minimal part of a distribution where A and B share the same binary output value. The denominator in equation (3.1) ensures the normalization of the distribution.
The range of the maximal extractable secret bit fraction is given by $\Lambda \in [\frac{1}{2}, 1]$ where 1/2 is related to the case of uniformly distributed parties A and B, means $P_{ABE}(a,b,e) = \frac{1}{4}\forall(a,b) \in \{0,1\}$. This is related to the worst case because it means that the honest parties are independent and hence share no (secret) correlations. The best scenario is given by $P_{ABE}(0,0,e) = P_{ABE}(1,1,e) = 1/2$. Then the secret bit fraction will give the value one.
It was also shown in [14] that a secret bit fraction bigger than 1/2 is always related to a positive secret key rate or, in other words, any candidate for bound information should have maximum secret bit fraction equal to 1/2.

Another property that this criterion makes use of is that every linear map $\mathcal{M} : \mathscr{H}_1 \otimes \mathscr{H}_2 \rightarrow \mathscr{H}_3$ with positive coefficients can be split like:

$$\mathcal{M} = \mathcal{U}\mathcal{M}' \tag{3.3}$$

with

- $\mathcal{M}' : \mathscr{H}_1 \rightarrow \mathscr{H}_3 \otimes \mathscr{H}_2$
- $\mathcal{U} : (\mathscr{H}_3 \otimes \mathscr{H}_2) \otimes \mathscr{H}_2 \rightarrow \mathscr{H}_3$ where $\mathcal{U}^{y_3}_{x_3x_2y_2} = \delta^{y_3}_{x_3}\delta_{x_2y_2}$

Where upper indices are outputs and lower indices are inputs and the number indicates the corresponding space. What $\mathcal{U}$ does is to compare the inputs x_2 and y_2 from $\mathscr{H}_2$ and only if they are the same it passes the input x_3 belonging to $\mathscr{H}_3$ and passes it to the output y_3, while $\mathcal{M}'$ makes a map from x_1 to x_3 and x_2, which are both used in $\mathcal{U}$.
Figure 3.1 shows our usage of this property: We take two distributions with the spaces $\mathscr{H}_A$ and $\mathscr{H}_B$, respectively, and $\mathscr{H}_E$ and $\mathscr{H}_K$ which are of no importance in this consideration. Now we apply each map as described in equation (3.3) on either space $\mathscr{H}_A$ or space $\mathscr{H}_B$. We can do the following assignments: $\mathscr{H}_1 \equiv \mathscr{H}_{A(B)} \otimes \mathscr{B}_{A(B)}$, $\mathscr{H}_2 \equiv \mathscr{H}_{A(B)}$ and $\mathscr{H}_3 \equiv \mathscr{B}_{A(B)}$

Summarizing this: we have two input distributions G_{ABE} and Q_{ABK} where we apply the maps $\mathcal{M}_A$ and $\mathcal{N}_B$ on each alphabet. These maps are each split in the operations defined by the map $\mathcal{U}$ that compares the two input alphabets and passes the binary aphabet and map $\mathcal{M}'$ which performs an arithmetic operation to calculate the binary output.

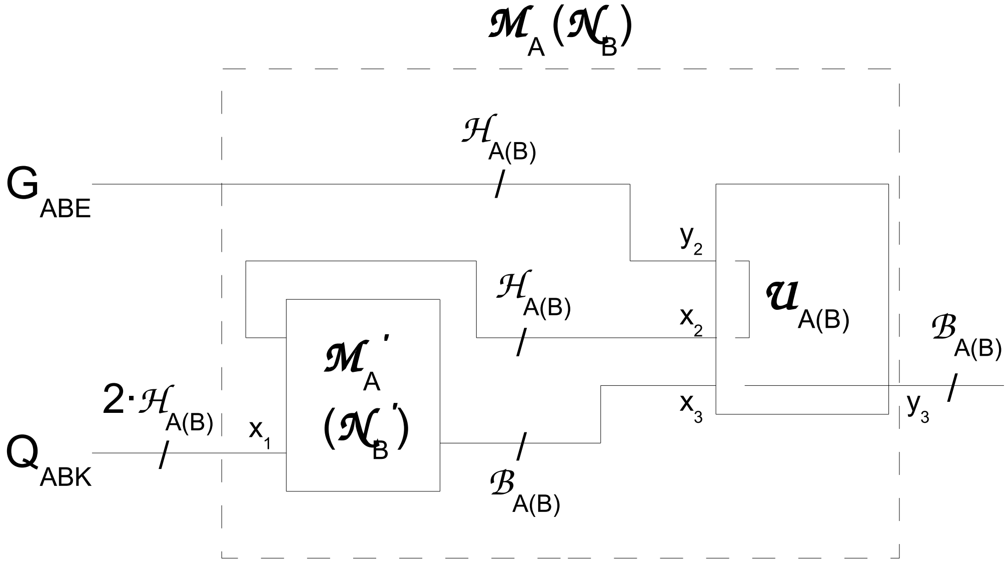


Figure 3.1: Data flow diagram of the maps $\mathcal{M}_A$ and $\mathcal{N}_B$

3.2 The criterion

Having introduced the previous concept, we are in position of presenting the criterion for non-distillability. The idea is to consider two distributions - the one, denoted by G_{ABE}, for which the presence of bound information has been conjectured and an arbitrary second probability distribution, $Q_{ABE'}$. Then we study the secret bit fraction of $Q_{ABE'}$ alone and in combination with the conjectured bound information distribution. One can then show that if G_{ABE} does not improve the secrecy of any arbitrary distribution, then it cannot be used for secret-key agreement, $S(X,Y\|Z) = 0$, and, thus, has bound information. An important aspect of the criterion is that these conditions can be mapped into a linear programming problem, which makes its numerical optimization feasible.

Let us now introduce the criterion. As stated in the previous chapter we know that a distribution is distillable if its maximal extractable secret bit fraction is bigger than 1/2. From that we can derive the following two conditions:

$$\Lambda\left[Q_{ABE'}\right] \leq \lambda_0 \tag{3.4}$$

$$\lambda\left[\mathcal{U}_A\mathcal{U}_B Q_{ABE'} \otimes G_{ABE}\right] > \lambda_0 \tag{3.5}$$

That is, if a distribution G_{ABE} is distillable, there exists another distribution $Q_{ABE'}$, with secret-bit fraction smaller than λ_0 (3.4), such that its secret bit fraction is increased when combined with G_{ABE}, see equation (3.5). Then, the distribution G_{ABE} is said to activate the distribution $Q_{ABE'}$.

To proof this we have to know that the distribution G_{ABE} is (secret-key) distillable if there exists n, such that $\Lambda[G_{ABE}^{\otimes n}] > \lambda_0$ for each $\lambda_0 \in [1/2, 1)$. If we consider the maps $\mathcal{M}_A$ and $\mathcal{N}_B$ according to equation (3.3) we can define $Q_{ABE'} = \mathcal{M'}_A \mathcal{N'}_B G_{ABE}^{\otimes(n-1)}$. Because Λ is defined by an optimization (see (3.2)) the following inequality holds:

$$\Lambda\left[G_{ABE}^{\otimes(n-1)}\right] \geq \Lambda\left[\mathcal{M'}_A \mathcal{N'}_B G_{ABE}^{\otimes(n-1)}\right] = \Lambda\left[Q_{ABE'}\right] \tag{3.6}$$

By the definition of n we know that $\Lambda\left[G_{ABE}^{\otimes(n-1)}\right] \leq \lambda_0$ which concludes inequality (3.4). The properties of the maps show that

$$\mathcal{U}_A \mathcal{U}_B Q_{ABE'} \otimes G_{ABE} = \mathcal{M}_A \mathcal{N}_B G_{ABE}^{\otimes n} \tag{3.7}$$

and we know that $\lambda\left[\mathcal{M}_A \mathcal{N}_B G_{ABE}^{\otimes n}\right] > \lambda_0$ which concludes inequality (3.5). This implies not only that the distribution G_{ABE} activates the distribution $Q_{ABE'}$, but moreover it activates itself.

Now, the aim of the criterion is to show that no such a distribution can exist, which can be seen as an optimization problem. The problem is that Eve's alphabet E' is unbounded which would lead to an endless search. However, as shown in [13], it is possible to (i) bound Eve's alphabet to a finite alphabet and (ii) map the optimization problem into a linear programming instance. These two properties make the optimization problem tractable using standard numerical techniques. Moreover in the following we only consider the case $\lambda_0 = 1/2$ that belongs to minimal distillability.
At the same time we have to check all possible pairs of maps $(\mathcal{M}_A^i, \mathcal{N}_B^i) : i = 1, \cdots, M$ that may improve the maximal extractable secret-bit fraction of $Q_{ABE'}$ above λ_0 which would violate equation (3.4).

For the linearization of equation (3.4) and (3.5) we rewrite them as

$$4\sum_{e'} \min_{a\in\{0,1\}} \left\{\left[\mathcal{M}_A^i \mathcal{N}_B^i Q_{ABE'}\right](a,a,e')\right\} - \sum_{a,b,e'} \left[\mathcal{M}_A^i \mathcal{N}_B^i Q_{ABE'}\right](a,b,e') \leq 0 \tag{3.8}$$

$$\begin{aligned} 4\sum_{e',e} \min_{a\in\{0,1\}} &\left\{\left[\mathcal{U}_A \mathcal{U}_B Q_{ABE'} \otimes G_{ABE}\right](a,a,e',e)\right\} - \\ &\sum_{a,b,e',e} \left[\mathcal{U}_A \mathcal{U}_B Q_{ABE'} \otimes G_{ABE}\right](a,b,e',e) > 0 \end{aligned} \tag{3.9}$$

Let us now define the dimension of Eve in G_{ABE} as d $(e = 1, ...d)$ and introduce the new functions:

$$s_i(e') = \begin{cases} 0 \text{ if } \sum_a (-1)^a [\mathcal{M}_A^i \mathcal{N}_B^i Q_{ABE'}](a, a, e') < 0 \\ 1 \text{ if } \sum_a (-1)^a [\mathcal{M}_A^i \mathcal{N}_B^i Q_{ABE'}](a, a, e') > 0 \end{cases} \tag{3.10}$$

$$r_e(e') = \begin{cases} 0 \text{ if } \sum_a (-1)^a [\mathcal{U}_A \mathcal{U}_B Q_{ABE'} \otimes G_{ABE}](a, a, e', e) < 0 \\ 1 \text{ if } \sum_a (-1)^a [\mathcal{U}_A \mathcal{U}_B Q_{ABE'} \otimes G_{ABE}](a, a, e', e) > 0 \end{cases} \tag{3.11}$$

For clarification one may have a closer look at the specific example $r_e(e') = 0$. Then

$$[\mathcal{U}_A \mathcal{U}_B Q_{ABE'} \otimes G_{ABE}](0, 0, e', e) < [\mathcal{U}_A \mathcal{U}_B Q_{ABE'} \otimes G_{ABE}](1, 1, e', e)$$

which is equal to say that the distribution has the smaller value with $a = 0$. This is directly related to $r_e(e') = 0$ as stated above. Therewith we can replace the *min* function by substituting a with $s_i(e')$ in equation (3.8) and with $r_e(e')$ in equation (3.9). This gives us

$$4 \sum_{e'} \left[\mathcal{M}_A^i \mathcal{N}_B^i Q_{ABE'}\right] (s_i(e'), s_i(e'), e') - \sum_{a,b,e'} \left[\mathcal{M}_A^i \mathcal{N}_B^i Q_{ABE'}\right] (a, b, e') \leq 0 \tag{3.12}$$

and

$$\begin{aligned} 4 \sum_{e',e} \Big[\mathcal{U}_A \mathcal{U}_B Q_{ABE'} \otimes G_{ABE}\Big] (r_e(e'), r_e(e'), e', e) - \\ \sum_{a,b,e',e} \Big[\mathcal{U}_A \mathcal{U}_B Q_{ABE'} \otimes G_{ABE}\Big] (a, b, e', e) > 0 \end{aligned} \tag{3.13}$$

respectively. Additionally we define the vector $\mathbf{k}(e')$ like the following:

$$\mathbf{k}(e') = [r_0(e'), r_1(e'), ..., r_d(e'), s_1(e'), ..., s_M(e')] \tag{3.14}$$

and rewrite the distribution $Q_{ABE'}$ like in table 3.1. Here one can see that several of the possible infinite outcomes of E' end up in the same vector $\mathbf{k}_j$ of the alphabet $j = 1, .., k$ with the dimension $k = 2^{d+M}$. If we now merge all coefficients with the same $\mathbf{k}_j$, like shown in equation (3.15) we will get our new finite i.e. bounded distribution Q_{ABK}.

$$Q_{ABK}(a, b, \mathbf{k}_j) = \sum_{e':\mathbf{k}(e')=\mathbf{k}_j} Q_{ABE'}(a, b, e') \tag{3.15}$$

Finally we have to adjust our summations in equations (3.12) and (3.13) from e' to the new variable $\mathbf{k}$ and conclude the equations for the algorithm:

$$\sum_{\mathbf{k}} \left(4 \cdot \left[\mathcal{M}_A^i \mathcal{N}_B^i Q_{ABK}\right] (k_{d+i}, k_{d+i}, \mathbf{k}) - \sum_{a,b} \left[\mathcal{M}_A^i \mathcal{N}_B^i Q_{ABK}\right] (a, b, \mathbf{k})\right) \leq 0 \tag{3.16}$$

$$\begin{aligned} \sum_{\mathbf{k},e} \Bigg(4 \cdot \Big[\mathcal{U}_A \mathcal{U}_B Q_{ABK} \otimes G_{ABE}\Big] (k_e, k_e, \mathbf{k}, e) - \\ \sum_{a,b} \Big[\mathcal{U}_A \mathcal{U}_B Q_{ABK} \otimes G_{ABE}\Big] (a, b, \mathbf{k}, e)\Bigg) > 0 \end{aligned} \tag{3.17}$$

$Q_{ABE'}$	$r_0(e')$	$r_1(e')$	$\cdots$	$r_d(e')$	$s_1(e')$	$\cdots$	$s_M(e')$	$\mathbf{k}_j$
q_{ab0}	0	0	$\cdots$	0	0	$\cdots$	0	
q_{ab1}	0	1	$\cdots$	0	1	$\cdots$	0	$\rightarrow \mathbf{k}_0$
$\vdots$	$\vdots$		$\ddots$				$\vdots$	$\vdots$
q_{abm}	0	0	$\cdots$	1	0	$\cdots$	0	
q_{abn}	0	1	$\cdots$	0	1	$\cdots$	0	$\rightarrow \mathbf{k}_0$
q_{abo}	1	0	$\cdots$	1	0	$\cdots$	0	
$\vdots$	$\vdots$		$\ddots$				$\vdots$	

Table 3.1: Relation between distribution $Q_{ABE'}$ and new variable $\mathbf{k}$ (illustrative)

For this transformation we still have to take into account the constraints coming from equations (3.10) and (3.11). Adjusted to our new notation we get:

$$\sum_a (-1)^a [\mathcal{M}_A^i \mathcal{N}_B^i Q_{ABK}](k_{d+i} \oplus a, k_{d+i} \oplus a, \mathbf{k}) < 0 \tag{3.18}$$

$$\sum_a (-1)^a [\mathcal{U}_A \mathcal{U}_B Q_{ABK} \otimes G_{ABE}](k_e \oplus a, k_e \oplus a, \mathbf{k}, e) < 0 \tag{3.19}$$

Let me refer to chapter 4.1 for another analysis of this set of equations.

The aim as mentioned above is to find the distribution Q_{ABK} by a linear programming such that we maximize equation (3.17) and also fulfill equations (3.16), (3.18) and (3.19). Furthermore we have the constraints from probability theory that $Q_{ABK} > 0$ and $\sum_{abk} Q_{ABK} = 1$.

If our maximization returns zero we can conclude that the distribution G_{ABE} does not activate any arbitrary distribution (including itself!) and hence its secret key rate is equal to zero. Now, if one is able to show that G_{ABE} has also positive intrinsic information, the existence of bound information can be established.

Remark If the maximization returns a positive value we cannot state anything because there might always be the case that the specific pairs of maps $(\mathcal{M}_A^i, \mathcal{N}_B^i)$ that have to be choosen in advance and show undistillability, were missing in the optimization.

4 Implementation and optimization

The implementation of the criterion explained in the previous chapter was done in MATLAB© aiming for the use of its internal *linprog* algorithm. This function is a linear programming algorithm that is defined by the formulation

$$\begin{aligned} \max_q f(q): \quad & A_{ineq} \cdot q \leq b_{ineq} \\ & A_{eq} \cdot q = b_{eq} \\ & lb \leq q \leq ub \end{aligned}$$

where capital letters represent a matrix and small letters a vector. Moreover *lb* and *ub* define bounds on the coefficients of vector q.

4.1 Analysis and implementation of the tool

The goal as already mentioned is to find the maximum over the distribution Q_{ABK} which needs to be fit to the vector q.
But first we consider the implementation of this distribution and how we programmed the maps. The distribution must be of three dimensions where the alphabets $\mathscr{H}_A$ and $\mathscr{H}_B$ must be twice the size of the given distribution G_{ABE} due to the binarization done by the maps $\mathcal{M}_A$ and $\mathcal{N}_B$. The approach was to write Alice and Bob's alphabet on the y- and x-axis, respectively, according to the graphical representation of the distributions given in this thesis. Eve is represented by the z-axis. That produces a three dimensional matrix where each slice for $Z = z$ shows a distribution over Alice's and Bob's alphabet.
We want to emphasize at this point that all coefficients are real and we do not have to care about complex numbers.

So we can create a two dimensional matrix for each value of Eve in the way shown in table 4.1 where $d_A = dim(\mathscr{H}_A)$ and $d_B = dim(\mathscr{H}_B)$ are the corresponding dimensions and α and β the binary outputs of the maps.

One can think of $Q_{ABK}(a, b, \mathbf{k}) \equiv Q^{\alpha\beta}_{ABK}(\alpha, \beta, a, b, \mathbf{k})$ to emphasize the binary output values. Hence we rewrite equations (3.17), (3.16), (3.18) and (3.19) following this new notation.

	$\beta \in \{0,1\}$	0	1
	$b \in \mathscr{H}_B$	$0 \;\cdots\; d_b$	$0 \;\cdots\; d_b$
$\alpha \in \{0,1\}$	$a \in \mathscr{H}_A$		
0	0 $\vdots$ d_a	Q_{ABK}	
1	0 $\vdots$ d_a		

Table 4.1: Q_{ABK} arrangement for the use in the program

4.1.1 Combination of both distributions

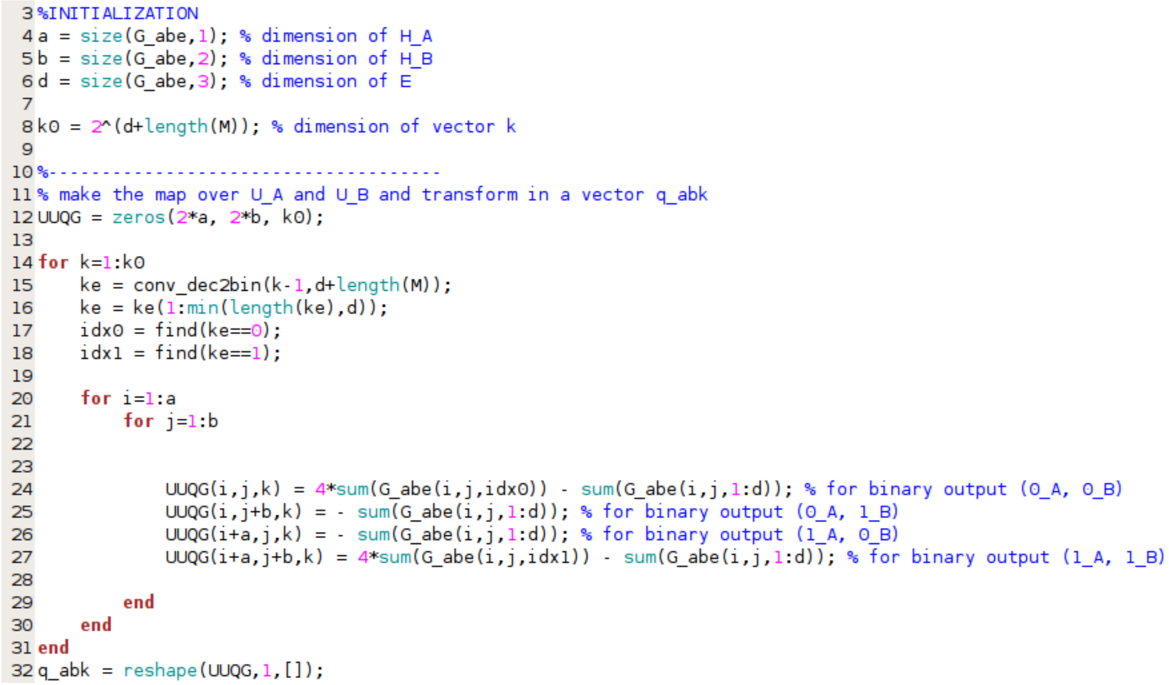
$$\max_{Q_{ABK}} \sum_{\mathbf{k},e} \Bigg(4 \cdot \sum_{a,b} \Big[\mathcal{U}_A \mathcal{U}_B Q^{\alpha\beta}_{ABK} \otimes G_{ABE} \Big](k_e, k_e, a, b, \mathbf{k}, e) \\ - \sum_{\alpha,\beta,a,b} \Big[\mathcal{U}_A \mathcal{U}_B Q^{\alpha\beta}_{ABK} \otimes G_{ABE} \Big](\alpha, \beta, a, b, \mathbf{k}, e) \Bigg) = 0 \tag{4.1}$$

Equation (4.1) takes each specific plane e in G_{ABE} and plane $\mathbf{k}$ in $Q^{\alpha\beta}_{ABK}$ and makes a product between all values a and b. Therefore we take the corresponding quarter of $Q^{\alpha\beta}_{ABK}$ characterized by α and β (see table 4.1) and multiply it with G_{ABE}. Finally we make the sum over the a and b values to get the probability.
The corresponding quarter of $Q^{\alpha\beta}_{ABK}$ for the first summand is depending on the value k_e that is deduced from the e-th coefficient in $\mathbf{k}$. The second summand takes consecutively all parts, i.e. for $(\alpha,\beta) = (0,0)$ then (0,1), (1,0) and (1,1).

Thus we can say that every quarter section in $Q^{\alpha\beta}_{ABK}$ will be substracted by its corresponding values in G_{ABE}. Moreover due to the tensor product we have to substract the sum over all e for each corresponding (a, b)-pair as given in figure 4.1 line 24-27.

How do we include the first summand? We know that we only have to care about the parts $(\alpha,\beta) = (0,0)$ or $(1,1)$ which depend on k_e. So we have to find for each $\mathbf{k}$ and e whether $k_e = 0$ or 1, then remember those indices (figure 4.1, line 15-18) and add the corresponding range (figure 4.1, line 24 and 27).

Finally we reshape the matrix to the vector to introduce it in the linear programming which is going to find the maximal values for the distribution $Q^{\alpha\beta}_{ABK}$.



```
%INITIALIZATION
a = size(G_abe,1); % dimension of H_A
b = size(G_abe,2); % dimension of H_B
d = size(G_abe,3); % dimension of E

k0 = 2^(d+length(M)); % dimension of vector k

%-----------------------------------
% make the map over U_A and U_B and transform in a vector q_abk
UUQG = zeros(2*a, 2*b, k0);

for k=1:k0
    ke = conv_dec2bin(k-1,d+length(M));
    ke = ke(1:min(length(ke),d));
    idx0 = find(ke==0);
    idx1 = find(ke==1);

    for i=1:a
        for j=1:b


            UUQG(i,j,k) = 4*sum(G_abe(i,j,idx0)) - sum(G_abe(i,j,1:d)); % for binary output (0_A, 0_B)
            UUQG(i,j+b,k) = - sum(G_abe(i,j,1:d)); % for binary output (0_A, 1_B)
            UUQG(i+a,j,k) = - sum(G_abe(i,j,1:d)); % for binary output (1_A, 0_B)
            UUQG(i+a,j+b,k) = 4*sum(G_abe(i,j,idx1)) - sum(G_abe(i,j,1:d)); % for binary output (1_A, 1_B)

        end
    end
end
q_abk = reshape(UUQG,1,[]);
```

Figure 4.1: Code implementation of equation (4.1)

4.1.2 Constraint on secret-bit fraction of $Q_{ABK}^{\alpha\beta}$

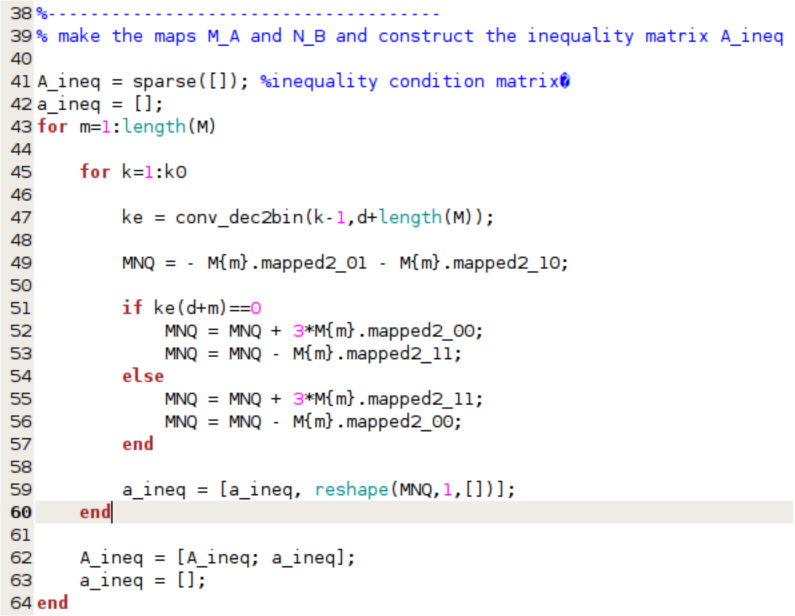
$$\begin{split}\sum_{\mathbf{k}}\Bigg(4\cdot\sum_{a,b}\left[\mathcal{M}_A^i\mathcal{N}_B^i Q_{ABK}^{\alpha\beta}\right](k_{d+i},k_{d+i},a,b,\mathbf{k})\\ -\sum_{\alpha,\beta,a,b}\left[\mathcal{M}_A^i\mathcal{N}_B^i Q_{ABK}^{\alpha\beta}\right](\alpha,\beta,a,b,\mathbf{k})\Bigg)<0\end{split}\tag{4.2}$$

Equation (4.2) is the first constraint on our optimized distribution. Here we have nearly an equivalent formalism to equation (4.1) except for the tensor product with G_{ABE}. The first summand depends in this case on the $(d+i)$-th coefficient in $\mathbf{k}$ that is related to the i-th pair of maps $\mathcal{M}_A$ and $\mathcal{N}_B$.

We will discuss the implementation of the maps in chapter 4.5. So far let us assume that they are predefined matrices of the size $2\cdot d_A\times 2\cdot d_B$ and stored in a structure data type with corresponding matrices to each binary output combination denoted by *mapped2_00, mapped2_01, etc*.

A linear map is characterized by a product of the mapping coefficient with the value. Hence we have to arrange the maps according to the part $(\alpha,\beta)=(0,0)$ or $(1,1)$ which is checked

in figure 4.2, line 51.
The rest is a straight forward imlementation of equation (4.2) that is closed by the reshape of the matrix to the vector notation. Finally we store everything in the matrix *A_ineq* which contains all constraints on the optimized vector.



```
38 %------------------------------------
39 % make the maps M_A and N_B and construct the inequality matrix A_ineq
40
41 A_ineq = sparse([]); %inequality condition matrix�
42 a_ineq = [];
43 for m=1:length(M)
44
45     for k=1:k0
46
47         ke = conv_dec2bin(k-1,d+length(M));
48
49         MNQ = - M{m}.mapped2_01 - M{m}.mapped2_10;
50
51         if ke(d+m)==0
52             MNQ = MNQ + 3*M{m}.mapped2_00;
53             MNQ = MNQ - M{m}.mapped2_11;
54         else
55             MNQ = MNQ + 3*M{m}.mapped2_11;
56             MNQ = MNQ - M{m}.mapped2_00;
57         end
58
59         a_ineq = [a_ineq, reshape(MNQ,1,[])];
60     end
61
62     A_ineq = [A_ineq; a_ineq];
63     a_ineq = [];
64 end
```

Figure 4.2: Code implementation of equation (4.2)

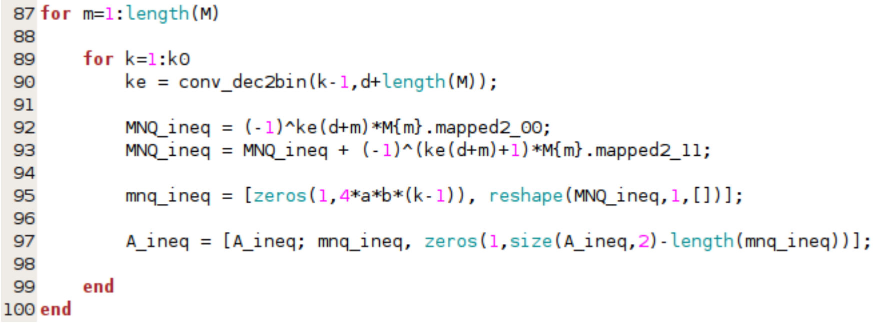
4.1.3 Additional constraint 1

$$\sum_{\alpha}(-1)^{\alpha}\sum_{a,b}\left[\mathcal{M}^{i}_{A}\mathcal{N}^{i}_{B}Q^{\alpha\beta}_{ABK}\right](k_{d+i}\oplus\alpha, k_{d+i}\oplus\alpha, a, b, \mathbf{k}) < 0 \tag{4.3}$$

Equation (4.3) initially represents equation (3.10) that assures the *min* function condition from equation (3.8). Here we are only looking at the parts $(\alpha,\beta) = (0,0)$ or (1,1). Hence it is enough to consider only one binary value, α in this case. Moreover we depend on the value k_{d+i} like in equation (4.2). The implementation of equation (4.3) is presented in figure 4.3, line 92 and 93.

After that we adapt the matrix (*MNQ_ineq*) to our *A_ineq* matrix which contains all the constraints for the linear programming. This is done by adding zeros before and after the reshaped vector of *MNQ_ineq* as seen in figure 4.3, line 95 and 97.

Finally this condition has to be fulfilled for every $i = 1, \cdots, M$ (in the code of figure 4.3 denoted as m) and every $\mathbf{k} = 1, \cdots, 2^{d+M}$.



```
for m=1:length(M)

    for k=1:k0
        ke = conv_dec2bin(k-1,d+length(M));

        MNQ_ineq = (-1)^ke(d+m)*M{m}.mapped2_00;
        MNQ_ineq = MNQ_ineq + (-1)^(ke(d+m)+1)*M{m}.mapped2_11;

        mnq_ineq = [zeros(1,4*a*b*(k-1)), reshape(MNQ_ineq,1,[])];

        A_ineq = [A_ineq; mnq_ineq, zeros(1,size(A_ineq,2)-length(mnq_ineq))];

    end
end
```

Figure 4.3: Code implementation of equation (4.3)

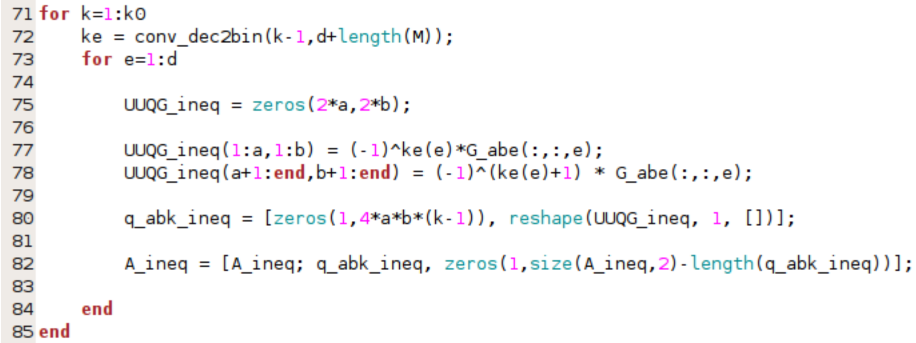
4.1.4 Additional constraint 2

$$\sum_{\alpha}(-1)^{\alpha}\sum_{a,b}\left[\mathcal{U}_A\mathcal{U}_B Q^{\alpha\beta}_{ABK}\otimes G_{ABE}\right](k_e\oplus\alpha, k_e\oplus\alpha, \mathbf{k}, e) < 0 \tag{4.4}$$

Equation (4.4) has the same functionality for equation (3.9) as (4.3) for (3.8) and is a straight forward implementation as done in the previous case.

The implementation of equation (4.4) is shown in figure 4.4, line 77 and 78 and has to be adapted to the condition matrix *A_ineq*, line 80 and 82.

In this case we have to consider all combinations of $e = 1, \cdots, d$ and $\mathbf{k} = 1, \cdots, 2^{d+M}$.

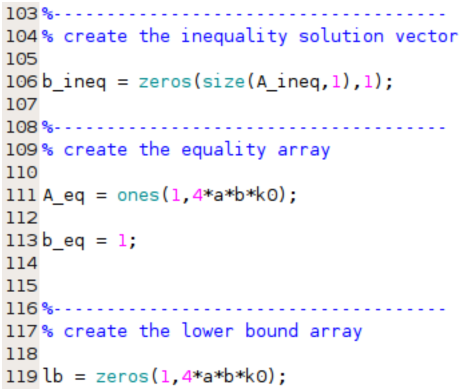


```
71 for k=1:k0
72     ke = conv_dec2bin(k-1,d+length(M));
73     for e=1:d
74 
75         UUQG_ineq = zeros(2*a,2*b);
76 
77         UUQG_ineq(1:a,1:b) = (-1)^ke(e)*G_abe(:,:,e);
78         UUQG_ineq(a+1:end,b+1:end) = (-1)^(ke(e)+1) * G_abe(:,:,e);
79 
80         q_abk_ineq = [zeros(1,4*a*b*(k-1)), reshape(UUQG_ineq, 1, [])];
81 
82         A_ineq = [A_ineq; q_abk_ineq, zeros(1,size(A_ineq,2)-length(q_abk_ineq))];
83 
84     end
85 end
```

Figure 4.4: Code implementation of equation (4.4)

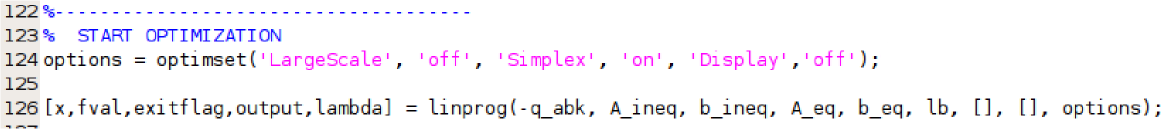
4.1.5 Remaining conditions and start of the optimization

Finally we have to add the bounds of probability theory that the sum over all values must be equal one and that all values are positive like shown in figure 4.5. Moreover we create the inequality vector *b_ineq* and then start the optimization (figure 4.6).



```
%------------------------------------
% create the inequality solution vector

b_ineq = zeros(size(A_ineq,1),1);

%------------------------------------
% create the equality array

A_eq = ones(1,4*a*b*k0);

b_eq = 1;


%------------------------------------
% create the lower bound array

lb = zeros(1,4*a*b*k0);
```

Figure 4.5: Code implementation of the bounds from probability theory



```
%------------------------------------
%  START OPTIMIZATION
options = optimset('LargeScale', 'off', 'Simplex', 'on', 'Display','off');

[x,fval,exitflag,output,lambda] = linprog(-q_abk, A_ineq, b_ineq, A_eq, b_eq, lb, [], [], options);
```

Figure 4.6: Start of the linear programming

4.2 Implementation of the maps $\mathcal{M}_A$ and $\mathcal{N}_B$

As we know from equation (3.2) the maps $\mathcal{M}_A$ and $\mathcal{N}_B$ perform a map to a binary alphabet that is needed to calculate the secret-bit fraction. That means we perform a binarization of our distribution Q_{ABK} to $\alpha \in \mathscr{B}_A$ and $\beta \in \mathscr{B}_B$.

If we consider each alphabet separately we can implement the maps as shown in table 4.2. There l is the probability of mapping $a = 0$ to $\alpha = 0$, m the probability of $a = 1 \rightarrow \alpha = 0$ etc. o maps $a = 0 \rightarrow \alpha = 1$ and so on. Following this notation we can now create a mapping vector for each alphabet and by a simple vector multiplication we get the mapping matrix that is used in the previous sections taking into account which alphabet is mapped to what binary output value.

α:	0				1			
a:	0	1	$\cdots$	d_A	0	1	$\cdots$	d_A
	l	m	$\cdots$	n	o		$\cdots$	

Table 4.2: General view of a binarization map acting on one alphabet.

4.3 Examined distributions

Now we introduce the distributions that we have examined with this tool. One is taken from [15] and is represented in table 4.3. Another one that was detected through investigations and also seemed quite promising is shown in table 4.4.

X Y (Z)	0	1
0	(1) $(1-\delta)/4+\delta/2$ (2) $(1-\delta)/8+\delta/2$ (3) $\delta/2$ (4) $\delta/2$	(1) $(1-\delta)/8$ (2) $(1-\delta)/8$
1	(3) $(1-\delta)/8$ (4) $(1-\delta)/8$	(1) $\delta/2$ (2) $\delta/2$ (3) $(1-\delta)/8+\delta/2$ (4) $(1-\delta)/4+\delta/2$

Table 4.3: Distribution D_1 to show bound information

Both are good candidates to find bound information because depending on the particular parameter, research showed that there are different limits for $S(X,Y||Z) > 0$ and $I_{form}(X,Y|Z)$ respectively $I(X,Y \downarrow Z) > 0$.

So for distribution D_1 one detected that the secret key rate is positive for $\delta \gtrsim 0.093$ but the intrinsic information is positive for the whole range of $\delta \in [0,1]$. That means we have the region $[0,0.093]$ where we assume to have bound information. Hence in this range the algorithm should give us a zero as the maximal extractable secret bit fraction, which would give us the proof for undistillability.

X Y (Z)	0	1
0	(1) $\beta/4$	(3) $(1-\beta)/8$
1	(2) $(1-\beta)/8$	(1) $(1-\beta)/4$ (2) $(1-\beta)/8$ (3) $(1-\beta)/8$

Table 4.4: Distribution D_2 to show bound information

For the second distribution D_2 we know that it is undistillable for $\beta < 1/3$ whereas the intrinsic information is positive for $\beta > 3 - 2\sqrt{2} \approx 0.17$. Hence the interesting range here is $[0.17, 1/3]$.

4.4 Problems

The implementation of the formulas could be done straight forward following [13]. But the matrices for the constraints of the linear programming reached too big dimensions because when aligning our three dimensional matrix Q_{ABK} we get a vector of length $2{\cdot}d_A{\times}2{\cdot}d_B{\times}2^{d+M}$. That gives for distribution $D_1 : 256 \cdot 2^M$ and $D_2 : 128 \cdot 2^M$. As one can see with the number of pairs of maps M our vector dimension i.e. the number of variables to be optimized increases exponentially.
Beginning with five pairs of maps, our q vector has length 8192 or 4096 respectively but with nine pairs we already end up in 131072 or 65536.
Moreover the linear programming needs to take into account the M inequalities from constraint (4.2), $d \cdot 2^{d+M}$ inequalities from (4.3) and $M \cdot 2^{d+M}$ inequalites from (4.4).This creates a matrix that is also increasing exponentially with the number of maps.
Hence due to memory limitations we are tied in the number of maps that we can introduce to the optimization.

Another difficulty that occurs directly from the mathematical description of the algorithm is the total amount of possibile maps one can perform from any alphabet to the binary one namely infinite ones.

4.5 Solutions and improvements

4.5.1 Maps of 100% and 0%

Primarily due to simplification we considered only maps with a probability weight of 100% or 0% to the binary output. For the tool we wanted to include as many pairs as possible, even though the dimensions of the matrices and hence the calculation time increases exponentially with this number. The maps for one alphabet $\mathscr{H}_A \rightarrow \mathscr{B}_A$ are described in table 4.5.

Let me remark here that the non-distillability criterion from chapter 3 does not need normalized values because of the denominator in equation (3.1).

That means we have to check $M = 2^4 \cdot 2^4$ pairs of maps to include all possibilites for $\mathcal{M}_A$ and $\mathcal{N}_B$. As mentioned above this is too much, so we had to improve the algorithm further. Our idea was to take only the best maps, i.e. those maps that give the lowest maximal values

$$\begin{array}{c|cc} \alpha: & \multicolumn{2}{c}{0} \\ \hline \text{a:} & 0 & 1 \\ \hline\hline 1. & 0 & 0 \\ 2. & 0 & 1 \\ 3. & 1 & 0 \\ 4. & 1 & 1 \end{array} \times \begin{array}{c|cc} \alpha: & \multicolumn{2}{c}{1} \\ \hline \text{a:} & 0 & 1 \\ \hline\hline 1. & 0 & 0 \\ 2. & 0 & 1 \\ 3. & 1 & 0 \\ 4. & 1 & 1 \end{array} \Longrightarrow \begin{array}{c|cc|cc} \alpha: & \multicolumn{2}{c|}{0} & \multicolumn{2}{c}{1} \\ \hline \text{a:} & 0 & 1 & 0 & 1 \\ \hline\hline 1. & 0 & 0 & 0 & 0 \\ 2. & 0 & 0 & 0 & 1 \\ 3. & 0 & 0 & 1 & 0 \\ \vdots & & \multicolumn{2}{c}{\cdots} & \vdots \\ 2^4. & 1 & 1 & 1 & 1 \end{array}$$

Table 4.5: Maps of alphabet $\mathscr{H}_A$ to $\mathscr{B}_A$ with probabilities of 100% and 0%

after the optimization.

One way to implement this is to introduce a loop that checks the output of the linear programming started with only one pair of maps. Thus the calculation time also decreases exponentially and we get a rough sketch about the quality of each pair. Then we take the best ones and start the main optimization where we can adjust the total number of used pairs and hence the calculation time.
It is necessary to start the optimization with several maps, because the criterion considers how one function can be maximized while its values are bounded by several other functions.

In figure (4.7) we illustrate the outcome of the optimization over the possible combinations of maps. We can see that the majority ends in the same large value. Those ones do not improve the tool and can thus be discarded. So we may choose only the best five pairs for the main optimization.

Following this improvement we could run a general simulation round through both distributions over the whole range of each parameter. These results are presented in chapter 4.6.1.

4.5.2 Decimal maps

Thereupon we expanded the maps further to introduce decimal mapping coefficients. I.e. maps to the binary output alphabet in the manner, given in table 4.6. This gives us a set of $11^4 = 14,641$ possible maps for the alphabet $\mathscr{H}_A$. The same number shall be applied to the other alphabet which leads us to a total amount of possible maps - to be checked: $M' = 11^4 \cdot 11^4 = 214,358,881$.

Refering to the fact that the criterion makes a normalization of the distribution itself and including the basic maps of table 4.5, we are able to exclude due to redundancy the maps

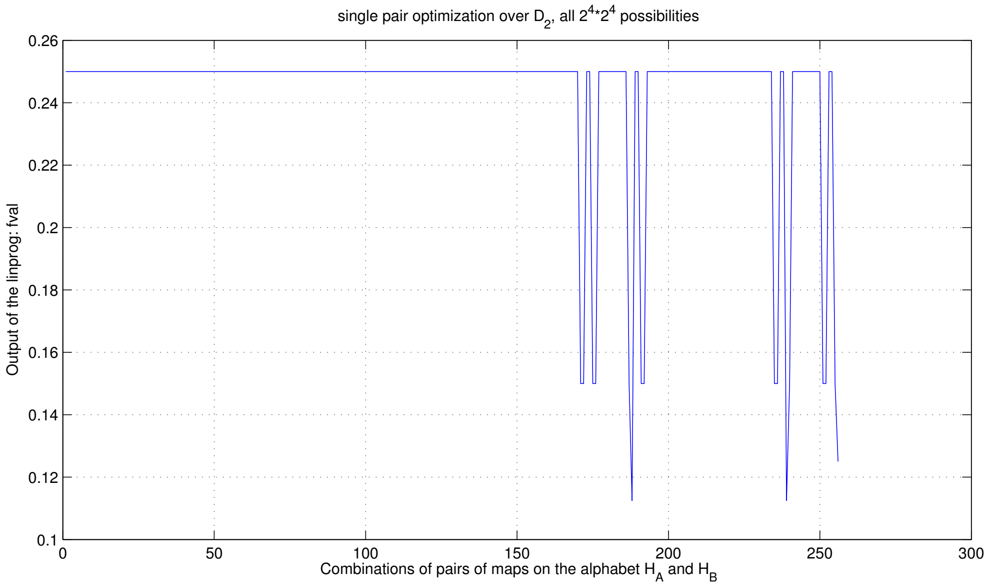


Figure 4.7: Single pair optimization over D_2 for all $2^4 \cdot 2^4$ possible combinations

that follow the condition:

$$P_{\alpha|a}(i,0) = P_{\alpha|a}(i,1) \text{ for } i = \{0,1\}$$

and the same for alphabet $\mathscr{H}_B$. This gives us a set of $m = 2^2 + 11 \cdot 10 = 114$ maps to $\alpha = 0$. Developing this strategy we have the same amount of maps to $\alpha = 1$ and the whole number of maps for the alphabet $\mathscr{H}_B$, too. That makes a total amount of possible pairs of maps: $M = m^4 = 168,896,016$ - to be checked individually. This represents a reduction to M' of 21%.

The conclusions and results of this version of the tool are presented in chapter 4.6.2.

4.5.3 Including the eavesdropper

Another idea for an improvement has been to include Eve's maps in the secret-bit fraction formula, but it has been shown in [14] that this has no influence on the secret-bit fraction:

Let $\Gamma_{\tilde{E}|E}$ be an arbitrary operation Eve may perform on the distribution. Then

$$\begin{aligned}
\lambda[\Gamma_{\tilde{E}|E} P_{ABE}] &= 2\sum_{e'} \min\left[\sum_{e} \Gamma_{\tilde{E}|E}(e',e) P_{ABE}(0,0,e), \sum_{e} \Gamma_{\tilde{E}|E}(e',e) P_{ABE}(1,1,e)\right] \\
&\geq 2\sum_{e',e} \Gamma_{\tilde{E}|E}(e',e) \min\left[P_{ABE}(0,0,e), P_{ABE}(1,1,e)\right] \\
&= 2\sum_{e} \min\left[P_{ABE}(0,0,e), P_{ABE}(1,1,e)\right] = \lambda[P_{ABE}]
\end{aligned}$$

α:	0		1	
a:	0	1	0	1
1.	0	0	0	0
2.	0	0	0	0.1
3.	0	0	0	0.2
$\vdots$		$\cdots$		$\vdots$
9.	0	0	0	1
10.	0	0	0.1	0
11.	0	0	0.1	0.1
$\vdots$		$\cdots$		$\vdots$
$(11^4 - 1)$.	1	1	1	0.9
11^4.	1	1	1	1

Table 4.6: Decimal maps for the binarization

The inequality comes from the *min* function and is independent on Eves' actions.

4.6 Results

4.6.1 Maps of 100% and 0%

Following the description in chapter 4.5.1 we implemented the program and could obtain the results described in this section.
For the distribution D_1 shown in table 4.3 we checked the range: $\delta = \{0.01, 0.02, \cdots, 0.10\}$ and for the distribution D_2 from table 4.4 we took: $\beta = \{0.1, 0.2, 0.3\}$.
The results for D_1 are presented in figure 4.8 that shows the solutions of the linear programming $fval$ and additionally the conditional mutual information of distribution D_1 for each δ. We can see that there is no exceptional behaviour in the curve, neither for the distillable region $\delta \in [0.093, 1]$, represented by $\delta = 0.1$, nor for the uncertain range $\delta \in [0, 0.093)$.
Our goal of reaching $fval = 0$ and hence showing $S(X, Y||Z) = 0$ could not be reached with this family of maps.

The results for D_2 were similar i.e. they did not return zero from the maximization. The specific values can be taken from table 4.7 (range of conjectured bound information: $\beta = [0.17, 0.33]$).

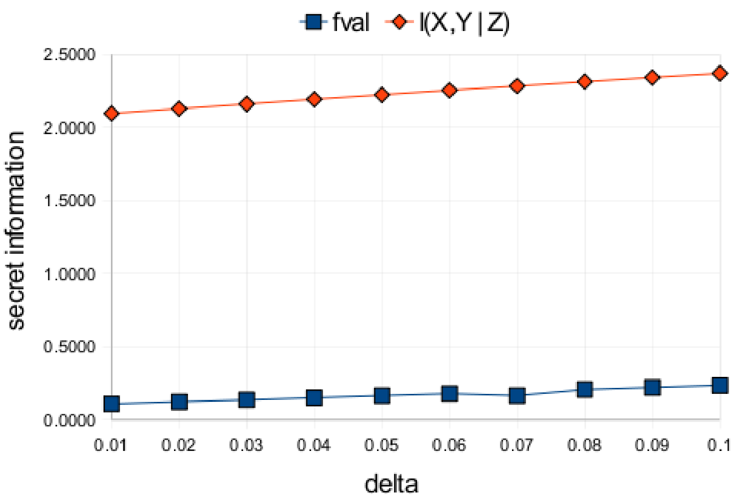


Figure 4.8: Outputs of the optimization for distribution D_1 over the uncertain range including the maps of table 4.5

β	0.1	0.2	0.3
$fval$	0.012	0.025	0.037

Table 4.7: Results of the optimization of distribution D_2

4.6.2 Decimal maps

After adapting the program to the specifications of chapter 4.5.2 we were able to measure the following time consumptions with the given server details:

- Servers capacities:
  - . Quadcore processors with 2.2 - 2.8 GHz
  - . Main memory per node 8 - 24 GB
  - . Architecture: 64 bit
- After 17 hours the server passed 0.2% of all pairs of maps.
- That makes a total calculation time of $17h/0.002 = 8500h \approx 350d$.
- This measurement was based on a loop with a backup in each pass. By the *profiler* function in MATLAB© we could detected that each backup takes $0.11ms$, which has no big influence at all on the complete optimization: $t = 0.11ms * 114^4/3600/24 = 0.21d$.

Concluding one can see that we were limited with the number of possible pairs that we could check. So we decided to examine only a suitable part in the remaining time and with the given capabilities.

Partial results

Even though we were not able to simulate a closed bigger range of pairs of maps we concentrated on simpler combinations between them. Therefore we concluded to use pairs where two maps are always equivalent:

(1) same maps inside each alphabet:

$$\begin{array}{lcll} a \to \alpha = 0 & \equiv & a \to \alpha = 1 & \text{and} \\ b \to \beta = 0 & \equiv & b \to \beta = 1 & \end{array}$$

(2) same maps to the same output value:

$$\begin{array}{lcll} a \to \alpha = 0 & \equiv & b \to \beta = 0 & \text{and} \\ a \to \alpha = 1 & \equiv & b \to \beta = 1 & \end{array}$$

Therefore we were able to reduce the amount of pairs of maps to $114^2 = 12996$. The results were quite similar for both distributions thus we consider only the results for distribution D_2.

In figure 4.9 we plot the outcomes of the optimization for each pair of maps based on distribution D_2 and following the restriction of (2). One can see that there exists a periodicity in the combination of maps that is due to the loops that start every pass with very low mapping coefficients.

From both cases (1) and (2) we obtained the same optimized value for the distribution D_2:

$$fval\,(\beta = 0.2) = 0.005$$

which is an improvement of 80% but still not enough to show our conjectured quantity. This value has been obtained with the maps:

α:	0		1	
a:	0	1	0	1
	0.3	0.7	0.3	0.7

β:	0		1	
b:	0	1	0	1
	0.3	0.7	0.3	0.7

A more specific simulation around the best mapping vector combination mentioned above, with a probability step size of 0.01 did not improve the result.

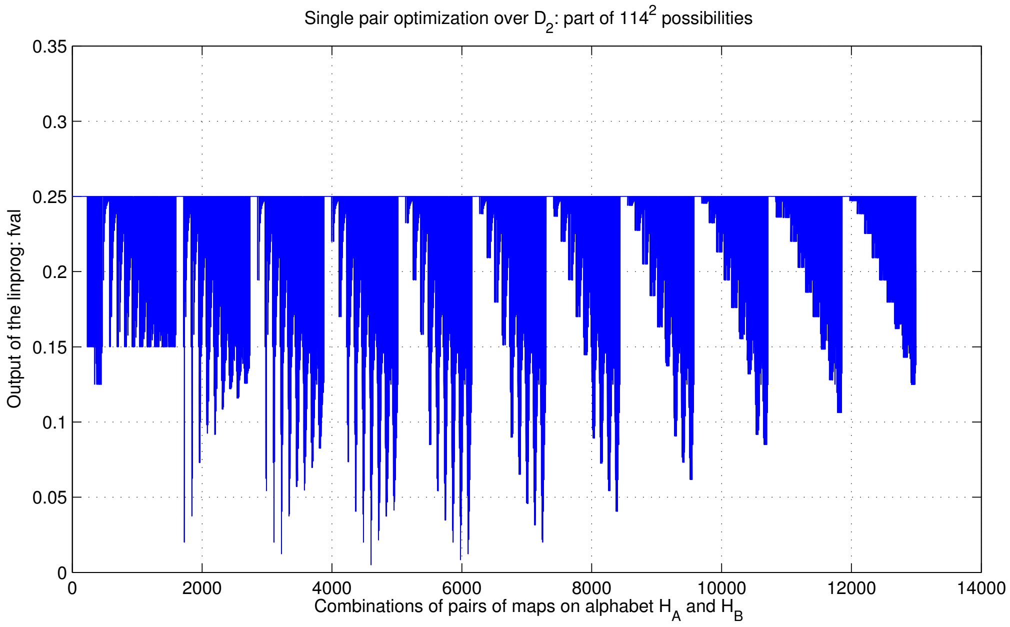


Figure 4.9: Optimization over the partial pairs of maps over the distribution D_2

5 Conclusion and outlook

The problem of secure communication is usually solved by making assumptions on the eavesdropper's computational power - the computational security. There is a however a stronger form of security, known as information-theoretical security, where a protocol can be shown to be secure using Information Theory terms. In this formalism, it is enough to consider the part of the secret-key agreement because there are protocols that ensure the secret communication given a secret-key. The standard key-agreement scenario consists of two honest parties, that communicate over an authenticated channel, and the wire-tapper, all of them sharing correlated random variables described by a joint probability distribution. One of the main questions is to understand how, if possible, the honest parties can distill a secret key out of their correlated variables. There exists strong evidence towards the existence of distributions that contain an irreversible form of secrecy, i.e. secrecy that cannot be distilled into a secret-key by local operations and public communication. This form of secrecy, known as bound information, has been proven so far for the multipartite case, with more than two honest parties, but is still an open question for the more natural bipartite case.

The main difficulty in proving the existence of bound information comes from the fact that, given an initial probability distribution, one has to show that there is no protocol leading to a secret key. Indeed, all the support so far to the existence of this quantity comes from probability distributions containing secret correlations that cannot be distilled into a secret key by any of the known protocols. The non-distillability criterion discussed in chapter 3 is a potentially promising approach to find this irreversible form of secrecy, as may allow proving the non-distillability of a distribution.

The main goal of this thesis was first to implement this criterion, to test it and then to simulate the most suitable distributions. From the theory we knew that the algorithm could be adapted to a linear programming algorithm that was already available in the MATLAB© repository. The adaptation did not turn out to be such a big problem, but it pointed out that the testing was the bigger challenge. The condition that a non-distillable distribution has a secret-bit fraction of 1/2 is necessary so we could not check the functionality by a distribution that is known to be non-distillable. Therefore we had to revise the code step by step.
Once we were certain that the program worked well we faced the next problem of the bi-

narization maps. With the number of maps we exceeded the memory of the simulation machines and also the suitable limit of simulation time. So it was necessary to reduce the number of pairs of maps. This was accomplished by the removal of redundant maps and a pre-optimization to detect the most suitable ones.
Due to the limited remaining time we decided to simulate a part of all possibilities to obtain meaningful results. Indeed we were able to show a big improvement towards the non-distillability of our distribution, but none of the obtained results was conclusive.

For a complete test it is an option to split the missing part of the maps in several smaller optimizations and combine those results to attempt to obtain the highest efficiency of the program in an iterative way. This can also be automated by another program to facilitate the process.
Another approach can be to find the best maps through another non-linear optimization. This is non-linear because we have to find two matrices for alphabet A and B at the same time. These results may be introduced as the starting point of the maximization of our linear programming.

We would also like to mention that it is also possible that the proposed criterion is in fact useless to prove the existence of bound information. The main idea of the criterion is to show that the initial distribution, with conjectured bound information, cannot improve the secrecy properties of any distribution. If this is the case, the distribution has to be non-distillable. Recall however that bound information was introduced as a cryptographic analog of bound entanglement, an irreversible form of quantum correlations observed in Quantum Information Theory. In the quantum case, all bound entangled states have been shown to improve the entanglement properties of another state. If the same was true for probability distributions, the analyzed criterion would be useless for the detection of bound information. This is however a theoretical open question in the classical case that deserves further investigation.

To conclude, the existence of bound information, conjectured in 2000 by Gisin and Wolf, is a nice and natural question in the key agreement scenario that remains open in spite of years of research. In this work, we have tested the first proposed criterion for the detection of non-distillable secret correlations. The obtained results somehow give more evidence for the existence of this quantity but, unfortunately, cannot solve the problem.

A Appendix: Conditional mutual information

The mutual information gives us some knowledge about the correlation between two parties within their distribution. It can be written in the following forms:

$$\begin{aligned} I(X,Y) &= \sum_{x\in X}\sum_{y\in Y} P(x,y)\log\frac{P(x|y)}{P(x)} \\ &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \\ &= H(X) + H(Y) - H(X,Y) \end{aligned}$$

The conditional entropy for the tripartite scenario can be derived as:

$$\begin{aligned} H(X,Y|Z) &= H(Y|Z,X) + H(X|Z) \\ &= H(X,Y,Z) - H(Z) \end{aligned}$$

And hence we can formulate the conditional mutual information as the correlation between two parties given the information of a third one:

$$I(X,Y|Z) = H(X|Z) - H(X|Y,Z)$$

here we used the formulas for the n-dimensional case:

$$H(X_1,...X_n) = \sum_{i=1}^{n} H(X_i|X_{i-1},...X_1) \tag{A.1}$$

$$H(X_1,...X_n|Y) = \sum_{i=1}^{n} H(X_i|Y;X_1,...X_{i-1}) \tag{A.2}$$

B Appendix: Introduction to quantum mechanics

Most of the distributions analyzed in this thesis are derived from measurements applied to tripartite quantum states. The very same concept of bound information was indeed proposed as a classical analog of bound entanglement, an irreversible form of quantum correlations appearing in quantum information theory. For the sake of completeness, we provide in this appendix a short introduction to the basic mathematical objects of quantum mechanics in general and, later, quantum information theory. This chapter is not thought to be a complete summary of quantum theory. Therefore we would like to refer those readers interested in the quantum formalism to [16]. Indeed, most of the discussion in the next lines follows this reference.

B.1 Postulates of quantum mechanics

In quantum mechanics on uses the *bra* $\langle\phi|$ and *ket* $|\phi\rangle$ notation to represent a quantum states, where ket is considered to be a columnvector and bra is the adjoint one, i.e. $\langle\phi| = (|\phi\rangle^*)^T$, both having complex elements.

B.1.1 State space

Postulate Associated to any isolated physical system is a complex vector space with inner product (that is, a Hilbert space) known as the *state space* of the system. The system is completely described by its *state vector*, which is a unit vector in the system's state space.

The simplest quantum mechanical system is the *qubit*, which corresponds to a two - dimensional Hilbert space. Suppose $|0\rangle$ and $|1\rangle$ form an orthonormal basis for that state space. Then an arbitrary state vector in the state space can be written as the superposition of the basis vectors

$$|\psi\rangle = a|0\rangle + b|1\rangle, \tag{B.1}$$

where $a, b \in \mathbb{C}$ and $|\psi\rangle$ is a new valid state of the system. This is main difference to a classical bit which can only be in the zero or one state. Thus the qubit system is located in a

2-dimensional state space with the *computational basis states* $|0\rangle$ and $|1\rangle$.
The condition that $|\psi\rangle$ is a unit vector, i.e. the *inner product* is one ($\langle\psi|\psi\rangle = 1$) which is known as the *normalization condition*, leads to the formulation that $|a|^2 + |b|^2 = 1$. We can talk about $|a|^2$ and $|b|^2$ as the probabilites related to the events that our qubit system is in state $|0\rangle$ or $|1\rangle$, respectively.

B.1.2 Evolution

Postulate The evolution of a *closed* quantum system is described by a *unitary transformation.* That is, the state $|\psi\rangle$ of the system at time t_1 is related to the state $|\psi'\rangle$ of the system at time t_2 by a unitary operator U which depends only on the times t_1 and t_2,

$$|\psi'\rangle = U|\psi\rangle. \tag{B.2}$$

A closed quantum system is defined by no interaction with its environment. This is a quite unrealistic assumption because all systems interact with each other but nevertheless there are systems that can be described to a good approximation as being closed. Some examples of such unitary operators are the well known Pauli matrices that describe logical operations in the quantum world. For example the Pauli matrix $X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ describes a NOT gate, because it transforms $|0\rangle \rightarrow |1\rangle$ and $|1\rangle \rightarrow |0\rangle$, thus it is also referred to as the *bit flip* matrix.

B.1.3 Measurements

Postulate A projective measurement is described by an *observable*, M, a Hermitian operator on the state space of the system being observed. The observable has a spectral decomposition,

$$M = \sum_m m P_m, \tag{B.3}$$

where P_m is the projector onto the eigenspace of M with eigenvalue m. The possible outcomes of the measurement correspond to the eigenvalues, m, of the observable. Upon measuring the state $|\psi\rangle$, the probability of getting result m is given by

$$p(m) = \langle\psi|P_m|\psi\rangle. \tag{B.4}$$

Given that outcome m occurred, the state of the quantum system immediately after the measurement is

$$\frac{P_m|\psi\rangle}{\sqrt{p(m)}} \tag{B.5}$$

In classical physics one can observe quantities like speed, energy, mass etc. without affecting the system. In quantum mechanics whenever one party measures the system he or she destroys it obtaining the desired measurement. This is known to be one of the basic differences between quantum and classical physics.

B.1.4 Composite systems

Postulate The state space of a composite physical system is the *tensor product* of the state spaces of the component physical systems. Moreover, if we have systems numbered 1 through n, and system number i is prepared in the state $|\psi_i\rangle$, then the joint state of the total system is $|\Psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle$.

B.2 Mixed states

As stated in Postulate 1, the state of a quantum system is described by a vector in a Hilbert space. However, in most practical situations, the preparation of a quantum system is not perfect, either because of limited resources or the presence of a noisy environment. The state then is no longer pure because of the noise and it should be described by means of a mixed state, also known as density operator.

B.2.1 Density operator

An imperfect preparation of the state of a quantum system implies that the quantum system becomes a mixture of several states $|\psi_i\rangle$ with different probabilities p_i. Then one uses the description of the *density operator* or *density matrix* ρ which is calculated over the *outer product* of each possible state weighted by its probability of occurrence:

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \tag{B.6}$$

This matrix representation of a system helps us to describe side-effects of noise in the channel.
Now we can state that a valid density operator has to satisfiy the two conditions:

1. *Trace condition*: $Tr(\rho) = 1$

   Having an ensemble of quantum states the following is true $Tr(\rho) = \sum_i p_i Tr(|\psi_i\rangle\langle\psi_i|) = \sum_i p_i = 1$

2. *Positivity condition*: ρ is a positive operator.

   Supposing $|\phi\rangle$ is an arbitrary vector in the state space $\langle\phi|\rho|\phi\rangle = \sum_i p_i \langle\phi|\psi_i\rangle\langle\psi_i|\phi\rangle = \sum_i p_i |\langle\phi|\psi\rangle|^2 \geq 0$

Remark that all postulates given in chapter B.1 can be reformulated in the form of the density matrix.

B.2.2 Pure and mixed states

Definition 7 *If we can write a state in the form $\rho = |\psi\rangle\langle\psi|$ then $Tr(\rho^2) = 1$ and we call the state* pure. *If $i > 1$ in* (B.6) *we call the state* mixed *and $Tr(\rho^2) < 1$.*

Let $|\psi\rangle$ be a state vector of our system. Then $Tr(\rho^2) = Tr(|\phi\rangle\langle\phi|\phi\rangle\langle\phi|) = Tr(|\phi\rangle\langle\phi|) = \langle\phi|\phi\rangle = 1$ where we used the normalization condition in equality two.
If the state is not pure we cannot write it in the state vector form. Instead we have: $\rho = \sum_i p_i \rho_i$. Moreover $Tr(\rho^2) = Tr\left((\sum_i p_i \rho_i)^2\right) = \sum_i p_i^2 Tr\left(\rho_i^2\right) < 1$, due to the linearity of the trace operation.

In the following we will refer to pure states by their state vector representation and to mixed states by the density matrix representation.

B.3 Entanglement and separability

The combination of the superposition principle with the tensor product structure leads to the appearance of entanglement. This is a very peculiar form of correlations, with no classical analog, that appear in the quantum states of composite systems. A key mathematical tool in the understanding of quantum entanglement is the *Schmidt decomposition.*

Definition 8 *Let $|\phi\rangle$ be a bipartite pure state of $\mathscr{H}_A \otimes \mathscr{H}_B$. Then we can represent the state in the Schmidt decomposition*

$$|\phi\rangle = \sum_{i=1}^{\substack{\min[dim(\mathscr{H}_A),\\ dim(\mathscr{H}_B)]}} \sqrt{\lambda_i} |\alpha_i\rangle \otimes |\beta_i\rangle \tag{B.7}$$

with $\lambda_i \geq 0$ being the Schmidt coefficients and $|\alpha_i\rangle$ and $|\beta_i\rangle$ orthonormal vectors in each space. The number of non-zero Schmidt coefficients is called the Schmidt rank of the state.

If we have a pure bipartite state with Schmidt rank one we call the state *product* or *separable* because we can write it as the product of two pure states in $\mathscr{H}_A$ and $\mathscr{H}_B$.

$$|\phi\rangle_{AB} = |\psi\rangle_A \otimes |\varphi\rangle_B \tag{B.8}$$

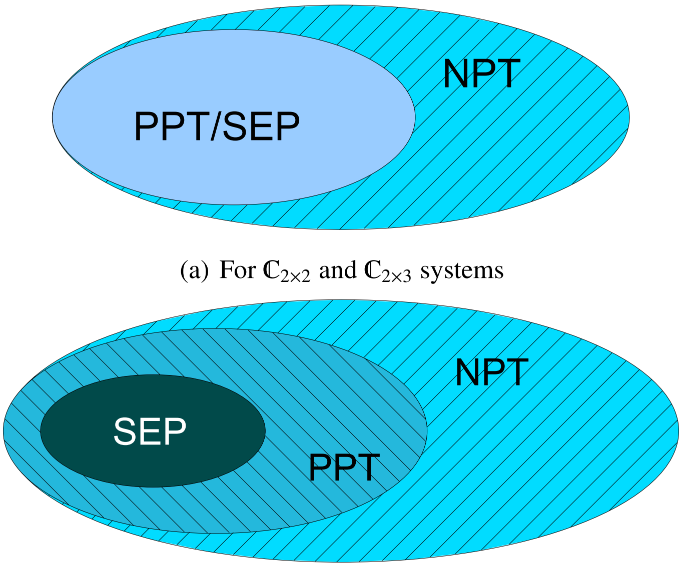
If the bipartite state has Schmidt rank greater than one we call the state *entangled* or *non-separable* because one is not able to write it as the tensor product of two pure states from each subspaces.
These definitions are generalized for mixed states as follows.

Definition 9 *Given a density matrix ρ_{AB} acting on $\mathscr{H}_A \otimes \mathscr{H}_B$. If we can write ρ_{AB} in the form*

$$\rho_{AB} = \sum_i \lambda_i \rho_i^A \otimes \rho_i^B \tag{B.9}$$

we call the state separable. *If it is impossible to write a state in the form (B.9) we call it* entangled.

It has been shown in [17] that positive eigenvalues of the partial transpose (PPT) of a state is a necessary condition for separability of a state. Moreover it is stated that in the cases $\mathbb{C}_{2\times 2}$ and $\mathbb{C}_{2\times 3}$ PPT is also a sufficient condition. This is illustrated in figure B.1 (a). But until now we do not have a useful tool for higher dimensional systems, i.e. $\mathbb{C}_{2\times k}$ with $k > 3$ and $\mathbb{C}_{i\times j}$ with $i, j \geq 3$ systems. Here the relation illustrated in figure B.1 (b) is valid.



(a) For $\mathbb{C}_{2\times 2}$ and $\mathbb{C}_{2\times 3}$ systems

(b) For $\mathbb{C}_{2\times k}$ with $k > 3$ and $\mathbb{C}_{i\times j}$ with $i, j \geq 3$ systems

Figure B.1: Relation of separability, PPT and entanglement

In chapter 3 we discuss another criterion for the separability of a system.

The most paradigmatic example of entangled pure states are the *Bell states*, that is the *maximally entangled states of two qubits*. An example of these states is:

$$|\Phi^+\rangle = (|00\rangle + |11\rangle)\frac{1}{\sqrt{2}} \tag{B.10}$$

This state represents the basic unit of bipartite entanglement and allows, for instance, quantum teleportation and secure quantum cryptography.

Entanglement distillability is a crucial question in the study of entangled states. Given an noisy entangled state ρ_{AB}, shared by two separated parties, we would like to know whether this state can be transformed into maximally entangled states of two qubits by local operations by the parties assisted by classical communication. Remarkably, there exist states that, depite being entangled, cannot be distilled into maximally entangled states. This phenomenon is called *bound entanglement* and gives a kind of irreversible form of entanglement.

Bibliography

[1] G , N. ; R , R. ; W , S.: Linking classical and quantum key agreement: Is there a classical analog to bound entanglement. In: *Algorithmica*, 2002, S. 309–559

[2] S , C. E.: A mathematical theory of communication, 1948, S. 623–656

[3] S , C. E.: Communication theory of secrecy systems. In: *Bell System Technical Journal* 28 (1949), S. 656–715

[4] V , G. S.: Cipher printing telegraph systems for secret secure key agreement in cryptography. In: *ETH dissertation No. 13138, Swiss Federal Institute of Technology* (1926), S. 109–115

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

[6] M , U. M. ; W , S.: Unconditionally secure key agreement and the intrinsic conditional information. In: *IEEE Transactions on Information Theory* 45 (1999), S. 499–514

[7] C ´ , I. ; K¨ , J.: Broadcast channels with confidential messages. In: *IEEE Transactions on Information Theory* 24 (1978), S. 339–348

[8] M , U. M.: The strong secret key rate of discrete random triples. In: *Communications and Cryptography: Two Sides of One Tapestry*, 1994, S. 271–285

[9] R , R. ; W , S.: New bounds in secret-key agreement: The gap between formation and secrecy extraction. In: *Proc. EUROCRYPT 2003 (Lecture notes in Computer Science)*, 2003, S. 562–577

[10] A , A. ; C , J. I. ; M , L.: Multipartite bound information exists and can be activated. In: *Physical Review Letters* 92 (2004), S. 107903

[11] C , D. ; P , S.: Classical analog of entanglement. In: *Physical Review Article* 65 (2002), Feb, Nr. 3, S. 032321

[12] M , L. ; A , A.: Multipartite secret correlations and bound information. In: *IEEE Transactions on Information Theory* 52 (2006), S. 4686

[13] M , L. ; W , A.: A non-distillability criterion for secret correlations. (2008)

[14] J , N. S. ; M , L.: Key distillation and the secret-bit fraction. In: *IEEE Transactions on Information Theory* 54 (2008), S. 680

[15] A , A. ; G , N. ; M , L.: From Bell's theorem to secure quantum key distribution. In: *Physical Review Letters* 97 (2006), S. 120405

[16] N , M. A. ; C , I. L.: *Quantum computation and quantum information*. Cambridge University Press, 2000

[17] H , M. ; H , P. ; H , R.: Separability of mixed states: necessary and sufficient conditions. In: *Physics Letters A* (1996)