Diploma Thesis

The Gap Between Intrinsic Information and the Secret-Key Rate

Juraj Skripsky

supervised by

Renato Renner
Prof. Dr. Ueli M. Maurer

Institute for Theoretical Computer Science
ETH Zürich, Switzerland

October 2002

Abstract

Perfectly secret message transmission can be realized with only partially secret and weakly correlated information shared by the parties as soon as this information allows for the extraction of information-theoretically secret bits. The rate at which such key bits can be generated is called *secret-key rate* and depends on the distribution modeling the parties', including the adversary's, knowledge. The best known upper bound for this rate has been the *intrinsic information.* We show the first distribution for which it is known that this bound is *not tight.* Making things even worse, the gap between the secret-key rate and intrinsic information can be arbitrarily large. Furthermore, we present new evidence for the existence of bound information, a quantity arising in a special case of non-tightness of the intrinsic information bound. For a distribution having a secret-key rate of zero, the remaining positive intrinsic information is called bound information. Its existence would disqualify the intrinsic information measure as a reliable indicator of the possibility of secret-key agreement.

When Eve is given some sort of *side information*, the secret-key rate decreases in general. We extend the model of secret-key agreement to an arbitrary number of participating parties in a natural way. The new model then allows us to describe and capture the effect of side information in a surprisingly simple yet accurate way. Using the measures resulting from its analysis, we derive a *new upper bound* on the secret-key rate (in the classical model), which is strictly stronger than the intrinsic information bound.

Contents

Chapter 1

Introduction

In modern cryptography there are mainly two security paradigms, namely *computational* and *information-theoretic* security. The latter is also called unconditional security. Computational security is based on the assumed hardness of certain computational problems (e.g., the integer-factoring or discrete-logarithm problems). However, since a computationally sufficiently powerful adversary can solve any computational problem, hence break any such system, and because no useful general lower bounds are known in complexity theory, computational security is always conditional and, in addition to this, endangered by the progress in the theory of efficient algorithms as well as in hardware engineering (e.g., quantum computing). Information-theoretic security on the other hand is based on probability theory and on the fact that an adversary's information is limited. Such a limitation can for instance come from noise in communication channels or from the laws of quantum mechanics.

Many different cryptographic settings based on noisy channels have been described and analyzed. Examples are Wyner's wire-tap channel [3], Csiszár and Körner's broadcast channel [5], or Maurer's model of key agreement from joint randomness [6], [8].

1.1 Contributions

We show the first distribution for which we know a proof for the fact that its secret key rate is strictly smaller than the corresponding intrinsic information. Knowing that the instrinic information bound is not tight, we demonstrate that the resulting gap can be arbitrarily large by generalizing the previous distribution. The resulting class of distributions can also be used to make the secret key rate arbitrarily small while keeping the intrinsic information constant at the same time. Another slightly modified version of the distribution is shown, which is believed to contain bound information.

We extend the secret-key agreement model to arbitrarily many parties and closely analyze the secret-key rates involving four parties. We use the resulting lower and upper bounds to prove a new upper bound for the secret key rate which is strictly stronger than the instrinsic information bound. An interesting special version of the new bound is used to prove the non-tightness of the intrinsic information bound for one of the distributions shown.

1.2 Organization of the Thesis

The thesis is organized as follows: In Chapter 2, we introduce the model of information-theoretically secure key agreement and its most interesting quantity, the secret-key rate. A gentle, informal derivation of a number of previously known upper bounds for the secret key rate, including the intrinsic information, follows in Chapter 3. In Chapter 4, we prove the non-tightness of the intrinsic information bound and present a distribution which is believed to contain bound information. Finally, in Chapter 5, we analyze the effect of side information given to Eve, leading to the extension of the secret-key agreement model to $m + n$ parties and, in the end, to a new upper bound for the secret key rate.

Chapter 2

The Model

Shannon [2] has defined an encryption scheme to be *perfectly secret* if the ciphertext does not reveal any information about the encrypted message. Such a system is unconditionally secure with respect to a ciphertext-only attack; in particular, an exhaustive search over the key space is no help in finding the cleartext. Shannon proved in the same paper that, unfortunately, such a high level of secrecy has its price: it is, roughly speaking, only possible between parties who share an information-theoretically secure key that is at least as long as the message to be encrypted. The so-called *one-time pad* [1], a computationally very simple encryption that just bit-wisely XORs the key to the message, on the other hand shows that perfectly secure encryption is possible between parties who do share a key of that length. Since we assume that insecure channels are always available, the one-time pad reduces the problem of information-theoretically secure encryption to information-theoretically secure *key agreement*, which we will consider in the following.

2.1 Secret-Key Agreement

We assume that two parties Alice and Bob, who are connected by an authentic but otherwise completely insecure channel, are willing to generate a secret key. More precisely, Alice and Bob want to compute, after some rounds of communication (where the random variable C summarizes the communication performed over the public channel), strings S_A and S_B, respectively, with the property that they are most likely both equal to a uniformly distributed string S about which the adversary Eve has virtually no information. More precisely,

$$H(S_A|C) = H(S_B|C) = 0, \tag{2.1}$$

$$P[S_A = S_B = S] \geq 1 - \epsilon, \tag{2.2}$$

$$H(S) = \log_2 |\mathcal{S}|, \tag{2.3}$$

and

$$I(S;C) \leq \epsilon \tag{2.4}$$

(where $\mathcal{S}$ is the range of S and $|\mathcal{S}|$ is its cardinality) should hold for some small ϵ. Note that these security conditions are information-theoretic (also called unconditional): Even an adversary with unlimited computer power must be unable to obtain useful information. In contrast to this, the Diffie-Hellman protocol [4], for instance, achieves the goal of key agreement by insecure communication only with respect to computationally bounded adversaries.

It is a straight-forward generalization of Shannon's impossibility result that information-theoretic secrecy cannot be generated in this setting, i.e., from authenticity only: Public-key systems are never unconditionally secure. Hence we have to assume an additional structure in the initial setting, for instance, some pieces of information given to Alice and Bob (and also Eve).

2.2 Correlated Random Variables

The general case where the information given to the three parties initially consists of the outcomes of some random experiment has been studied intensively [6], [8], [9]. Here, it is assumed that Alice, Bob, and Eve have access to realizations of random variables X, Y, and Z, respectively, jointly distributed according to P_{XYZ}. Conditions (2.1) and (2.4) now have to be replaced by modified versions introducing X, Y, and Z:

$$H(S_A|XC) = H(S_B|YC) = 0, \tag{2.5}$$

$$I(S;ZC) \leq \epsilon. \tag{2.6}$$

It was shown that if the setting is modified this way, then secret-key agreement is often possible.

2.3 Secret-Key Rate

Shannon's results about perfect secrecy have shown us that the key used for encryption has to be at least as long as the message and that it cannot be reused for subsequent messages. Therefore in any realistic scenario, very long secret keys will have to be provided. In the key agreement model, the key length equals $H(S)$ and never exceeds neither $H(X)$ nor $H(Y)$. So, in order to generate long secret keys, we need to have random variables with high entropy.

Nevertheless, it seems reasonable to restrict further analyses to "simple" distributions. Motivated by models such as discrete memoryless channels and channels previously considered in information theory, we make the following natural assumption: The random variables used for key agreement are created by *repeating the random experiment* generating XYZ many times independently. Alice, Bob and Eve receive $X^N = [X_1, \ldots, X_N]$, $Y^N = [Y_1, \ldots, Y_N]$ and $Z^N = [Z_1, \ldots, Z_N]$, respectively, where

$$P_{X^N Y^N Z^N} = (P_{XYZ})^N.$$

For such a scenario of independent repetitions of a random experiment, the quantity that appears to be of most interest is the *secret-key rate*. It is a measure for the efficiency of the optimum key agreement protocol for a given distribution P_{XYZ} and relates the length of the resulting key to the number of random experiments needed to generate it.

Definition 1. [6], [9] The *secret-key rate of* X *and* Y *with respect to* Z, denoted $S(X;Y||Z)$, is the maximum rate R at which Alice and Bob can agree on a secret key S about which Eve has virtually no information, i.e. it is the maximum R, so that for every $\epsilon > 0$ there exists a protocol for sufficiently large N satisfying (2.2), (2.3), (2.5) and (2.6) with X, Y and Z replaced by X^N, Y^N and Z^N, respectively, achieving

$$\frac{1}{N}H(S) \geq R - \epsilon. \tag{2.7}$$

○

In general, the secret-key rate is very difficult to calculate, as it involves finding the best one of all (infinitely many) possible protocols. The best approach is to approximate the secret-key rate by finding good lower and upper bounds. It is the main focus of this thesis to analyze and discuss previously known bounds, as well as derive better ones.

2.3.1 A Lower Bound

The following theorem states a nontrivial lower bound on the secret-key rate. It was proven by Csiszár and Körner [5] (in a slightly different context) and later again by Wolf [9]. If it is either the case that Eve has less information about Y than Alice or, by symmetry, less information about X than Bob, then such a difference of information can be exploited.

Theorem 2. [5], [9] *The secret-key rate of* X *and* Y *with respect to* Z *is lower bounded by*

$$S(X;Y||Z) \geq max\{I(X;Y) - I(X;Z), I(X;Y) - I(Y;Z)\}.$$

2.4 The Satellite Scenario

The following realistic special scenario was proposed in [6] and completely analyzed in [7]. Assume that a satellite sends out random bits at very low signal power and that Alice, Bob and Eve receive these bits over independent binary-symmetric channels with error probabilities α, β, and ϵ, respectively. In general, we have to assume that Eve may have a better antenna than the legitimate partners and hence a possibly substantially lower error rate. It is a surprising fact that secret-key agreement is always possible in this scenario (unless Eve has a noiseless access to the satellite bits or either Alice or Bob obtains no information at all about these bits).

2.4.1 Phases of Secret-Key Agreement Protocols

We sketch a protocol for secret-key agreement in the satellite scenario. Such a protocol is often interpreted as consisting of three phases [9], namely *advantage distillation*, *information reconciliation*, and *privacy amplification.*

Advantage Distillation. The adversary possibly has an initial advantage over the legitimate partners in terms of the error probabilities. Coping with this situation is the objective of this phase. Alice and Bob create an advantage over the opponent by exchanging information about their bits over the public channel with the goal of identifying bits that are correct with a higher probability than others. Although Alice's and Bob's bit error probabilities can be much worse on the average than Eve's, they will be much better when the average is taken over the more reliable bits only.

Information Reconciliation. Alice and Bob do not generally share a mutual string after the previous phase. In order to guarantee that the two strings are identical, we apply (possibly interactive) error-correction techniques.

Privacy Amplification. Alice and Bob have agreed on a mutual string about which Eve has a possibly considerable amount of information consisting of both a priori knowledge but also of information leaked during the previous two phases. Therefore, the partially-secret string must be transformed into a (shorter) highly-secret string.

Chapter 3

Upper Bounds for the Secret-Key Rate

In this chapter, we will introduce a few upper bounds for the secret-key rate, starting with the simple ones and concluding with the best direct bound[1] known so far, namely the *intrinsic information.*

Each bound will be introduced based on intuition. References to the respective proofs will be given. As a next step, a distribution will be shown and analyzed with respect to its secret-key rate and the given formula will be evaluated. The distributions are chosen in a way that they not only demonstrate the non-tightness of the bound under scrutiny, but also nicely illustrate the reason for its failure to give the optimum result. This insight will then lead to a refinement of the bound.

3.1 Simple Bounds

A first trivial bound. As the secret-key rate depends on the amount of information Alice and Bob have in common, the following bound seems intuitive:

$$S(X;Y||Z) \leq I(X;Y) \tag{3.1}$$

Its proof is not completely obvious and can be found in [6].

Now consider the distribution (to be normalized) shown in the table below. The table has to be read as follows: Each number in the table body represents the probability that the triple of random variables X, Y and Z takes on the corresponding values (with Z's value given in parentheses). For instance, the entry "**1** (0)" for $X = Y = 0$ means $P_{XYZ}(0,0,0) = \frac{1}{2}$ (normalized) (and $P_{XYZ}(0,0,1) = 0$). The fields corresponding to XYZ triples having probability zero are left blank.

X Y (Z)	0	1
0	**1** (0)	
1		**1** (1)

[1] A direct bound's evaluation does not involve calculating another secret-key rate.

Bound $I(X;Y)$ evaluates to 1. This might suggest that Alice and Bob can transfer up to one bit per random experiment in perfect secrecy. This is of course not the case, as the value of Z is always identical to both X and Y. Eve has complete knowledge of Alice's and Bob's information. Hence secret-key agreement is not possible. Its rate is 0.

Taking Z into account. The relationship of Z and the information shared by X and Y is completely ignored by $I(X;Y)$, making it a bound which is far too optimistic. Evidently, a good upper bound has to take the information available to Eve into account. A natural way to do this is to condition the mutual information of X and Y on Z. As was proven in [6], the conditional mutual information is indeed an upper bound:

$$S(X;Y||Z) \leq I(X;Y|Z) \tag{3.2}$$

This could also be written as $S(X;Y||Z) \leq I(XZ;YZ) - H(Z)$. Roughly speaking, this translates to the statement that even under the assumption that Alice and Bob could learn Z, the remaining shared information is an upper limit on the information they can share in secrecy. Applying $I(X;Y|Z)$ to the distribution above yields 0, the result we were hoping for.

Unfortunately, (3.2) introduces a problem of its own. Let us analyze the following distribution:

X Y (Z)	0	1
0	**1** (0)	**1** (1)
1	**1** (1)	**1** (0)

Bound $I(X;Y|Z)$ evaluates to 1 for this distribution. But looking at the probabilities only with respect to the values of X and Y, it should be clear that X and Y are statistically independent. Thus the secret-key rate is 0.

Taking a second look at the distribution, we see that in fact all three random variables are pairwise statistically independent. Nevertheless they are closely related and any two of them uniquely determine the third one as $Z = X \oplus Y$ holds. This correlation between X and Y brought in by Z is what is being picked up by $I(X;Y|Z)$. Alice and Bob are not able to take advantage of this correlation as they do not know the value of Z.

Combining both bounds. While our first bound captures the structure of this distribution correctly, bound (3.2) fails to detect the independence of X and Y. This problem can be seen as a consequence of an unexpected property of mutual information, namely that $I(X;Y|Z) > I(X;Y)$ is possible. A simple yet effective solution to this problem is to combine both previous upper bounds into one [6]:

$$S(X;Y||Z) \leq \min\{I(X;Y), I(X;Y|Z)\} \tag{3.3}$$

This gives us the best of both worlds and handles the distributions shown above correctly. The bound is tight in those cases.

3.2 Intrinsic Conditional Information

When we arrange both distributions from Section 3.1 side-by-side in one single larger distribution as shown below, a interesting new problem arises. We end up in a situation where secret-key agreement is not possible, yet both $I(X;Y)$ and $I(X;Y|Z)$ (and hence also the minimum of the two) are positive.

X Y (Z)	0	1	2	3
0	**1** (0)	**1** (1)		
1	**1** (1)	**1** (0)		
2			**2** (2)	
3				**2** (3)

Bound $\min\{I(X;Y), I(X;Y|Z)\}$ evaluates to $\frac{1}{2}$. It is destined to fail for this distribution, which is a mixture of the two "pure" ones studied in Section 3.1. Bound $I(X;Y)$ handles the "subdistribution" in the upper left quadrant of the table correctly but fails in the lower right quadrant. For $I(X;Y|Z)$, it is the other way round. Both do fail somewhere. What we would like to do is to apply both $I(X;Y)$ and $I(X;Y|Z)$ at the same time, the former in the upper left quadrant and the latter in the lower right quadrant. Of course we cannot just arbitrarily split a distribution into components and treat each one of them separately. But there is a different approach which achieves something very similar without destroying the structure of the distribution.

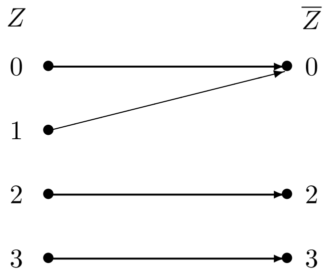
Let us analyze the distribution from Eve's point of view. For every single random experiment, her random variable Z takes on one symbol out of four. Whenever that symbol is either 2 or 3, she has complete knowledge of the symbol received by Alice and Bob. For the values 0 and 1, the situation is slightly more complicated. Eve knows that Alice's and Bob's symbols are restricted to 0 and 1. As X and Y are statistically independent when restricted to the range $\{0, 1\}$, the distinction between 0 and 1 is useless to Alice and Bob. Consequently, the distinction between 0 and 1 is irrelevant to Eve as well. Eve can replace every occurence of either 0 or 1 by the single symbol 0 without loss of essential information (and so could Alice and Bob). This can be interpreted as a (completely local) creation of a new random variable $\overline{Z}$. The net result is that Eve has transformed the distribution P_{XYZ} into $P_{XY\overline{Z}}$:

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

Finally, applying $I(X;Y|Z)$ (or rather $I(X;Y|\overline{Z})$) to $P_{XY\overline{Z}}$ yields the value 0.

How can we generalize the idea, which the above analysis is based on, to arbitrary distributions? What we are looking for is a concept for the transformation of a random variable. We would like this transformation to be simple to describe and analyze yet universal enough to allow for various symbol replacement strategies. There is a concept in information theory which is very well suited for this purpose: the channel.

The description of a channel consists of the conditional probability distribution $P_{\overline{Z}|Z}$, where $P_{\overline{Z}|Z}(\overline{z}, z) = \alpha$ means that the symbol z is transformed into symbol $\overline{z}$ with probability α. The channel corresponding to the replacement strategy used in the example above would look like this (all transition probabilities are 1):



Upon receiving her symbol, Eve sends it through such a channel and replaces it with this channel's output. There is nothing that prevents Eve from locally generating a new random this way. Given a distribution P_{XYZ}, one has to find the optimum replacement strategy or its corresponding channel. The optimum channel makes the resulting conditional information as small as possible. Therefore the search for this channel has to be guided by a minimization process. This immediately leads us to a new information measure:

Definition 3. [8] For a distribution P_{XYZ}, the *intrinsic conditional information between X and Y, when given Z*, denoted by $I(X;Y\!\downarrow\! Z)$, is

$$I(X;Y\!\downarrow\! Z) := \min_{XY \to Z \to \overline{Z}} I(X;Y|\overline{Z})$$

(where the minimum is taken over all Markov chains $XY \to Z \to \overline{Z}$[2]). ○

In contrast to $I(X;Y|Z)$, the intrinsic information $I(X;Y\downarrow Z)$ measures only the remaining conditional mutual information between X and Y (possibly reduced by giving Z), but not the additional information between X and Y brought in by Z. As the considerations made in the example above already suggest, the intrinsic information is an upper bound to the secret-key rate.

Theorem 4. [8] *The secret-key rate of X and Y with respect to Z is upper bounded by*

$$S(X;Y||Z) \leq I(X;Y\!\downarrow\! Z).$$

[2] This restriction makes sure that the channel $Z \to \overline{Z}$ does not depend in any way on X or Y, i.e. that $I(XY;\overline{Z}|Z) = 0$ holds.

Chapter 4

The Bound Given by Intrinsic Information Is not Tight

Intrinsic information is the best upper bound for the secret-key rate previously known. Supported by some evidence, it was believed that this bound is not tight in general, i.e. that $S(X;Y||Z) < I(X;Y\!\downarrow\! Z)$ is possible. But no proof for this claim was known. We will show a distribution for which not only we have a proof that there is a gap, but for which we can also give a simple explanation for how this gap comes about. Generalizing this distribution in a natural way, we will illustrate that the situation is even worse than previously thought. The gap not only exists but it can even be arbitrarily large.

Finally, we take a look at a very special case of non-tightness, where the secret-key rate is zero while the intrinsic information remains positive. What the intrinsic information measures in such a hypothetical scenario is called “bound information” [10] as it cannot be used for secret-key agreement. The existence of such an effect would be rather unfortunate, as it would disqualify the bound as a reliable indicator for the possibility or impossibility of secret-key agreement. We will construct a distribution (again based on the first one) for which it seems very plausible that it suffers from this effect — or rather defect.

4.1 A Distribution with $S(X;Y||Z) < I(X;Y\!\downarrow\! Z)$

At first glance, the following distribution looks very similar to the one in Section 3.2. But there is a subtle difference. In the fields of the lower right quadrant, Eve now receives the symbols 0 and 1 (instead of 2 and 3). As a consequence, Eve is no longer able to tell whether we are in the upper left or in the lower right quadrant. As we will see later, this has interesting implications for the intrinsic information of the distribution.

X Y (Z)	0	1	2	3
0	**1** (0)	**1** (1)		
1	**1** (1)	**1** (0)		
2			**2** (0)	
3				**2** (1)

Let us first take a look at the secret-key rate. As already mentioned, Eve has lost the ability to distinguish between upper left and lower right quadrant. In fact, it seems that this is the only bit of information unavailable to Eve, but at the same time available to both Alice and Bob. This suggests a secret-key rate of 1.

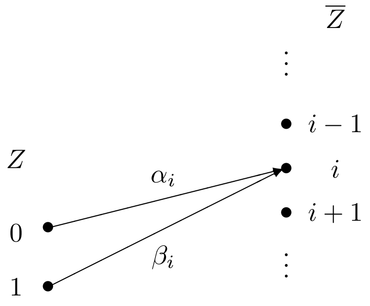
However, we are facing a dilemma when searching for the optimum channel which minimizes the conditional information. In one region of the distribution, we want to "blur" the correlation brought in by Z's symbols by mixing them. In the other region, we want to keep the symbols unaltered. But both regions contain exactly the same symbols! Since Eve never knows which region we are in, she does not know which channel she should use. Mixing the symbols means destroying the "artifical" correlation in the upper left corner at the cost of losing the ability to distinguish the symbols in the lower right corner. Keeping the symbols unchanged preserves both the ability to distinguish as well as the "artifical" correlation. Doing the right thing in one region inevitably means doing the wrong thing in the other. So Eve makes a mistake no matter what channel she uses.

The conditional information $I(X;Y|\overline{Z})$ evaluates to 1.5 for the output of both the mixing and the preserving channel.

4.1.1 Proof of $I(X;Y\downarrow Z) = 1.5$

The distribution is "balanced" in such a way that optimizing the channel for one region has an perfectly compensating effect for the other one. As we will show, $I(X;Y|\overline{Z})$ equals 1.5 *for all possible channels* $Z \rightarrow \overline{Z}$!

Proof. Consider an arbitrary channel $Z \rightarrow \overline{Z}$ with output symbols $1, 2, \ldots, n$ and transition probabilities $P_{\overline{Z}|Z}(i, 0) = \alpha_i$ and $P_{\overline{Z}|Z}(i, 1) = \beta_i$ (where $1 \leq i \leq n$ and $\sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \beta_i = 1$):



We are interested in the conditional information which results from sending Z through the channel above:

$$I(X;Y|\overline{Z}) = \sum_{i=1}^{n} (P_{\overline{Z}}(i) \cdot I(X;Y|\overline{Z}=i))$$

In order to calculate $I(X;Y|\overline{Z}=i)$, we need the distribution $P_{XY|\overline{Z}=i}$:

X Y	0	1	2	3
0	α_i	β_i		
1	β_i	α_i		
2			$2\alpha_i$	
3				$2\beta_i$

$$\begin{aligned}
I(X;Y|\overline{Z}=i) &= H(X|\overline{Z}=i) - H(X|Y,\overline{Z}=i) \\
&= H(\left[\frac{\alpha_i+\beta_i}{4s}, \frac{\alpha_i+\beta_i}{4s}, \frac{2\alpha_i}{4s}, \frac{2\beta_i}{4s}\right]) - \frac{1}{2}h(\frac{\alpha_i}{s}) \\
&= \left(1 + \frac{1}{2}H(\left[\frac{\alpha_i+\beta_i}{2s}, \frac{\alpha_i+\beta_i}{2s}\right]) + \frac{1}{2}H(\left[\frac{2\alpha_i}{2s}, \frac{2\beta_i}{2s}\right])\right) - \frac{1}{2}h(\frac{\alpha_i}{s}) \\
&= \left(1.5 + \frac{1}{2}h(\frac{\alpha_i}{s})\right) - \frac{1}{2}h(\frac{\alpha_i}{s}) \\
&= 1.5
\end{aligned}$$

where $s := \alpha_i + \beta_i$ and $h(p)$ is the binary entropy function[1]. Note that $I(X;Y|\overline{Z}=i)$ is independent of i. Therefore we have

$$\begin{aligned}
I(X;Y|\overline{Z}) &= \sum_{i=1}^{n} (P_{\overline{Z}}(i) \cdot I(X;Y|\overline{Z}=i)) \\
&= 1.5 \cdot \sum_{i=1}^{n} P_{\overline{Z}}(i) \\
&= 1.5
\end{aligned}$$

as the sum of $P_{\overline{Z}}(i)$ over all symbols of $\overline{Z}$ is 1.
When $I(X;Y|\overline{Z})$ equals 1.5 for all possible channels $Z \to \overline{Z}$, then we have

$$I(X;Y\downarrow Z) = \min_{XY\to Z\to\overline{Z}} I(X;Y|\overline{Z}) \;= 1.5$$

□

[1] The binary entropy function is defined as $h(p) := H([p, 1-p])$.

4.1.2 Proof of $S(X;Y||Z) = 1.0$

We first prove that $S(X;Y||Z) \geq 1$ holds by describing a protocol which achieves this rate. Subsequently we will prove that this rate cannot be exceeded, i.e. $S(X;Y||Z) \leq 1$. Both inequalities together then imply $S(X;Y||Z) = 1$.

Proof of $S(X;Y||Z) \geq 1$. Alice and Bob can agree on 1 bit per random experiment using following simple transformation. They both locally generate a new random variable by dividing the received value by two, i.e. $X' := \lfloor X/2 \rfloor$ and $Y' := \lfloor Y/2 \rfloor$. The shared secret key is then $S = X'^N = Y'^N$. No communication is needed. It can easily be verified that the protocol satisfies all requirements of the secret-key agreement model. □

Proof of $S(X;Y||Z) \leq 1$. The non-existence of a protocol having some property is much harder to prove than its existence. We need to make use of following inequality[2]:

$$S(X;Y||Z) \leq S(X;Y||[Z,U]) + H(U|Z). \tag{4.1}$$

It gives a limit on the difference between the secret-key rates $S(X;Y||Z)$ and $S(X;Y||[Z,U])$, i.e. where Eve is given extra information in the form of a random variable U. What it says is that supplying Eve with n bits of additional information will never reduce the secret-key rate by more than n bits. By defining $U := \lfloor X/2 \rfloor$ we get

$$\begin{aligned} S(X;Y||Z) &\leq S(X;Y||[Z,U]) + H(U|Z) \\ &\leq I(X;Y\downarrow[Z,U]) + H(U|Z) \\ &= 1 \end{aligned}$$

where the minimum 0 for the intrinsic information formula can be reached by using a channel equivalent to the one used in Section 3.2. □

[2] This new upper bound for the secret-key rate will be proven in Section 5.3.3 as Theorem 14.

4.2 Stretching the Gap between $S(X;Y||Z)$ and $I(X;Y\!\downarrow\! Z)$

The value given by intrinsic information can be slightly off when compared to the secret-key rate. It would be interesting to know how substantial this slight difference can be in the worst case. Unfortunately, it can be arbitrarily large, as the following class of distributions proves.

It is possible to widen the gap by generalizing the distribution which allowed us to prove its existence in the first place. This is done by increasing the number of symbols Eve receives to n (and adjusting Alice's and Bob's symbol count accordingly to $2n$), while retaining the basic structure of the distribution. The upper left quadrant contains a block of equal probabilities. Its inner structure is partly accessible to Eve with lines of equal symbols running diagonally through it, yet completely unavailable to Alice and Bob. The lower right quadrant is filled only at its diagonal with distinct symbols for all three parties.

X Y (Z)	1	2	$\cdots$	n	$n+1$	$n+2$	$\cdots$	$2n$
1	**1** (1)	**1** (2)	$\cdots$	**1**(n)				
2	**1** (2)	**1** (1)	**1** (2)	$\cdots$				
$\cdots$	$\cdots$	**1** (2)	**1** (1)	**1** (2)				
n	**1**(n)	$\cdots$	**1** (2)	**1** (1)				
$n+1$					**n** (1)			
$n+2$						**n** (2)		
$\cdots$							$\cdots$	
$2n$								**n**(n)

The resulting class of distributions $(P_{XYZ})_{(n)}$ shares the interesting features of its original. The secret-key rate is still limited to 1 (again achievable by quadrant detection). And the intrinsic information still picks up the correlation brought in by Z in the upper left. Increasing n increases the intrinsic information while having no effect on the secret-key rate:

$$\begin{aligned} I(X_{(n)};Y_{(n)}\!\downarrow\! Z_{(n)}) &= 1+\frac{1}{2}\log_2 n, \\ S(X_{(n)};Y_{(n)}||Z_{(n)}) &= 1. \end{aligned}$$

With a subtle modification, we can keep $I(X_{(n)};Y_{(n)}\downarrow Z_{(n)})$ constant and make $S(X_{(n)};Y_{(n)}||Z_{(n)})$ arbitrarily small, reversing the situation. We add a blank symbol $\triangle$ to the distribution, which is received by all three parties with probability $p(n)$. Everytime the blank symbol occurs, no information can be transferred between Alice and Bob, thereby lowering the secret-key rate. The presence of the new symbol scales down both $I(X_{(n)};Y_{(n)}\downarrow Z_{(n)})$ and $S(X_{(n)};Y_{(n)}||Z_{(n)})$ by the same factor. A appropriately chosen $p(n)$ can now keep $I(X_{(n)};Y_{(n)}\!\downarrow\! Z_{(n)})$ constant and make $S(X_{(n)};Y_{(n)}||Z_{(n)})$ arbitrarily small.

4.3 A Distribution with Bound Information (most probably)

A special case of non-tightness of intrinsic information, that has especially significant implications is *bound information* [10]. Whenever the secret-key rate is zero for a distribution, while its intrinsic information is positive, the difference of this two measures is called bound information as it cannot be used for secret-key agreement. This means we cannot depend on intrinsic information to reliably classify distributions with respect to the possibility and impossibility of secret-key agreement.

We present a distribution along with an informal argument of why it seems very plausible that secret-key agreement is not possible (while it can be shown that the intrinsic information is positive). Again starting with the distribution from Section 3.2 and adding positive probabilities, we arrive at the following candidate for a distribution containing bound information:

X Y (Z)	0	1	2	3
0	**1** (0)	**1** (1)	**2** (2)	**2** (3)
1	**1** (1)	**1** (0)	**2** (4)	**2** (5)
2	**2** (6)	**2** (7)	**2** (0)	
3	**2** (8)	**2** (9)		**2** (1)

Whenever the outcome of the random experiment lies in one of the newly filled quadrants, the received symbols are useless to Alice and Bob, because Eve has an important advantage here. She has perfect knowledge of the values of both X and Y. All that is left are the fields which resemble the original distribution. Transferring one bit by distinguishing between upper left and lower right quadrant is the only hope for Alice and Bob. But this is no longer possible, because reducing X and Y to the new random variables $\widetilde{X} := \lfloor X/2 \rfloor$ and $\widetilde{Y} := \lfloor Y/2 \rfloor$ results in a scenario where the intrinsic information and hence the secret-key rate is 0.

$\widetilde{X}$ $\widetilde{Y}$ (Z)	0	1
0	**4** (0/1)	**8** (2–5)
1	**8** (6–9)	**4** (0/1)

In Section 5.3, we will derive a new upper bound for $S(X;Y||Z)$ which gives further support to this argument.

4.4 Discussion

In the light of its non-tightness, it is worth discussing the quality of the intrinsic information as an upper bound for the secret-key rate. We know that there is a gap between $I(X;Y\!\downarrow\!Z)$ and $S(X;Y||Z)$ in general. But despite its shortcomings, intrinsic information is a natural choice for the analysis of the secret-key rate. It is important not to consider it a tailer-made measure for the secret-key rate which just happens to be “not quite good enough” and needs to be optimized. Instead, it is a more general concept which captures the essential properties for a wide range of distributions in a simple yet accurate way.

Another aspect of intrinsic information, which underscores its importance, is the fact that it is a lower bound on what is called “information of formation”, the amount of secret key bits required to generate a certain distribution by public discussion. Inverting the process of secret-key agreement, information of formation is the rate at which secret bits are required to synthesize a distribution which is, in terms of the provided privacy, at least as good as P_{XYZ} from Alice’s and Bob’s point of view.

Chapter 5

The Effect of Adversarial Side Information

How does the secret-key rate change when Eve is given side information in form of an additional random variable U (which is jointly distributed with X, Y and Z)? Inequality (4.1) already gave us a first quantitative answer to this question by measuring how much information Eve gains at most when it is given the extra variable. As we will show, this upper bound can be improved.

5.1 Alice's and Bob's Loss vs. Eve's Gain

Inequality (4.1) relates the two secret-key rates $S(X;Y||Z)$ and $S(X;Y||[Z,U])$ by measuring the amount of additional information available to Eve when she receives the value of U. This is a rather pessimistic analysis as it implicitly assumes that Eve can fully exploit that information in order to reduce the rate. But clearly only those parts of information contained in U (or $[Z,U]$), which are also known to both Alice and Bob, will result in a change to the secret-key rate[1].

So instead of looking at the advantage Eve gains through U in the worst case (from Alice's and Bob's perspective), we take a closer look at Alice's and Bob's loss caused by U. The idea is to conceptually split U into two parts. While one part will contain precious information used by Alice and Bob during key agreement, the other one will consist solely of information unrelated to X and Y and thereby completely useless to Eve. By properly relating U to X and Y, it is possible to isolate the part harmful to Alice and Bob. Motivated by these considerations, we will derive a new upper bound for the secret-key rate by informal argumentation. A formal proof follows.

A bound measuring the loss suffered by Alice and Bob should always provide us with an rate equal to or lower than (4.1) for the same U. As (4.1) can be derived from the new bound, the latter is indeed a strict improvement.

[1] E.g. for a U independent of X, Y and Z (i.e. $I(U;XYZ) = 0$), the secret-key rate remains the same not matter if Eve knows U or not.

5.2 Extending Secret-Key Agreement to $m + n$ Parties

The model of secret-key agreement can easily be extended to arbitrarily many parties. Besides being interesting in its own right, we introduce this extension because it will later help us to capture the effect of side information in a surprisingly simple and intuitive way.

We let the number of both the parties seeking key agreement as well as their adversaries be freely chosen. The random variables X, Y and Z are replaced by X_1 through X_m and Z_1 through Z_n. All of the m parties have access to a public broadcast channel. By communicating over this channel, they try to generate a secret key based on their "personal" random variables. We want this key to be identical for all m parties and perfectly secret with respect to the n separately operating adversaries. The knowledge of each adversary is restricted to his random variable and the communication transcript C. This leads to the following definition for the extended secret-key rate.

Definition 5. The *secret-key rate of* X_1 *through* X_m *with respect to* Z_1 *through* Z_n, denoted $S(X_1; X_2; \ldots; X_m || Z_1 || Z_2 || \ldots || Z_n)$, is the maximum rate R at which m parties can agree on a secret key S about which no single one of the n adversaries has any information, i.e. it is the maximum R, so that for every $\epsilon > 0$ there exists a protocol for sufficiently large N satisfying (2.3),

$$H(S_i | X_i C) = 0,$$

$$P[S_1 = \ldots = S_m = S] \geq 1 - \epsilon,$$

and

$$I(S; Z_j C) \leq \epsilon \tag{5.1}$$

(where $1 \leq i \leq m$ and $1 \leq j \leq n$) and achieving (2.7). ○

Note that $I(S; Z_i) \leq \epsilon$ for all i is a weaker requirement than $I(S; [Z_1, \ldots, Z_n]) \leq \epsilon$. Accordingly, we have

$$S(X_1; \ldots; X_m || Z_1 || \cdots || Z_n) \geq S(X_1; \ldots; X_m || [Z_1, \ldots, Z_n]), \tag{5.2}$$

i.e. the rate of secret-key agreement subject to n separate adversaries is in general larger than the rate involving only one adversary knowing the values of all Z_j.

An equally intuitive relationship between secret-key rates involving a fixed set of agreement seeking parties is

$$\begin{aligned} S(X_1; \ldots; X_m) \geq S(X_1; \ldots; X_m || Z_1) \geq \cdots \\ \cdots \geq S(X_1; \ldots; X_m || Z_1 || \cdots || Z_n) \end{aligned} \tag{5.3}$$

as taking an adversary out of the picture never reduces the secret-key rate ((5.1) for $1 \leq j \leq t$ always implies (5.1) for $1 \leq j \leq t - 1$).

Finally, it is worth mentioning that while the extension of the secret-key rate to more parties follows in a quite straight-forward manner, this does not apply to its lower and upper bounds. In particular, it is not clear how or if the concept of intrinsic information can be generalized to $m + n$ or even only $3 + 1$ or $2 + 2$ parties.

5.2.1 Analysis of $S(X;Y||U||Z)$

The following simple upper bound follows directly from (5.3):

$$S(X;Y||U||Z) \leq \min\{S(X;Y||U), S(X;Y||Z)\}$$

As we have already seen in (5.2), the conspiracy of two adversaries, able to exploit their collective knowledge, never increases the secret-key rate:

$$S(X;Y||U||Z) \geq S(X;Y||[Z,U])$$

To achieve a rate which is often higher than the lower bound given above, we can use the following protocol: Using separate instances of X and Y, Alice and Bob generate two keys S_U respectively S_Z, about which U respectively Z (together with the communication transcript C) give no information. By XORing both keys, Alice and Bob calculate the final key $S := S_U \oplus S_Z$ completely unknown to both adversaries.

Theorem 6. *The secret-key rate $S(X;Y||U||Z)$ is lower bounded by*

$$\begin{aligned} S(X;Y||U||Z) &\geq \frac{S(X;Y||U) \cdot S(X;Y||Z)}{S(X;Y||U) + S(X;Y||Z)} \\ &\geq \frac{1}{2} \min\{S(X;Y||U), S(X;Y||Z)\}. \end{aligned}$$

Proof. Let us first analyze the secrecy of S, the final key. Alice and Bob use N_Z and N_U instances of X and Y to generate the intermediate keys S_Z and S_U respectively, having the following properties:

$$\begin{aligned} I(S_Z; [[Z_1, \ldots, Z_{N_Z}], C]) &\leq \epsilon, \\ I(S_U; [[Z_{N_Z+1}, \ldots, Z_{N_Z+N_U}], C]) &\leq \epsilon. \end{aligned}$$

The secrecy of the resulting key is guaranteed by the security of the one-time pad [1]:

$$I(S_Z \oplus S_U; [Z_1, \ldots, Z_{N_Z+N_U}], C]) \leq \epsilon.$$

Let $R_Z := S(X;Y||Z)$, $R_U := S(X;Y||U)$, and $R := S(X;Y||U||Z)$. As the optimum choice for N_Z and N_U results in equally-sized keys S_Z and S_U[2], it is easy to check that

$$\frac{1}{R} \leq \frac{1}{R_Z} + \frac{1}{R_U} \leq 2 \cdot \max\left\{\frac{1}{R_Z}, \frac{1}{R_U}\right\}$$

holds for the resulting rate R. The theorem follows immediately. □

[2] Before XORing S_Z and S_U, the longer key has to be cut down to the size of the shorter one in order to guarantee the secrecy of the result.

The following theorem is a straight-forward generalization of Theorem 2:

Theorem 7. *The secret-key rate $S(X;Y||U||Z)$ is lower bounded by*

$$\begin{aligned} S(X;Y||U||Z) \quad &\geq \quad I(X;Y) - \min\{\max\{I(X;U), I(X;Z)\}, \\ &\qquad\qquad \max\{I(Y;U), I(Y;Z)\}\}. \end{aligned}$$

Proof. The proof resembles the one for Theorem 2 as given in [9]. Alice and Bob transfer one message for information reconciliation. Afterwards, they both share identical strings. They apply privacy amplification to eliminate the remaining information between their strings and U, Z and the previously transferred message. As was shown in [9], this can be done without knowing the actual values of U and Z. Alice and Bob only need to know an upper bound of the information available to the adversaries through their random variables. This bound is given by $I(X;U)$ and $I(X;Z)$ (or $I(Y;U)$ and $I(Y;Z)$). Making the strings secret with respect to the adversary having more information then automatically makes the final key secret with respect the other one as well. □

5.2.2 Analysis of $S(X;Y;U||Z)$

The following simple upper bound follows directly from (5.3):

$$S(X;Y;U||Z) \leq S(X;Y;U)$$

Because one single party having the knowledge of two agreement seeking parties is always able to simulate the latter ones, we have

$$S(X;Y;U||Z) \leq \min\{S([X,Y];U||Z), S([X,U];Y||Z), S([Y,U];X||Z)\}. \quad (5.4)$$

Consider the protocol described in the proof of Theorem 6. The same idea can be used to construct a protocol for the scenario of this section. First, a "coordinator" is chosen. For both partners, the coordinator separately agrees on a secret key using distinct sequences of his random variable's instances. The coordinator combines both keys by XORing and broadcasts the result. One of the keys is picked as the final one.

The best rate achievable with this protocol and its security can easily be calculated and proven along the lines of Theorem 6.

Theorem 8. *The secret-key rate $S(X;Y;U||Z)$ is lower bounded by*

$$\begin{aligned} S(X;Y;U||Z) \quad &\geq \quad \max\left\{\frac{S(X;Y||Z)\cdot S(X;U||Z)}{S(X;Y||Z)+S(X;U||Z)},\right. \\ &\qquad\qquad \frac{S(Y;X||Z)\cdot S(Y;U||Z)}{S(Y;X||Z)+S(Y;U||Z)}, \\ &\qquad\qquad \left.\frac{S(U;X||Z)\cdot S(U;Y||Z)}{S(U;X||Z)+S(U;Y||Z)}\right\} \\ &\geq \quad \tfrac{1}{2}\max\{\min\{S(X;Y||Z), S(X;U||Z)\}, \\ &\qquad\qquad \min\{S(Y;X||Z), S(Y;U||Z)\}, \\ &\qquad\qquad \min\{S(U;X||Z), S(U;Y||Z)\}\}. \end{aligned}$$

The following theorem can be interpreted as a generalization of Theorem 2:

Theorem 9. *The secret-key rate $S(X;Y;U||Z)$ is lower bounded by*

$$\begin{aligned} S(X;Y;U||Z) \quad \geq \quad \max\{ & S(X;Y||Z) - \min\{H(X|U), H(Y|U)\}, \\ & S(X;U||Z) - \min\{H(X|Y), H(U|Y)\}, \\ & S(Y;U||Z) - \min\{H(Y|X), H(U|X)\}\}. \end{aligned}$$

Proof. The protocol can the decomposed into three phases, of which the last two again resemble the protocol described in [9] for the proof of Theorem 2. In the first phase, two parties agree on a secret key. Information reconciliation with respect to the third party follows. It is given just enough information to be able to calculate the shared key. Due to the information leaked in this step, the key is not completely secret anymore. Hence privacy amplification has to be applied as the last phase of the protocol to compress the key to a perfectly secret one. □

Finally, note that

$$S(X;Y;U||Z) > S(X;Y||Z)$$

is possible, a fact which may seem counter-intuitive at first. Consider the distribution

$$\begin{aligned} X &= R_1 \\ Y &= R_2 \\ U &= R_1, R_2 \\ Z &= \triangle \end{aligned}$$

where R_1 and R_2 are independently, uniformly distributed binary random variables. We have

$$S(X;Y;U||Z) = 1 > S(X;Y||Z) = 0$$

as the third party can act as a "coordinator" between Alice and Bob by sending $M := R_1 \oplus R_2$. This way, all three parties can agree on one secret bit per random experiment.

5.3 An Improved Upper Bound for the Secret-Key Rate

5.3.1 Intuitive Derivation

Motivated by the considerations made in Section 5.1, we will now informally derive a new upper bound which is strictly stronger than (4.1). More specifically, we want to find a good correction term c for

$$S(X;Y||Z) \leq S(X;Y||[Z,U]) + c.$$

If Eve gains or loses knowledge of U, how large is the resulting difference for the secret-key rate? A good analysis should relate U to all three random variables X, Y, and Z.

It is possible to split c into two components. One component of c will obviously have to quantify the information contained in U, which is also part of both X and Y. Additionally, we want to reduce this quantity by the amount of information already known to Eve through Z. A natural measure satisfying both requirements is $S(X;Y;U||Z)$. One might be tempted to think that $S(X;Y;U||Z)$ already captures the maximum difference possible between $S(X;Y||Z)$ and $S(X;Y||[Z,U])$. But this is not the case as the following distribution illustrates:

$$\begin{aligned} X &= R_1 \oplus R_2 \\ Y &= R_1 \oplus R_2 \\ Z &= R_1 \\ U &= R_2 \end{aligned}$$

where R_1 and R_2 are independently, uniformly distributed binary random variables. Clearly we have

$$S(X;Y||Z) = 1 > S(X;Y||[Z,U]) + S(X;Y;U||Z) = 0.$$

The effect shown here is rather unexpected: Z and U contain parts of information which are independent of both X and Y, when isolated. Putting Z and U together suddenly gives us information about X and Y. Therefore, the second component of c needs to describe the effect of splitting or joining Z and U. The rates corresponding to the scenarios "in-between" those operations are $S(X;Y||U||Z)$ and $S(X;Y||[Z,U])$ and their difference is what we are looking for. The resulting expression can be simplified, leading to a suprising new upper bound:

$$\begin{aligned} S(X;Y||Z) &\leq S(X;Y||[Z,U]) + S(X;Y;U||Z) + \\ &\quad (S(X;Y||U||Z) - S(X;Y||[Z,U])) \\ &= S(X;Y;U||Z) + S(X;Y||U||Z) \end{aligned} \tag{5.5}$$

Note that for the case that Z is constant, there is an interesting similarity between this bound and a classic formula from information theory:

$$S(X;Y) \leq S(X;Y;U) + S(X;Y||U)$$

looks like the secret-key agreement analog to

$$I(X;Y) = I(X;Y;U) + I(X;Y|U).$$

5.3.2 Proof

In order to prove (5.5), we consider N independent realizations of (X,Y,Z,U), where the random variables are jointly distributed according to P_{XYZU}.

Let $R := S(X;Y||Z)$. Then there exists, for all $\epsilon > 0$ and sufficiently large N, a protocol (with communication C) so that Alice and Bob (knowing X^N and Y^N) end up with S_A and S_B, respectively, and so that there exists a S satisfying

$$\begin{aligned} H(S) = \log_2 |\mathcal{S}| &\geq N(R-\epsilon), && (5.6) \\ P[S_A = S_B = S] &\geq 1-\epsilon, && (5.7) \\ I(S;[Z^N,C]) &\leq \epsilon. && (5.8) \end{aligned}$$

The following lemma quantifies the information between Alice's and Bob's strings.

Lemma 10. *Let S_A and S_B be defined as above. Then we have*

$$I(S_A; S_B) \geq (1-2\epsilon)\log_2|\mathcal{S}| + (1-\epsilon)\log_2(1-\epsilon) - h(\epsilon).$$

Proof. Let $\mathcal{E}$ denote the event $(S_A \neq S \vee S_B \neq S)$, let $\overline{\mathcal{E}}$ be its complement and $\chi_{\mathcal{E}}$ its characteristic random variable satisfying $\chi_{\mathcal{E}} = \mathcal{E}$ if $\mathcal{E}$ occurs and $\chi_{\mathcal{E}} = \overline{\mathcal{E}}$ otherwise. We then have

$$\begin{array}{rcl} H(S_A) & \geq & H(S_A|\chi_{\mathcal{E}}) \\ & \geq & P_{\chi_{\mathcal{E}}}(\overline{\mathcal{E}}) \cdot H(S_A|\chi_{\mathcal{E}} = \overline{\mathcal{E}}) \\ & \geq & (1-\epsilon) \cdot H_\infty(S_A|\chi_{\mathcal{E}} = \overline{\mathcal{E}}) \\ & \geq & (1-\epsilon)(-\log_2(1/|\mathcal{S}|(1-\epsilon))) \\ & = & (1-\epsilon)(\log_2|\mathcal{S}| + \log_2(1-\epsilon)), \end{array}$$

$$\begin{array}{rcl} H(S_A|S_B) & \leq & H(S_A|S_B\chi_{\mathcal{E}}) + H(\chi_{\mathcal{E}}) \\ & \leq & P_{\chi_{\mathcal{E}}}(\mathcal{E})\log_2|\mathcal{S}| + h(\epsilon) \\ & \leq & \epsilon\log_2|\mathcal{S}| + h(\epsilon) \end{array}$$

and therefore

$$\begin{array}{rcl} I(S_A; S_B) & = & H(S_A) - H(S_A|S_B) \\ & \geq & (1-2\epsilon)\log_2|\mathcal{S}| + (1-\epsilon)\log_2(1-\epsilon) - h(\epsilon). \end{array}$$

□

Two Special Lower Bounds

We derive lower bounds for $S(X;Y;U||Z)$ and $S(X;Y||U||Z)$ for the special case that $(X,Y,U,Z) = (S_A, S_B, [U^N, C], [Z^N, C])$.

Corollary 11. *Let S_A, S_B, $[U^N, C]$, and $[Z^N, C]$ be defined as above. Then we have*

$$S(S_A; S_B; [U^N, C]||[Z^N, C]) \geq I(S_A; S_B) - H(S_A|[U^N, C]) - \epsilon$$

for all $\epsilon > 0$ and sufficiently large N.

Proof. Using Theorems 9 and 2, we find

$$\begin{array}{rcl} S(S_A; S_B; [U^N, C]||[Z^N, C]) & \geq & S(S_A; [U^N, C]||[Z^N, C]) - H(S_A|S_B) \\ & \geq & I(S_A; [U^N, C]) - I(S_A; [Z^N, C]) \\ & & \quad -H(S_A|S_B) \\ & \geq & I(S_A; [U^N, C]) - H(S_A|S_B) - \epsilon \\ & = & I(S_A; S_B) - H(S_A|[U^N, C]) - \epsilon. \end{array}$$

□

Corollary 12. *Let S_A, S_B, $[U^N, C]$, and $[Z^N, C]$ be defined as above. Then we have*

$$S(S_A; S_B || [U^N, C] || [Z^N, C]) \geq I(S_A; S_B) - \max\{I(S_A; [U^N, C]), \epsilon\}$$

for all $\epsilon > 0$ and sufficiently large N.

Proof. Using Theorem 7, we find

$$\begin{aligned} S(S_A; S_B || [U^N, C] || [Z^N, C]) &\geq I(S_A; S_B) \\ &\qquad - \max\{I(S_A; [U^N, C]), I(S_A; [Z^N, C]\} \\ &\geq I(S_A; S_B) - \max\{I(S_A; [U^N, C]), \epsilon\}. \end{aligned}$$

□

Putting Everything Together

Now using Lemma 10 and Corollaries 11 and 12, we can finally prove (5.5).

Theorem 13. *Let X, Y, Z, and U be jointly distributed random variables. Then we have*

$$S(X; Y || Z) \leq S(X; Y; U || Z) + S(X; Y || U || Z).$$

Proof. Let S_A, S_B, $[U^N, C]$, and $[Z^N, C]$ be defined as above. Then we have

$$\begin{aligned} S(S_A; S_B || [U^N, C] || [Z^N, C]) + \\ S(S_A; S_B; [U^N, C] || [Z^N, C]) &\geq 2 \cdot I(S_A; S_B) - \epsilon \\ &\qquad - \max\{H(S_A), H(S_A | [U^N, C]) + \epsilon\} \\ &\geq 2 \cdot I(S_A; S_B) - \epsilon \\ &\qquad - \max\{H(S_A), H(S_A) + \epsilon\} \\ &= 2 \cdot I(S_A; S_B) - H(S_A) - 2\epsilon \\ &\geq (R - \Theta(\epsilon))N - 2\epsilon \end{aligned}$$

for all $\epsilon > 0$ and sufficiently large N. The first inequality follows from Corollaries 11 and 12, the last from Lemma 10 and (5.6). Since N realizations of (X, Y, Z, U) are required to achieve one realization of S_A, S_B, $[Z^N, C]$, and $[U^N, C]$), we have

$$\begin{aligned} S(X; Y || U || Z) + S(X; Y; U || Z) &\geq (S(S_A; S_B || [U^N, C] || [Z^N, C]) + \\ &\qquad S(S_A; S_B; [U^N, C] || [Z^N, C]))/N \\ &\geq R - \Theta(\epsilon). \end{aligned}$$

This concludes the proof because $\epsilon > 0$ is arbitrary. □

5.3.3 A Specialized Version

We can apply Theorem 13 to $(X, Y, Z, [Z, U])$ for any given X, Y, Z, and U. An informal interpretation along with an interesting application of the resulting upper bound for $S(X;Y||Z)$ can be found in Section 4.1.2.

Theorem 14. *Let X, Y, Z, and U be jointly distributed random variables. Then we have*

$$S(X;Y||Z) \leq S(X;Y||[Z,U]) + H(U|Z).$$

Proof. Note first that

$$S(X;Y||[Z,U]||Z) = S(X;Y||[Z,U])$$

follows from (5.2)[3] and (5.3). Furthermore, we have

$$\begin{array}{rcl} S(X;Y;[Z,U]||Z) & \leq & S([X,Y];[Z,U]||Z) \\ & \leq & I([X,Y];[Z,U]|Z) \\ & = & I([X,Y];U|Z) + I([X,Y];Z|Z) \\ & = & I([X,Y];U|Z) \\ & \leq & H(U|Z) \end{array}$$

where the first inequality is a direct consequence of (5.4).
Applying Theorem 13 to (X, Y, Z, U') (where $U' := [Z, U]$), we now conclude

$$\begin{array}{rcl} S(X;Y||Z) & \leq & S(X;Y||[Z,U]||Z) + S(X;Y;[Z,U]||Z) \\ & \leq & S(X;Y||[Z,U]) + H(U|Z). \end{array}$$

□

5.3.4 Applications

In Section 4.1.2, we used Theorem 14 to prove $S(X;Y||Z) \leq 1$.

An interesting application of the more general Theorem 13 is on the distribution shown in Section 4.3, which is believed to contain *bound information.* We define $U := \lfloor X/2 \rfloor$. A proof for $S(X;Y;U||Z) = 0$ should not be hard to find. Unfortunately, proving $S(X;Y||Z||U) = 0$ looks equally difficult as proving $S(X;Y||Z) = 0$ itself:

$$S(X;Y||Z||U) > 0 \iff (S(X;Y||Z) > 0 \ \wedge \ S(X;Y||U) > 0)$$

follows from (5.3) and Theorem 6. As $S(X;Y||U)$ is positive, we have

$$S(X;Y||Z||U) > 0 \iff S(X;Y||Z) > 0$$

for this distribution.

Nevertheless, it seems very plausible that the rate $S(X;Y||Z||U)$ is zero. As we have argued in Section 4.3, the only hope for secret-key agreement is by distinction of upper left and lower right quadrant. But in the scenario corresponding to $S(X;Y||Z||U)$, exactly this information is known to the party receiving U.

[3] An adversary knowing $[[Z,U],Z]$ is equivalent to one knowing “only” $[Z,U]$.

Chapter 6

Open Problems

These are some examples of open questions left. They provide motivation for further research in the field.

- To prove the existence of bound information for the candidate distribution.
- Is there an intrinsic information analog for $S(X;Y;U||Z)$ or $S(X;Y||U||Z)$?
- To find better lower and upper bounds for $S(X;Y;U||Z)$ and $S(X;Y||U||Z)$.
- To analyze the effect of two or even more adversarial side information variables.
- To specialize the new upper bound to allow for the systematic decomposition of large distributions into small ones (which can be analyzed more easily).
- To relate secret-key rates with four parties to secret-key rates involving even more parties (maybe leading to a stepwise refinement to the new upper bound).
- To apply the new upper bound to distributions corresponding to the satellite scenario by finding a good choice for U.

Bibliography

[1] G. S. Vernam, Cipher printing telegraph systems for secret wire and radio telegraphic communications, *Journal of the American Institute for Electrical Engineers*, Vol. 55, pp. 109–115, 1926.

[2] C. E. Shannon, Communication theory of secrecy systems, *Bell System Technical Journal*, Vol. 28, pp. 656–715, 1949.

[3] A. D. Wyner, The wire-tap channel, *Bell System Technical Journal*, Vol. 54, No. 8, pp. 1355–1387, 1975.

[4] W. Diffie and M. E. Hellman, New directions in cryptography, *IEEE Transactions on Information Theory*, Vol. 22, No. 6, pp. 644–654, 1976.

[5] I. Csiszár and J. Körner, Broadcast channels with confidential messages, *IEEE Transactions on Information Theory*, Vol. IT-24, pp. 339–348, 1978.

[6] U. Maurer, Secret-key agreement by public discussion from common information, *IEEE Transactions on Information Theory*, Vol. 39, No. 3, pp. 733–742, 1993.

[7] U. Maurer and S. Wolf, Towards characterizing when information-theoretic key agreement is possible, *Advances in Cryptology—ASIACRYPT '96*, Lecture Notes in Computer Science, Vol. 1163, pp. 196–209, Springer-Verlag, 1996.

[8] U. Maurer and S. Wolf, Unconditionally secure key agreement and the intrinsic conditional information, *IEEE Transactions on Information Theory*, Vol. 45, No. 2, pp. 499–514, 1999.

[9] S. Wolf, *Information-theoretically and computationally secure key agreement in cryptography*, ETH dissertation No. 13138, Swiss Federal Institute of Technology (ETH Zurich), May 1999.

[10] N. Gisin, R. Renner, and S. Wolf, Bound information: the classical analog to bound entanglement, in *Proceedings of 3ecm*, Birkhäuser Verlag, 2000.