# The Intrinsic Conditional Mutual Information and Perfect Secrecy

Ueli Maurer and Stefan Wolf

Institute for Theoretical Computer Science
ETH Zürich
CH-8092 Zürich, Switzerland
E-mail addresses: {maurer,wolf}@inf.ethz.ch

September 26, 1996

**Abstract**

This paper is concerned with secret key agreement by public discussion: two parties Alice and Bob and an adversary Eve have access to independent realizations of random variables $X$, $Y$, and $Z$, respectively, with joint distribution $P_{XYZ}$. The secret key rate $S(X;Y||Z)$ has been defined as the maximal rate at which Alice and Bob can generate a secret key by communication over an insecure, but authenticated channel such that Eve's information about this key is arbitrarily small. We define a new conditional mutual information measure, the intrinsic conditional mutual information, denoted by $I(X;Y{\downarrow}Z)$, and show that it is an upper bound on $S(X;Y||Z)$. The special scenarios where $X$, $Y$, and $Z$ are generated by sending a binary random variable $R$, for example a signal broadcast by a satellite, over independent channels, or where $Z$ is generated by sending $X$ and $Y$ over erasure channels, are analyzed. In the first scenario it can be shown, even for continuous random variables, that the secret key rate is strictly positive if and only if the intrinsic conditional mutual information between $X$ and $Y$ when given $Z$ is strictly positive. For the second scenario a new protocol is presented which allows secret key agreement even when all the previously known protocols fail.

**Keywords:** Cryptography, Secret key agreement, One-time pad, Perfect secrecy.

## 1 Introduction

Perfectly secure key agreement has been studied recently by several authors [10],[3],[7],[1],[5],[8]. Two possible approaches are based on quantum cryptography (e.g., see [1]) and on the exploitation of the noise in communication channels. In contrast to quantum cryptography, which is expensive to realize, noise is a natural property of every physical communication channel. In [7] and in [8] it has been illustrated how such noise in communication channels can be used for unconditionally secure secret key agreement and, furthermore, that it is advantageous to combine error control coding and cryptographic coding in a communication system. In particular it has been shown that if Alice, Bob, and Eve receive noisy versions of a binary random variable, for example of a signal broadcast by a satellite, then Alice and Bob can agree, by discussion over a public channel, on a secret key about which Eve has only a negligible amount of information if and only if Alice's and Bob's information have some arbitrarily weak statistical dependence with the satellite's signal, and if Eve has some minimal uncertainty about this signal.

It is a classical cryptographic problem of transmitting a message $M$ from a sender (referred to as Alice) to a receiver (Bob) over an insecure communication channel such that an enemy (Eve) with access to this channel is unable to obtain useful information about $M$. In the classical model of a cryptosystem (or cipher) introduced by Shannon [9], Eve has perfect access to the insecure channel; thus she is assumed to receive an identical copy of the ciphertext $C$ received by the legitimate receiver Bob, where $C$ is obtained by Alice as a function of the plaintext message $M$ and a secret key $K$ shared by Alice and Bob. Shannon defined a cipher system to be perfect if $I(M;C) = 0$, i.e., if the ciphertext gives no information about the plaintext or, equivalently, if $M$ and $C$ are statistically independent. When a perfect cipher is used to encipher a message $M$, an enemy can do no better than guess $M$ without even looking at the ciphertext $C$. Shannon proved the pessimistic result that perfect secrecy can be achieved only when the secret key is at least as long as the plaintext message or, more precisely, when $H(K) \geq H(M)$.

For this reason, perfect secrecy is often believed to be impractical. In [7] this pessimism has been relativized by pointing out that Shannon's apparently innocent assumption that, except for the secret key, the enemy has access to precisely the same information as the legitimate receiver, is very restrictive and that indeed in many practical scenarios, especially if one considers the fact that every transmission of data is ultimately based on the transmission of an analog signal subject to noise, the enemy has some minimal uncertainty about the signal received by the legitimate receiver(s).

Wyner [10] and subsequently Csiszár and Körner [3] considered a scenario in which the enemy Eve is assumed to receive messages transmitted by the sender Alice over a channel that is noisier than the legitimate receiver Bob's channel. The assumption that Eve's channel is worse than the main channel is unrealistic in general. It was shown in [7] that this assumption can be unnecessary if Alice and Bob can also communicate over a completely insecure (but authenticated) public channel.

For the case where Alice, Bob, and Eve have access to repeated independent realizations of random variables $X$, $Y$, and $Z$, respectively, with joint distribution $P_{XYZ}$, the rate at which Alice and Bob can generate a secret key by public discussion over an insecure channel is defined in [7] as follows.

**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 $S$ such that the rate at which Eve obtains information about $S$ is arbitrarily small. In other words, it is the maximal $R$ such that for every $\varepsilon > 0$ 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 $X^N = [X_1, \ldots, X_N]$ and $Y^N = [Y_1, \ldots, Y_N]$, respectively, have the same key $S$ with probability at least $1 - \varepsilon$, satisfying

$$\frac{1}{N} I(S; VZ^N) \leq \varepsilon \qquad \text{and} \qquad \frac{1}{N} H(S) \geq R - \varepsilon \ ,$$

where $V$ denotes the collection of messages sent over the insecure channel by Alice and Bob, and where $Z^N = [Z_1, \ldots, Z_N]$.

The following lower bound for $S(X;Y||Z)$ is proved in [7], and follows from a result by Csiszár and Körner [3]:

$$\max\{I(X;Y) - I(X;Z), I(Y;X) - I(Y;Z)\} \leq S(X;Y||Z) \ . \qquad (1)$$

As already mentioned, it has been first shown by an example in [7] that the secret key rate $S(X;Y||Z)$ can be strictly positive even when both $I(X;Z) > I(X;Y)$ and $I(Y;Z) > I(Y;X)$.

We give a brief outline of this paper. In Section 2 we define a new conditional mutual information measure and show that this measure gives an improved upper bound for the secret key rate. In the later sections we address the problem whether secret key agreement is always possible when this new upper bound is strictly positive. We consider this in the scenarios where $X$, $Y$, and $Z$ are generated by sending a binary random variable over independent channels, and where $Z$ is generated by sending $X$ and $Y$ over erasure channels. We show the somewhat surprising fact that a new protocol can be more powerful than the previously known protocols in the latter scenario.

## 2 The intrinsic conditional mutual information

The following upper bound on the secret key rate has been proved in [7]:

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

Trying to reduce the quantity $I(X;Y|Z)$ in this upper bound, the adversary Eve can send the random variable $Z$ over a channel, characterized by $P_{\overline{Z}|Z}$, in order to generate the random variable $\overline{Z}$, and hence

$$S(X;Y||Z) \leq S(X;Y||\overline{Z}) \leq I(X;Y|\overline{Z}) \tag{2}$$

holds for every such $\overline{Z}$. This motivates the following definition of the intrinsic conditional mutual information between $X$ and $Y$ when given $Z$, which is the minimum of $I(X;Y|\overline{Z})$, taken over all random variables $\overline{Z}$ that can be obtained by sending $Z$ over a channel, characterized by $P_{\overline{Z}|Z}$.

**Definition 2** For a distribution $P_{XYZ}$, the *intrinsic conditional mutual information between $X$ and $Y$ when given $Z$*, denoted by $I(X;Y\downarrow Z)$, is

$$I(X;Y\downarrow Z) := \min\left\{ I(X;Y|\overline{Z}) \ : \ P_{XY\overline{Z}} = \sum_{z \in \mathcal{Z}} P_{XYZ} \cdot P_{\overline{Z}|Z} \right\} \ ,$$

where the minimum is taken over all possible conditional distributions $P_{\overline{Z}|Z}$.

**Theorem 1** *For arbitrary random variables $X$, $Y$, and $Z$, we have*

$$S(X;Y||Z) \leq I(X;Y\downarrow Z) \ . \tag{3}$$

**Proof:** The bound (3) follows from (2). □

Theorem 1 implies in particular that secret key agreement can be possible only if $I(X;Y\downarrow Z) > 0$. The intrinsic conditional information has the following properties:

$$\begin{aligned} 0 \ \leq \ I(X;Y\downarrow Z) \ &\leq \ I(X;Y) \ , \\ I(X;Y\downarrow Z) \ &\leq \ I(X;Y|Z) \ , \\ \text{and} \quad I(X;Y\downarrow Z) \ &\leq \ I(X;Y\downarrow \overline{Z}) \ , \end{aligned}$$

where $\overline{Z}$ is generated from $Z$ by an arbitrary channel. The following example shows that the intrinsic conditional information can be equal to 0 (and secret key agreement is impossible) even when both $I(X;Y|Z) > 0$ and $I(X;Y) > 0$. Let $\mathcal{X} = \mathcal{Y} = \mathcal{Z} = \{0,1,2,3\}$,

$$P_{XYZ}(0,0,0) = P_{XYZ}(0,1,1) = P_{XYZ}(1,0,1) = P_{XYZ}(1,1,0) = \frac{1}{8} ,$$

and

$$P_{XYZ}(2,2,2) = P_{XYZ}(3,3,3) = \frac{1}{4} .$$

Then $I(X;Y) = 1/2$ and $I(X;Y|Z) = 1/2$ (note that $Z = X \oplus Y$ if $X, Y \in \{0,1\}$), but $I(X;Y{\downarrow}Z) = 0$. To see this, consider the random variable $\overline{Z}$, generated by sending $Z$ over the channel characterized by

$$P_{\overline{Z}|Z}(0,0) = P_{\overline{Z}|Z}(0,1) = P_{\overline{Z}|Z}(1,0) = P_{\overline{Z}|Z}(1,1) = \frac{1}{2}$$

and

$$P_{\overline{Z}|Z}(2,2) = P_{\overline{Z}|Z}(3,3) = 1 .$$

Intuitively, giving the side information $Z$ "destroys" all the information between $X$ and $Y$, but generates new mutual information (that cannot be used to generate a secret key). 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 $Y$ (when given $Z$), but *not* the *additional* information between $X$ and $Y$ when giving $Z$. A graphical representation of $I(X;Y{\downarrow}Z)$ is described in Appendix A.

## 3 Secret key agreement with general random variables

The intrinsic conditional mutual information $I(X;Y{\downarrow}Z)$ defined above gives a new upper bound for the secret key rate, and in particular, secret key agreement is impossible unless $I(X;Y{\downarrow}Z) > 0$. It appears plausible that this condition is also sufficient for a positive secret key rate. This is a fundamental open question, and we will address this problem in the following sections.

As a preparation we prove three lemmas that we will use later when considering different scenarios. All these lemmas follow immediately from the definition of the secret key rate. Lemma 2 states that Alice and Bob cannot increase the secret key rate when they discard a realization of $[X,Y]$ if $X$ (or $Y$) does not lie in a specified subset $\hat{\mathcal{X}}$ ($\hat{\mathcal{Y}}$) of the range $\mathcal{X}$ ($\mathcal{Y}$). In this case we say that Alice and Bob obtain new random variables by *restriction of the ranges*. Lemma 3 states that processing $X$ and $Y$ does not help increasing the secret key rate.

**Lemma 2** *Let $X$, $Y$, and $Z$ be random variables with ranges $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ and joint distribution $P_{XYZ}$. For $\hat{\mathcal{X}} \subset \mathcal{X}$ and $\hat{\mathcal{Y}} \subset \mathcal{Y}$, we define a new random experiment with random variables $\hat{X}$ and $\hat{Y}$ (with ranges $\hat{\mathcal{X}}$ and $\hat{\mathcal{Y}}$, respectively). If $\Omega$ is the event that $X \in \hat{\mathcal{X}}$ and $Y \in \hat{\mathcal{Y}}$, then the joint distribution of $\hat{X}$ and $\hat{Y}$ with $Z$ is defined as follows:*

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

*for all $(x,y,z) \in \hat{\mathcal{X}} \times \hat{\mathcal{Y}} \times \mathcal{Z}$. (This is a normed probability distribution for $\hat{\mathcal{X}} \times \hat{\mathcal{Y}} \times \mathcal{Z}$.) Then*

$$S(X;Y||Z) \geq P_{XYZ}[\Omega] \cdot S(\hat{X};\hat{Y}||Z) ,$$

*In other words, the secret key rate cannot be increased by restricting the ranges of $X$ and $Y$.*

**Lemma 3** *Let $X$, $Y$, $Z$, $\overline{X}$, and $\overline{Y}$ be random variables with distribution*

$$P_{XYZ\overline{X}\,\overline{Y}} = P_{XYZ} \cdot P_{\overline{X}|X} \cdot P_{\overline{Y}|Y} \ ,$$

*where $P_{\overline{X}|X}$ and $P_{\overline{Y}|Y}$ are arbitrary conditional probability distributions. Then $S(\overline{X};\overline{Y}||Z) \leq S(X;Y||Z)$.*

**Definition 3** We say that $\overline{X}$ *and* $\overline{Y}$ *are generated from $X$ and $Y$ with positive probability* if one can obtain from $X$ and $Y$ random variables $\hat{X}$ and $\hat{Y}$ by restriction of the ranges (see above), and the random variables $\overline{X}$ and $\overline{Y}$ by sending $\hat{X}$ and $\hat{Y}$ over two channels, specified by $P_{\overline{X}|\hat{X}}$ and $P_{\overline{Y}|\hat{Y}}$.

Lemma 4 states that if Eve has access to a random variable $U$ (in addition to $Z$) that can be interpreted as side information provided by an oracle, then the secret key rate is not greater than in the original situation.

**Lemma 4** *Let $X$, $Y$, $Z$, and $U$ be arbitrary random variables. Then*

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

We will in particular make use of the following corollary which is an immediate consequence of the three Lemmas 2, 3, and 4.

**Corollary 5** *Let $X$, $Y$, $Z$, and $U$ be arbitrary random variables, and let $\overline{X}$ and $\overline{Y}$ be generated from $X$ and $Y$ with positive probability. Then $S(\overline{X};\overline{Y}||[Z,U]) > 0$ implies $S(X;Y||Z) > 0$. In other words, secret key agreement by public discussion with respect to $X$, $Y$, and $Z$ is possible if the same holds for $\overline{X}$, $\overline{Y}$, and $[Z,U]$.*

## 4 The scenario of independent channels

The scenario where $X$, $Y$, and $Z$ are discrete random variables generated by sending a binary random variable $R$ over three arbitrary independent channels $C_A$, $C_B$, and $C_E$ (see Figure 1), was considered and completely analyzed in [8]. It was shown that secret key agreement is possible if and only if $I(X;Y|Z) > 0$, which is therefore equivalent to $I(X;Y{\downarrow}Z) > 0$ in this scenario (one can even show that $I(X;Y|Z) = I(X;Y \downarrow Z)$ always holds here). The given proof does not apply to the realistic case where Eve's channel is an analog channel, because the resulting random variable is not discrete. In Appendix B we prove a similar result for *continuous* random variables generated from independent binary-input channels.

We state the main theorem and two lemmas of [8] which will be required in the following sections, and we sketch the proof of the main result in order to present the protocol for achieving a positive secret key rate, and some further techniques.

**Theorem 6** *Let $R$ be a binary random variable, and let $X$, $Y$, and $Z$ be discrete random variables (with ranges $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, respectively), generated by sending $R$ over independent channels $C_A$, $C_B$, and $C_E$, i.e., $P_{XYZ|R}(x,y,z,r) = P_{X|R}(x,r) \cdot P_{Y|R}(y,r) \cdot P_{Z|R}(z,r)$ for all $x \in \mathcal{X}$, $y \in \mathcal{Y}$, $z \in \mathcal{Z}$, and $r \in \{0,1\}$. Then the secret key rate is strictly positive, i.e., $S(X;Y||Z) > 0$, if and only if $I(X;Y|Z) > 0$.*

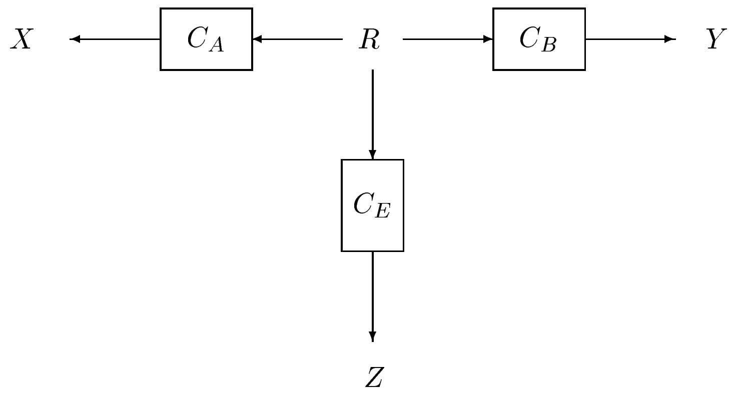


Figure 1: The scenario of three independent channels.

The proof of this theorem (see [8]) is roughly as follows. From $I(X;Y|Z) > 0$ one can conclude that there exists a channel, characterized by $P_{\overline{R}|R}$, such that the resulting random variable $\overline{R}$ is binary and symmetrically distributed, and that there exist symmetric binary random variables $\overline{X}$ and $\overline{Y}$ which can be generated from $X$ and $Y$ with positive probability, such that $\overline{X}$ and $\overline{Y}$ can be interpreted as generated by sending $\overline{R}$ over independent binary symmetric channels $\overline{C_A}$ and $\overline{C_B}$ with identical crossover probabilities smaller than $1/2$. Moreover, there exists side information $U$ such that $[Z,U]$ is equivalent to a random variable which is generated from $\overline{R}$ by an erasure channel $\overline{C_E}$ with positive erasure probability, and which is independent of the pair $(\overline{C_A}, \overline{C_B})$. This situation is illustrated in Figure 2.

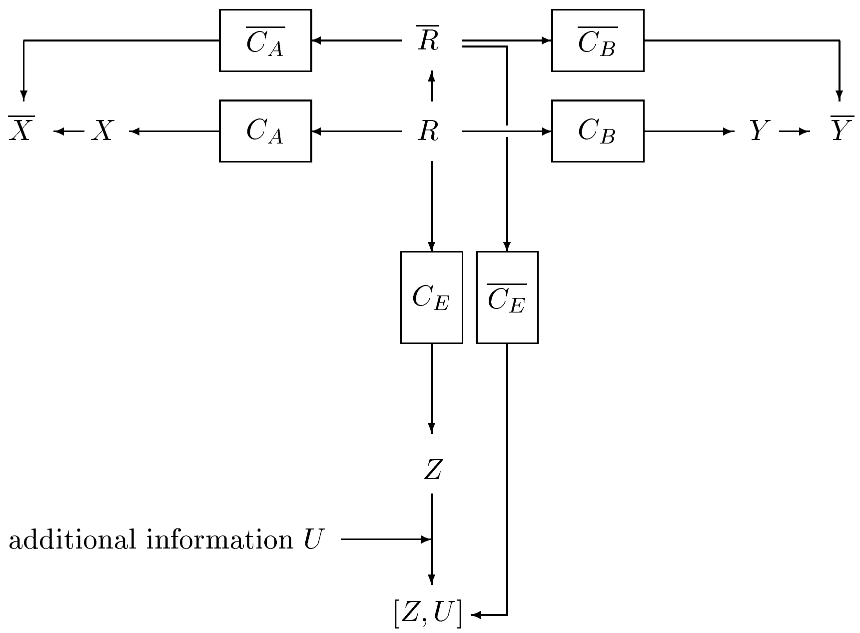


Figure 2: The situation in the proof of Theorem 6

According to Corollary 5 it suffices to show that $S(\overline{X};\overline{Y}||[Z,U]) > 0$. We state two lemmas from [8]. Lemma 7 states that in the scenario with $\overline{X}$, $\overline{Y}$, and $[Z, U]$, there exists, for all $N$, a protocol that allows Alice to send a bit such that Bob's error probability about the bit is asymptotically smaller (for $N \to \infty$) than Eve's error probability (even when she uses the optimal strategy), and Lemma 8 states that this is sufficient for $S(X;Y||Z) > 0$.

**Lemma 7** *Let $R$, $X$, and $Y$ be symmetric binary random variables such that $X$ and $Y$ can be interpreted as generated by sending $R$ over independent binary symmetric channels $C_A$ and $C_B$ with identical error probability smaller than $1/2$. Let $Z$ be generated by sending $R$ over an erasure channel $C_E$, independent of the pair $(C_A, C_B)$, with positive erasure probability. Then, for every $N$, there exists a protocol, using public discussion over an insecure channel, that allows Alice (who knows $X^N = [X_1, \ldots, X_N]$) to send a bit to Bob and Eve (who know $Y^N$ and $Z^N$, respectively) such that Bob's error probability $\beta_N$ and Eve's error probability $\gamma_N$ (for all possible strategies) when guessing this bit satisfy (for sufficiently large $N$)*

$$\beta_N \leq b^N \ , \quad \gamma_N \geq c^N$$

*for some constants $b$ and $c$ with $b < c$.*

The considered protocol (which was presented already in [7]) works as follows (for fixed $N$).

**Protocol A.** Alice randomly chooses a bit $C$ and sends $[C \oplus X_1, C \oplus X_2, \ldots, C \oplus X_N]$ over the public channel. Bob computes $[(C \oplus X_1) \oplus Y_1, \ldots, (C \oplus X_N) \oplus Y_N]$ and accepts if and only if this is equal to either $[0, 0, \ldots, 0]$ or $[1, 1, \ldots, 1]$. In other words, Alice and Bob make use of a repeat code of length $N$ with only two codewords $[0, 0, \ldots, 0]$ and $[1, 1, \ldots, 1]$.

It is obvious that Eve's optimal strategy in the described scenario is to compute the block $[(C \oplus X_1) \oplus Z_1, \ldots, (C \oplus X_N) \oplus Z_N]$ and guess $C$ as 0 if at least half of the bits in this block are 0, and as 1 otherwise.

This protocol is computationally efficient, but it is not efficient in terms of the size of the generated secret key. For special cases there exist protocols that are much more efficient with respect to the effective key generation rate [6].

Because of its importance, we also restate the proof of the following lemma from [8], which completes the proof of Theorem 6.

**Lemma 8** *Let $X$, $Y$, and $Z$ be arbitrary random variables, and let $C$ be a bit, randomly chosen by Alice. Assume that for all $N$, Alice can generate a message $M$ from $X^N = [X_1, \ldots, X_N]$ and $C$ (and possibly some random bits) such that with some probability $p_N > 0$, Bob (who knows $M$ and $Y^N$) accepts and can compute a bit $C'$ such that* $\text{Prob}[C \neq C'] \leq b^N$ *for some $b \geq 0$. If in addition, given that Bob accepts, for every strategy for guessing $C$ when given $M$ and $Z^N$ the average error probability $\gamma_N$ is at least $c^N$ for some $c > b$ and for sufficiently large $N$, then $S(X;Y||Z) > 0$.*

**Proof:** According to inequality (1) it suffices to show that Alice and Bob can, for some $N$, construct random variables $\hat{X}$ and $\hat{Y}$ from $X^N$ and $Y^N$ by exchanging messages over an insecure, but authenticated channel, such that

$$I(\hat{X};\hat{Y}) - I(\hat{X};\hat{Z}) > 0 \tag{4}$$

with $\hat{Z} = [Z^N, V]$, where $V$ is the collection of all messages sent over the public channel.

Let $\hat{X}$ and $\hat{Y}$ be defined as follows. If Bob accepts, let $\hat{X} = C$ and $\hat{Y} = C'$, and if Bob (publicly) rejects, let $\hat{X} = \hat{Y} =$"reject". We show that (4) holds for sufficiently large $N$. If Bob accepts then

$$H(C|C') \leq h(b^N) \leq 2b^N \cdot \log_2(1/b^N) = 2b^N \cdot N \cdot \log_2(1/b) < c^N$$

for sufficiently large $N$ (where $h(p) = -p\log_2 p - (1-p)\log_2(1-p)$ is the binary entropy function, the first inequality follows from Jensen's inequality, and the reason for the second inequality is that $-p\log_2 p \geq -(1-p)\log_2(1-p)$ for $p \leq 1/2$). On the other hand

$$H(C|\hat{Z}) = \sum_{\hat{z} \in \mathcal{Z}^N \times \mathcal{M}} P_{\hat{Z}}(\hat{z}) \cdot H(C|\hat{Z} = \hat{z}) = E[h(q(\hat{Z}))] \geq E[q(\hat{Z})] = \gamma_N \geq c^N,$$

where $q(\hat{z})$ is the probability of guessing $C$ incorrectly with the optimal strategy (i.e., $q(\hat{z}) \leq 1/2$), given that $\hat{Z} = \hat{z}$. Given that Bob publicly rejects, we have $H(\hat{X}|\hat{Y}) = H(\hat{X}|\hat{Z}) = H(\hat{X}|V) = 0$. From $p_N > 0$ we conclude that $I(\hat{X};\hat{Y}) - I(\hat{X};\hat{Z}) > 0$. □

## 5 Towards the general scenario: Protocol A is not optimal

In this section $X$ and $Y$ are completely general random variables, and Eve obtains her information from a random variable that is generated by sending $X$ and $Y$ over erasure channels. The advantage of considering such a scenario is that is it less difficult to analyze than the completely general situation. Moreover, many more general situations can be reduced, by the methods of Corollary 5, to such a scenario (with respect to the question whether secret key agreement is possible). Later in this section we will prove the somewhat surprising fact that the described Protocol A is not optimal. A new protocol, Protocol B, works in many situations in which Protocol A fails. We will consider the following two scenarios.

**Scenario 1.** The random variables $X$ and $Y$ are binary and distributed according to

$$P_{XY}(0,0) = P_{XY}(1,1) = \frac{1-\alpha}{2}\ , \quad P_{XY}(0,1) = P_{XY}(1,0) = \frac{\alpha}{2} \tag{5}$$

for some $\alpha < 1/2$. The random variable $Z$ is generated by sending $[X,Y]$ over an erasure channel with positive erasure probability $1-r$.

**Scenario 2.** The random variables $X$ and $Y$ are distributed as in Scenario 1, and $Z = [Z_X, Z_Y]$, where $Z_X$ and $Z_Y$ are generated by sending $X$ and $Y$, respectively, over two independent erasure channels with positive erasure probabilities.

Prior to this we show that under a condition which appears to be satisfied normally if $X$, $Y$, and $Z$ are completely general random variables with $I(X;Y{\downarrow}Z) > 0$, there exist random variables $\overline{X}$ and $\overline{Y}$, which can be generated from $X$ and $Y$ with positive probability, and side information $U$ such that $\overline{X}$, $\overline{Y}$, and $[Z,U]$ correspond to one of the special scenarios 1 or 2. Corollary 5 then implies that $S(X;Y||Z) > 0$ if $S(\overline{X};\overline{Y}||[Z,U]) > 0$.

**Lemma 9** *Let $I(X;Y) > 0$. Then Alice and Bob can generate symmetric binary random variables $\overline{X}$ and $\overline{Y}$ from $X$ and $Y$ with positive probability such that $\overline{X}$ and $\overline{Y}$ have a symmetric joint distribution as given in (5) for some $\alpha < 1/2$.*

**Proof:** Because $X$ and $Y$ are not statistically independent there exist $x, x' \in \mathcal{X}$ and $y, y' \in \mathcal{Y}$ satisfying

$$P_{X|Y}(x,y) > P_X(x) > P_{X|Y}(x,y') \tag{6}$$

and

$$P_{Y|X}(y,x) > P_Y(y) > P_{Y|X}(y,x') \ .$$

Alice and Bob can generate random variables $\hat{X}$ and $\hat{Y}$ by restricting the ranges of $X$ and $Y$ to $\{x,x'\}$ and $\{y,y'\}$, respectively. Then $\hat{X}$ is sent over the following channel (we can assume without loss of generality that $P_{\hat{X}}(x) \geq 1/2$):

$$P_{\overline{X}|\hat{X}}(0,x) = \frac{1}{2 \cdot P_{\hat{X}}(x)} \ , \quad P_{\overline{X}|\hat{X}}(1,0) = 1 - P_{\overline{X}|\hat{X}}(0,x) \ , \quad P_{\overline{X}|\hat{X}}(1,x') = 1 \ .$$

It is obvious that $P_{\overline{X}}(0) = P_{\overline{X}}(1) = 1/2$. The symmetrically distributed random variable $\overline{Y}$ can be obtained from $\hat{Y}$ in an analogous way. Then $\overline{X}$ and $\overline{Y}$ are distributed according to (5) with $\alpha = 1 - P_{\hat{X}\hat{Y}}(x,y)/(2 \cdot P_{\hat{X}}(x) \cdot P_{\hat{Y}}(y))$ if $P_{\hat{Y}}(y) \geq 1/2$, and $\alpha = P_{\hat{X}\hat{Y}}(x,y')/(2 \cdot P_{\hat{X}}(x) \cdot P_{\hat{Y}}(y'))$ if $P_{\hat{Y}}(y) < 1/2$. In both cases we have $\alpha < 1/2$ because of (6). □

If, for such $\overline{X}$ and $\overline{Y}$, there exists $z \in \mathcal{Z}$ such that the conditional probabilities $P_{\overline{X}\,\overline{Y}|Z=z}(i,j)$ are positive for all $(i,j) \in \{0,1\}^2$, then there exists side information $U$ such that $U$ equals $[\overline{X}, \overline{Y}]$ with some probability (that depends on $[\overline{X}, \overline{Y}]$), but where $U$ contains no information about $\overline{X}$ or $\overline{Y}$ otherwise, and such that $[Z, U]$ can be interpreted as being obtained by sending $[\overline{X}, \overline{Y}]$ over an erasure channel with positive erasure probability. In other words, the probability that $U = [\overline{X}, \overline{Y}]$ must be such that the probabilities that $[\overline{X}, \overline{Y}, Z] = [i,j,z]$ (and that $U$ contains no information about $[\overline{X}, \overline{Y}]$) are proportional to the probabilities of the events $[\overline{X}, \overline{Y}] = [i,j]$.

We conclude that very general situations can be reduced to the special scenario in which $Z$ is obtained by sending $[X,Y]$ over an erasure channel, and which is less difficult to analyze. In an analogous way, general scenarios can be reduced to the (asymmetric) situation where Eve obtains her information about $X$ and $Y$ from independent erasure channels. However, it appears to be difficult to decide in general which reduction leads to the strongest results.

### 5.1 Analysis of Scenario 1

We consider the symmetric scenario where $X$ and $Y$ are distributed according to (5) with $\alpha < 1/2$, and $Z$ is obtained by sending $[X,Y]$ over an erasure channel with erasure probability $1-r$. We now derive a condition for when Protocol A allows secret key agreement. Bob's conditional error probability when guessing the bit sent by Alice, given that he accepts, is

$$\beta_N = \frac{1}{p_N} \cdot \alpha^N \leq \left(\frac{\alpha}{1-\alpha}\right)^N \ ,$$

where $p_N = \alpha^N + (1-\alpha)^N$ is the probability that Bob accepts the received block. Given that Bob accepts, Eve (using the optimal strategy) guesses the bit sent by Alice correctly unless she receives $N$ times the erasure symbol $\Delta$. In the latter case her error probability is $1/2$, and is independent of her strategy. Hence Eve's error probability, given that Bob accepts, is

$$\gamma_N = \frac{1}{2} \cdot (1-r)^N \ .$$

Using Lemma 8, we conclude that Protocol A works and allows the generation of a secret key if

$$1 - r > \frac{\alpha}{1-\alpha} .$$

The next theorem shows that no other protocol can improve this bound, i.e., that Protocol A is optimal here, and that we have found another situation for which the conditions $I(X;Y{\downarrow}Z) > 0$ and $S(X;Y||Z) > 0$ are equivalent.

**Theorem 10** *In the described scenario, the following three conditions are equivalent:*

1. $I(X;Y{\downarrow}Z) > 0$ ,
2. $1 - r > \alpha/(1-\alpha)$ ,
3. $S(X;Y||Z) > 0$ .

**Proof:** It remains to show that $I(X;Y{\downarrow}Z) = 0$ if $1 - r \leq \alpha/(1-\alpha)$. The random variable $\overline{Z}$ can be generated from $Z$ by the following channel: if $Z = \Delta$, $Z = [0,1]$, or $Z = [1,0]$, then $\overline{Z} = \overline{\Delta}$. When $Z = [0,0]$ or $Z = [1,1]$, then $\overline{Z}$ is defined to be equal to $Z$ with probability

$$\frac{1-2\alpha}{r(1-\alpha)} \quad (\leq 1) ,$$

and equal to $\overline{\Delta}$ otherwise.

We show that $I(X;Y|\overline{Z}) = 0$. It is obvious that $I(X;Y|\overline{Z} = [0,0]) = I(X;Y|\overline{Z} = [1,1]) = 0$. Because of

$$P_{XY\overline{Z}}(0,1,\overline{\Delta}) = P_{XY\overline{Z}}(1,0,\overline{\Delta}) = \frac{\alpha}{2}$$

and

$$\begin{aligned} P_{XY\overline{Z}}(0,0,\overline{\Delta}) &= P_{XY\overline{Z}}(1,1,\overline{\Delta}) \\ &= \frac{1-\alpha}{2} \cdot (1-r) + \frac{1-\alpha}{2} \cdot r \cdot \left(1 - \frac{1-2\alpha}{r(1-\alpha)}\right) = \frac{\alpha}{2} , \end{aligned}$$

$X$ and $Y$ also are independent when given $\overline{Z} = \overline{\Delta}$, i.e., $I(X;Y|\overline{Z} = \overline{\Delta}) = 0$. □

A similar pessimistic result can even be shown in the general scenario of arbitrary random variables $X$, $Y$, and $Z$. We show that $I(X;Y{\downarrow}Z) = 0$ if Eve knows $X$ and $Y$ precisely with some positive probability, and if the common distribution of $X$ and $Y$ is too close to an "independent distribution". We have to define an appropriate measure for the "deviation from independence" of the common distribution $P_{XY}$ of the two random variables $X$ and $Y$.

**Definition 4** Let $X$ and $Y$ be random variables with ranges $\mathcal{X}$ and $\mathcal{Y}$, respectively, and common distribution $P_{XY}$. Let

$$F(P_{XY}) := \min_{Q_{XY}} \left( \max_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \left( \frac{P_{XY}(x,y)}{Q_{XY}(x,y)} \right) \cdot \max_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \left( \frac{Q_{XY}(x,y)}{P_{XY}(x,y)} \right) \right) ,$$

where the minimum is taken over all probability distributions $Q_{XY}$ for which $X$ and $Y$ are statistically independent, and where we set $0/0 := 1$ and $c/0 := \infty$ for $c > 0$. The *deviation* $d_{ind}(P_{XY})$ *of* $P_{XY}$ *from independence* is defined as

$$d_{ind}(P_{XY}) := 1 - \frac{1}{F(P_{XY})} ,$$

where we set $1/\infty := 0$.

For every $P_{XY}$,

$$0 \leq d_{ind}(P_{XY}) \leq 1 \ , \tag{7}$$

where equality on the left side of (7) holds if and only if the random variables $X$ and $Y$ are independent, and equality on the right side of (7) holds if and only if there exist $x$, $x'$, $y$, and $y'$ such that $P_{XY}(x, y) = 0$, but $P_{XY}(x, y') > 0$ and $P_{XY}(x', y) > 0$. Furthermore,

$$d_{ind}(P_{XY}) \leq 1 - \frac{\min P_{XY}}{\max P_{XY}} \ .$$

This can be seen by taking the uniform distribution for $Q_{XY}$. When $X$ and $Y$ are distributed according to (5), then

$$d_{ind}(P_{XY}) = 1 - \frac{\alpha}{1-\alpha} = \frac{1-2\alpha}{1-\alpha} \ .$$

The next theorem implies that secret key agreement is impossible if the probability that Eve reliably knows $X$ and $Y$ equals or exceeds $d_{ind}(P_{XY})$.

**Theorem 11** *Let $X$ and $Y$ be arbitrary random variables with common distribution $P_{XY}$, and let $Z$ be generated by sending $[X, Y]$ over an erasure channel with erasure probability $1 - r$. Then*

$$r \geq d_{ind}(P_{XY})$$

*implies $I(X; Y \downarrow Z) = 0$.*

**Proof:** Let $r \geq d_{ind}(P_{XY}) = 1 - 1/F(P_{XY})$, i.e., $F(P_{XY}) \leq 1/(1-r)$. Then, from the definition of $F(P_{XY})$, we conclude that there exists a distribution $Q_{XY}$, corresponding to an independent distribution of $X$ and $Y$, such that

$$(1 - r) \cdot P_{XY}(x, y) \leq \lambda \cdot Q_{XY}(x, y) \leq P_{XY}(x, y) \tag{8}$$

for all $x \in \mathcal{X}$, $y \in \mathcal{Y}$, and for some $0 \leq \lambda \leq 1$. We define the random variable $\overline{Z}$, which can be obtained from $Z$, as follows. If $Z = \Delta$, then $\overline{Z} = \overline{\Delta}$. Because $P_{XYZ}(x, y, \Delta) = (1-r) \cdot P_{XY}(x, y)$, and because of (8), $\overline{Z}$ can be defined to be equal to $\overline{\Delta}$ with some conditional probability when given $Z = [x, y]$, and $\overline{Z} = Z$ otherwise, such that

$$P_{XY\overline{Z}}(x, y, \overline{\Delta}) = \lambda \cdot Q_{XY}(x, y) \ . \tag{9}$$

This can be done for all pairs $(x, y)$, and (9) implies $P_{XY|\overline{Z}=\overline{\Delta}} = Q_{XY}$, i.e., that $X$ and $Y$ are independent when given $\overline{Z} = \overline{\Delta}$. Hence

$$I(X; Y \downarrow Z) \leq I(X; Y | \overline{Z}) = P_{\overline{Z}}(\overline{\Delta}) \cdot I(X; Y | \overline{Z} = \overline{\Delta}) = 0 \ .$$

□

## 5.2 Analysis of Scenario 2

We show that Protocol A (using a repeat code) is not optimal in Scenario 2 by presenting a different protocol that allows secret key agreement in many situations for which Protocol A fails.

Let $1-r_X$ and $1-r_Y$ be the erasure probabilities of the (independent) erasure channels generating $Z_X$ and $Z_Y$ from $X$ and $Y$, respectively. We assume that $r_Y \geq r_X$. For fixed $\alpha$ and $r_Y$, we prove three different upper bounds on $r_X$ such that secret key agreement is possible if $r_X$ is smaller than at least one of these bounds.

According to the lower bound (1) the secret key rate is positive if $I(X;Y) > I(X;Z)$. This condition is equivalent to $H(X|Y) < H(X|Z)$, and to

$$h(\alpha) < (1-r_X)(1-r_Y) + (1-r_X)r_Y h(\alpha) \ . \tag{10}$$

The following lemma gives a first upper bound for $r_X$ which depends on $r_Y$ and $\alpha$, and such that if $r_X$ is smaller than this bound, then $S(X;Y||Z) > 0$.

**Lemma 12** *In Scenario 2, $S(X;Y||Z)$ is strictly positive if*

$$r_X < 1 - \frac{h(\alpha)}{1-r_Y+r_Y h(\alpha)} \ . \tag{11}$$

**Proof:** It is easy to see that (11) is equivalent to (10). □

If Lemma 12 does not apply, in some cases one can prove that secret key agreement is nevertheless possible by using Protocol A. When the block length is $N$, the probability $p_{10}$ that Bob accepts and receives the bit sent by Alice incorrectly, and that Eve receives this bit correctly is upper bounded by $\alpha^N$.

On the other hand, the probability $p_{01}$ that Bob accepts and receives the correct bit, and that Eve receives the bit incorrectly, satisfies

$$p_{01} \geq \frac{1}{2}(1-\alpha)^N(1-r_X)^N(1-r_Y)^N \ .$$

The reason for this is that if Eve receives only erasure symbols, her error probability about the bit sent by Alice is, independent of her strategy, equal to $1/2$. Finally, the probability $p_{11}$ that Bob accepts, and that both Bob and Eve receive the bit incorrectly satisfies

$$p_{11} \leq \alpha^N(1-r_X)^N \ .$$

Hence Bob's error probability is of order $O(\alpha^N)$, whereas Eve's error probability is of order $\Omega(((1-\alpha)(1-r_X)(1-r_Y))^N + (\alpha(1-r_X))^N)$. From this and from Lemma 8 we can conclude that Protocol A works if and only if

$$\alpha < (1-\alpha)(1-r_X)(1-r_Y) \ , \tag{12}$$

and the following lemma is proved.

**Lemma 13** *In Scenario 2, the secret key rate is strictly positive if*

$$r_X < 1 - \frac{\alpha}{(1-\alpha)(1-r_Y)} \ . \tag{13}$$

Note that this bound can be fulfilled only if $1-r_Y > \alpha/(1-\alpha)$. This is the same condition as in Theorem 10 of the previous section.

**Corollary 14** *In Scenario 2, the secret key rate is strictly positive if*

$$r_X < 1 - \min\left\{\frac{h(\alpha)}{1 - r_Y + r_Y h(\alpha)}\;,\;\frac{\alpha}{(1-\alpha)(1-r_Y)}\right\}\;. \tag{14}$$

We remark that each of the expressions in (14) can dominate the other. If $r_Y$ is constant, and $\alpha \to 0$, the second expression is smaller, whereas if $r_Y = \alpha/(1-\alpha)$, the second expression equals 1, and the first is smaller than 1 for all $\alpha < 1/2$.

Intuitively, the repeat-code protocol (Protocol A) does not appear to be very appropriate in a situation where Eve has perfect access to $X$ or $Y$ with some positive probability, because revealing one bit of a repeat code block means revealing the entire block. It is therefore conceivable that a protocol using blocks which contain a certain fraction (less than half) of incorrect bits is better here, although the bit reconciliation effect between Alice and Bob is weaker in such a protocol. The advantage is that if Eve reliably knows one bit (or a small number of bits) of a block, she does not automatically know the whole block. We will show that the following protocol is in fact better here.

**Protocol B.** Bob randomly chooses a bit $C$, and computes a random $N$-bit block $[C_1, \ldots, C_N]$ such that $tN$ of the bits are equal to $C$, and $(1-t)N$ of the bits are equal to $\overline{C} := 1-C$ (where $t > 1/2$ is a parameter, and $tN$ is an integer). As in Protocol A, Bob computes $[C_1 \oplus Y_1, \ldots, C_N \oplus Y_N]$ and sends this block over the public channel. Alice computes $[(C_1 \oplus Y_1) \oplus X_1, \ldots, (C_N \oplus Y_N) \oplus X_N]$ and accepts only if this equals $[0, 0, \ldots, 0]$ or $[1, 1, \ldots, 1]$.

It is advantageous for Alice and Bob when Bob, and not Alice, is the sender of the bit in Protocol B.

The analysis of this protocol in Scenario 2 is quite technical, and is given in Appendix C. There we prove the following theorem, which also gives an upper bound for $r_X$ when given $\alpha$ and $r_Y$. We only mention here the surprising fact that typically, $t$ has to be chosen only slightly larger than $1/2$ (whereas it is obvious that the choice $t = 1/2$ is completely useless).

**Theorem 15** *Protocol B works in Scenario 2, and thus $S(X;Y||Z) > 0$, if*

$$r_X < \frac{2\left(1 - \frac{\alpha}{1-\alpha}\right)^2 (1 - r_Y)}{5 - 4r_Y} \tag{15}$$

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

$$r_X < (1 - r_Y)\left(1 - \frac{\alpha}{1-\alpha} - \frac{1 - r_Y}{2} - \frac{1}{8}\right) \tag{16}$$

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

Theorem 15 shows that, in Scenario 2, Protocol B is normally better than Protocol A, which is therefore not optimal. It is easy to see that the upper bounds of (15) and (16) are generally greater than the bound given by (14). We consider two examples.

If $r_Y$ is constant, and $\alpha \to 1/2$, then the bound given in (11) tends to 0 much faster than (15) (which applies in this situation). The bound of (13) is even negative. On the other hand, if $\alpha = 3/7$, and $r_Y \to 1$, then (11) is smaller than (16) (which applies here). The bound (13) is negative again.

Finally, we give a pessimistic bound on $r_X$ for Scenario 2. As in the previous section we derive a condition here for the fact that $I(X;Y{\downarrow}Z) = 0$.

**Lemma 16** *In Scenario 2,* $I(X;Y{\downarrow}Z) = 0$ *if*

$$r_X \geq \frac{(1-r_Y)(1-2\alpha)}{r_Y\alpha - (1-r_Y)(1-2\alpha) + 2(1-r_Y)\sqrt{1-2\alpha}\cdot(1-\alpha)} \ . \tag{17}$$

The proof of Lemma 16 is given in Appendix D. Of course the bound on $r_X$ given in (17) is greater than the bounds (14) and (15) (or (16), respectively) for all possible choices of $\alpha$ and $r_Y$.

Unfortunately, the conditions for $I(X;Y{\downarrow}Z) = 0$ and for the presented protocols for secret key agreement to be successful are not completely complementary. Hence it is still unknown whether $I(X;Y{\downarrow}Z) > 0$ implies $S(X;Y||Z) > 0$ in Scenario 2.

# References


[1] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, Experimental quantum cryptography, *Journal of Cryptology*, Springer Verlag, Vol. 5, No. 1, pp. 3-28, 1992.

[2] T. M. Cover and J. A. Thomas, Elements of information theory, Wiley Series in Telecommunications, 1992.

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

[4] W. Feller, An introduction to probability theory and its applications, 3rd edition, Vol. 1, Wiley International, 1968.

[5] M. J. Fischer and R. N. Wright, Bounds on secret key exchange using a random deal of cards, *Journal of Cryptology*, Springer Verlag, Vol. 9, No. 2, pp. 71-99, 1996.

[6] U. M. Maurer, Protocols for secret key agreement based on common information, *Advances in Cryptology - CRYPTO '92*, Lecture Notes in Computer Science, Springer-Verlag, Vol. 740, pp. 461-470, 1993.

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

[8] U. M. Maurer and S. Wolf, Towards characterizing when information-theoretic secret key agreement is possible, to appear in *Advances in Cryptology - ASIACRYPT '96*, Lecture Notes in Computer Science, Springer-Verlag, 1996.

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

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

## Appendix A: A graphical representation of information-theoretic quantities involving three random variables

Let $X$ and $Y$ be random variables. Then the quantities $H(XY)$, $H(X)$, $H(Y)$, $H(X|Y)$, $H(Y|X)$, and $I(X;Y)$ can be graphically represented (see Fig. 3). Note that the union of all inner regions corresponds to $H(XY)$.

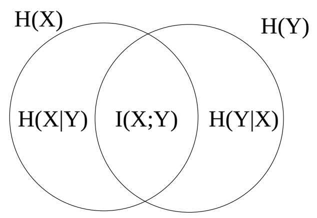


Figure 3: Two random variables

The case of three random variables $X$, $Y$, and $Z$ is more complicated. Assume first that $I(X;Y) \geq I(X;Y|Z)$. Let

$$R(X;Y;Z) := I(X;Y) - I(X;Y|Z)$$

(one can easily verify that $R(X;Y;Z)$ is symmetric in its three arguments). It is obvious that a simple graphical representation is possible (see Fig. 4).

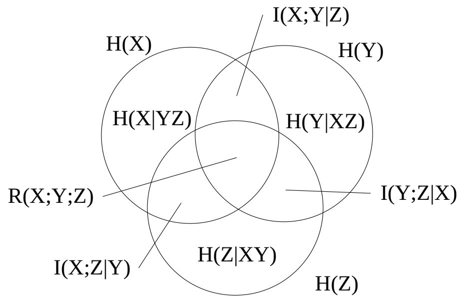


Figure 4: Three random variables

If $I(X;Y) < I(X;Y|Z)$, such a simple representation does not exist because $R(X;Y;Z) < 0$, and the overlapping region would be negative. For example when $X$ and $Y$ are independent bits, and $Z = X \oplus Y$, then $I(X;Y) = I(X;Z) = I(Y;Z) = 0$, but $I(XY;Z) = 1$.

We are now interested in a representation of $I(X;Y \downarrow Z)$. When given arbitrary $X$, $Y$, and $Z$ (i.e., even when $R(X;Y;Z) < 0$), we consider all the random variables $\overline{Z}$ that can be generated by sending $Z$ over a channel $P_{\overline{Z}|Z}$. Then $I(X;\overline{Z}) \leq I(X;Z)$ and $I(Y;\overline{Z}) \leq I(Y;Z)$. The particular $\overline{Z}$ which minimizes $I(X;Y|\overline{Z})$ fulfills also $R(X;Y;\overline{Z}) \geq 0$. This means that the above representation is always possible with this $\overline{Z}$ (even when this is not the case for $Z$). In this representation, $I(X;Y{\downarrow}Z)$ can be directly associated with one of the regions (see Fig. 5). The random variable $\overline{Z}$ is the one that maximally reduces the size of this region.

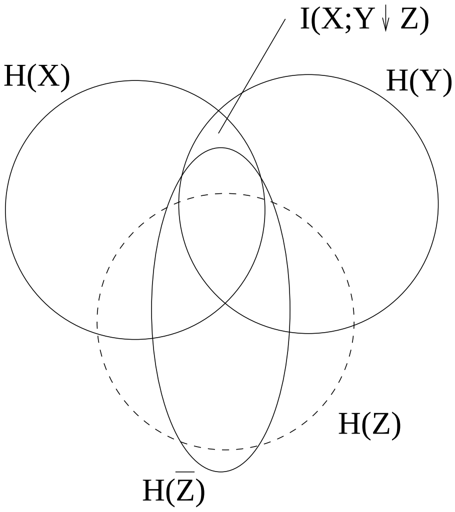


Figure 5: Visualization of $I(X;Y\downarrow Z)$

# Appendix B: Continuous random variables from independent channels

Here we show that the result of Theorem 6 also holds when the random variables that are generated from $R$ are not discrete. For example, this is the case in the realistic scenario where Eve receives her information about $R$ from a Gaussian channel.

Let $X$, $Y$, and $Z$ be continuous random variables, and let $f_{XYZ}, f_{X|Y}, \ldots$ be the probability density functions (we assume that such functions exist). The differential entropy of $X$, the conditional differential entropy of $X$ when given $Y$, and the mutual information between $X$ and $Y$ are defined as follows (see for example [2]):

$$\begin{aligned} h(X) &= -\int f_X \cdot \log f_X \, dx \\ h(X|Y) &= -\int f_{XY} \cdot \log f_{X|Y} \, dx \, dy \\ I(X;Y) &= h(X) - h(X|Y) = \int f_{XY} \cdot \log \frac{f_{XY}}{f_X \cdot f_Y} \, dx \, dy \end{aligned}$$

The conditional information between $X$ and $Y$ when given $Z$ can be defined in analogy to the case of discrete random variables as follows:

$$\begin{aligned} I(X;Y|Z) &= h(X|Z) - h(X|YZ) \\ &= \int f_{XYZ} \cdot \log \frac{f_{XY|Z}}{f_{X|Z} \cdot f_{Y|Z}} \, dx \, dy \, dz \\ &= \int I(X;Y|Z=z) \cdot f_Z(z) \, dz \; . \end{aligned}$$

As in Section 4 we assume that $X$, $Y$, and $Z$ are generated by sending a binary random variable $R$ over independent channels, i.e.,

$$f_{XYZ|R} = f_{X|R} \cdot f_{Y|R} \cdot f_{Z|R} \; , \tag{18}$$

or equivalently $f_{X|RYZ} = f_{X|R}$ , $f_{Y|RXZ} = f_{Y|R}$ , and $f_{Z|RXY} = f_{Z|R}$ .

**Theorem 17** *Let $R$ be a binary random variable, and let $X$, $Y$, and $Z$ be (real-valued) random variables with probability density function $f_{XYZ}$ and conditional density $f_{XY|Z}$. If (18) holds, then secret key agreement is possible, i.e., $S(X;Y||Z) > 0$, if $I(X;Y|Z) > 0$.*

**Proof:** We assume $I(X;Y|Z) > 0$, and conclude the following two statements:

1. We have $0 < P_R(0) < 1$, and Alice and Bob can generate binary random variables $\overline{X}$ and $\overline{Y}$ from $X$ and $Y$ with positive probability such that

$$P_{\overline{X}|R}(0,0) > P_{\overline{X}|R}(1,0) \tag{19}$$

and

$$P_{\overline{X}|R}(0,1) < P_{\overline{X}|R}(1,1) \tag{20}$$

(as well as the corresponding inequalities when replacing $\overline{X}$ by $\overline{Y}$) hold.

2. The random variable $Z$, together with some specific additional information $U$, corresponds to a random variable $\overline{Z}$ obtained by sending $R$ through a symmetric binary erasure channel with positive erasure probability.

Theorem 6 and Corollary 5 show that the statements 1 and 2 together imply $S(X;Y||Z) > 0$.

*Proof of 1.* Obviously $0 < P_R(0) < 1$ holds. We show that

$$\text{Prob}[f_{X|R=0}(x) \neq f_{X|R=1}(x)] > 0 \ . \tag{21}$$

Otherwise, if $f_{X|R=0}(x) = f_{X|R=1}(x)$ with probability 1, then

$$\begin{aligned} f_{XY|Z=z} &= f_{XY|R=0} \cdot P_{R|Z=z}(0) + f_{XY|R=1} \cdot P_{R|Z=z}(1) \\ &= f_{X|R=0} \cdot f_{Y|R=0} \cdot P_{R|Z=z}(0) + f_{X|R=1} \cdot f_{Y|R=1} \cdot P_{R|Z=z}(1) \\ &= f_{X|R=0} \cdot (f_{Y|R=0} \cdot P_{R|Z=z}(0) + f_{Y|R=1} \cdot P_{R|Z=z}(1)) = f_{X|Z=z} \cdot f_{Y|Z=z} \end{aligned}$$

with probability 1. Hence $I(X;Y|Z = z) = 0$ for all $z$, and $I(X;Y|Z) = 0$, which is a contradiction. Therefore (21) holds. We define

$$A_0 := \{x \mid f_{X|R=0}(x) > f_{X|R=1}(x)\}$$

and

$$A_1 := \{x \mid f_{X|R=0}(x) < f_{X|R=1}(x)\} \ .$$

Then $A_0$ and $A_1$ are disjoint measurable sets, with

$$P_{X|R=0}(A_0) > P_{X|R=1}(A_0)$$

and

$$P_{X|R=0}(A_1) < P_{X|R=1}(A_1)$$

(where $P_{X|R=0}(A_0)$ stands for $\int_{A_0} f_{X|R=0}\, dx$). If for example $P_{X|R=0}(A_0) = P_{X|R=1}(A_0)$, then

$$\int_{A_0} f_{X|R=0}(x) - f_{X|R=1}(x)\, dx = 0$$

(and the same holds for $A_1$, because the $f_{X|R=i}$ are densities of normed probability measures). It is a well-known fact from measure theory that the integral of a strictly positive function on a set with non-vanishing measure is also strictly positive, and hence $A_0$ and $A_1$ would be null sets, which is a contradiction to (21). For the random variable $Y$, two sets $B_0$ and $B_1$ can be defined similarly.

In analogy to the case of discrete random variables (see Section 3), Alice and Bob can obtain new random variables $\hat{X}$ and $\hat{Y}$ by restriction of the ranges of $X$ and $Y$ to $A_0 \cup A_1$ and $B_0 \cup B_1$, respectively, and send these random variables $\hat{X}$ and $\hat{Y}$ over two channels in order to generate binary random variables $\overline{X}$ and $\overline{Y}$ such that $\overline{X} = 0$ if $\hat{X} \in A_0$ and $\overline{X} = 1$ if $\hat{X} \in A_1$ (and analogously for $\overline{Y}$). It is obvious that (19) and (20) hold, as well as the corresponding inequalities for $\overline{Y}$.

*Proof of 2.* From

$$I(X;Y|Z) = \int I(X;Y|Z=z) \cdot f_Z(z)\, dz > 0$$

we conclude that there is a measurable set $A$ with

$$\mu(A) > 0 \tag{22}$$

(where $\mu$ denotes the Lebesgue-measure of $\mathbf{R}$) and

$$I(X;Y|Z=z) > 0 \quad \text{for all} \quad z \in A\ . \tag{23}$$

Because of (23) we have both $f_{Z|R=0}(z) > 0$ and $f_{Z|R=1}(z) > 0$ for all $z \in A$. (If for example $f_{Z|R=0}(z) = 0$, then

$$\begin{aligned} f_{XY|Z=z} &= f_{XY|R=0} \cdot P_{R|Z=z}(0) + f_{XY|R=1} \cdot P_{R|Z=z}(1) \\ &= f_{XY|R=1} = f_{X|R=1} \cdot f_{Y|R=1} = f_{X|Z=z} \cdot f_{Y|Z=z} \end{aligned}$$

and $I(X;Y|Z=z) = 0$.) For every $n$, let $A_n$ be the (measurable) set of all $z$ in $A$ such that $f_{Z|R=0}(z) \geq 1/n$ and $f_{Z|R=1}(z) \geq 1/n$. Then $A = \cup A_n$, and (22) implies

$$0 < \mu(A) = \mu(\cup A_n) \leq \sum_n \mu(A_n)\ .$$

We conclude that there exists $n_0$ such that $\mu(A_{n_0}) > 0$.

Let $U$ be a random variable such that $U = R$ with probability 1 if $z \notin A_{n_0}$, and with probability

$$\frac{f_{Z|R=i}(z) - 1/n_0}{f_{Z|R=i}(z)}$$

if $z \in A_{n_0}$ and $R = i$ (and such that otherwise, $U$ gives no information about $R$). The random variable $Z$, together with this side information $U$, corresponds to a random variable $\overline{Z}$, generated from $R$ by a symmetric binary erasure channel with erasure probability $\mu(A_{n_0})/n_0 > 0$. □

## Appendix C: Analysis of Protocol B in Scenario 2

Let the protocol parameter $t$ be fixed, and let

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

We first compute the conditional probability $\beta_N$ that Alice receives the bit sent by Bob incorrectly, given that she accepts:

$$\beta_N = \frac{\alpha^{tN}(1-\alpha)^{(1-t)N}}{(1-\alpha)^{tN}\alpha^{(1-t)N} + \alpha^{tN}(1-\alpha)^{(1-t)N}} \leq \left(\frac{\alpha}{1-\alpha}\right)^{N/(2K)}\ . \tag{24}$$

Eve's conditional error probability $\gamma_N$, given that Alice accepts, is lower bounded by $1/2$ times the probability that Eve receives exactly $sN$ of the $tN$ correct bits of Bob's block (more precisely, that she receives the corresponding realizations of $Y$ from the erasure channel, and erasure symbols for the other $(t-s)N$ realizations of $Y$ that also correspond to correct bits in Bob's block), and exactly the same number of incorrect bits, and that she learns nothing about Alice's block (i.e., about all the realizations of $X$) because she receives only erasure symbols from that channel. This is a lower bound for $\gamma_N$ because in this case, Eve's error probability for guessing Bob's bit is equal to $1/2$, and is independent of her strategy. This holds for all possible $s$, and hence the maximum of this probability, taken over all $0 \leq s \leq 1-t$, gives also a lower bound.

$$\gamma_N \geq \frac{1}{2}\cdot \max_{0\leq s\leq(1-t)} \left\{ \binom{tN}{sN}(r_Y)^{sN}(1-r_Y)^{(t-s)N}\binom{(1-t)N}{sN}(r_Y)^{sN}(1-r_Y)^{(1-t-s)N}\right\}\cdot(1-r_X)^N. \tag{25}$$

The next lemma gives a simpler lower bound that can be derived from the bound in (25) by determining its asymptotic behaviour.

**Lemma 18** *The lower bound (25) implies that*

$$\gamma_N^{2K/N} \geq 1 - \frac{1}{4K} - \frac{1}{16(1-r_Y)K} - 2Kr_X \tag{26}$$

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

**Proof:** First note that $r_Y/2 \leq 1-t$ means that $s := r_Y/2$ is a possible choice (in fact, this is the optimal choice). From Stirling's formula (see for example [4]) we can conclude that

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

for some constant $C$. The binomial coefficients in (25) can be replaced by this expression, and a straightforward computation leads to the following asymptotic behaviour of the lower bound on $\gamma_N$.

$$\begin{aligned}
\gamma_N^{2K/N} &\geq \left(1-\frac{1}{2K}\right)^K \cdot \left(1+\frac{1}{2K}\right)^K \cdot \cdots \\
&\qquad \cdots \cdot \left(1+\frac{1}{4(1-r_Y)K}\right)^{(1-r_Y)K} \cdot \left(1-\frac{1}{4(1-r_Y)K}\right)^{(1-r_Y)K} \cdot (1-r_X)^{2K} \\
&\geq \left(1-\frac{1}{4K}\right) \cdot \left(1-\frac{1}{16(1-r_Y)K}\right) \cdot (1-2Kr_X) \\
&\geq 1 - \frac{1}{4K} - \frac{1}{16(1-r_Y)K} - 2Kr_X
\end{aligned}$$

for sufficiently large $N$. $\square$

The bound (26) in the above lemma holds for all $K$ that correspond to a protocol parameter $t$ which satisfies $r_Y/2 \leq 1-t$. This is equivalent to

$$\frac{1}{2K} \leq 1 - r_Y \ . \tag{27}$$

The idea of the proof of Theorem 15 is to find the best choice for $K$ (i.e., the best choice of $t$ in Protocol B) with respect to the fixed parameters $\alpha$ and $r_Y$, and such that (27) holds. This optimal choice of $K$ leads to an upper bound on $r_X$, such that if $r_X$ is smaller than this bound, then Protocol B works for secret key agreement. This is exactly the upper bound stated in the theorem.

**Proof of Theorem 15:** According to (24) and (26), Protocol B (with parameter $t$) works for secret key agreement if

$$\gamma_N^{2K/N} \geq 1 - \frac{1}{4K} - \frac{1}{16(1-r_Y)K} - 2Kr_X > \frac{\alpha}{1-\alpha} \geq \beta_N^{2K/N} \ , \tag{28}$$

and if the condition (27) also holds. The reason is that (28) implies that Eve's error probability about the bit sent by Bob is asymptotically larger than Alice's error probability for $N \to \infty$. Lemma 8 states that this is sufficient for the possibility of secret key agreement by public discussion. Let $\delta := 1 - \alpha/(1-\alpha)$. Then (28) is satisfied if

$$r_X < \frac{\delta}{2K} - \frac{1}{8K^2}\left(1 + \frac{1}{4(1-r_Y)}\right) \ . \tag{29}$$

This bound depends on $K$, and from (29) we can determine the optimal choice for $K$ (and hence the optimal choice of the protocol parameter $t$). The only restriction is that the choice must be compatible with (27). It is easy to see that the expression on the right of (29) is maximal for

$$K = K_0 := \frac{1}{\delta} \cdot \left(\frac{1}{2} + \frac{1}{8(1-r_Y)}\right) \ .$$

It is somewhat surprising that if $\delta$ is small and $r_Y \approx 1$ (i.e., in a situation which is not advantageous to Alice and Bob) $K$ must be large, and this means that $t$ is only slightly greater than $1/2$ (whereas the choice $t = 1/2$ is obviously the worst possible choice). Taking $K = K_0$ is compatible with (27) if $\delta \geq 5/4 - r_Y$. Then the condition (29) is

$$r_X < \frac{2\delta^2(1-r_Y)}{5-4r_Y} \ .$$

If $\delta > 5/4 - r_Y$, the condition (27) is not fulfilled for $K = K_0$. For $K = K_0' := 1/(2-2r_Y)$ (the smallest choice for $K$ that satisfies (27)) the right side of (29) equals

$$(1-r_Y)\left(\delta - \frac{1-r_Y}{2} - \frac{1}{8}\right) \ .$$

$\square$

The bounds of Theorem 15 are not tight. The reason for this is that we have compared Alice's error probability with Eve's conditional error probability, given that Alice's bit is correct. Eve's error probability, given that *Alice accepts*, is greater, because, given that *Alice does not receive the correct bit*, it is more likely that Eve's bit is also incorrect. However, it is difficult to compute the exact error probability because it is complicated to determine Eve's optimal strategy of guessing the bit.

It is conceivable that the above results can be improved using a block protocol where both Alice and Bob (and not only Bob) have a block that is not composed by $N$ times the same bit. Such a protocol appears to be much more difficult to analyze.

## Appendix D: Proof of Lemma 16

We show that if (17) is satisfied, then a channel, characterized by $P_{\overline{Z}|Z}$, can be constructed such that $I(X;Y|\overline{Z}) = 0$. The only $z \in \mathcal{Z}$ with $I(X;Y|Z=z) > 0$ is $z = [\Delta,\Delta]$, and the event $Z = [\Delta,\Delta]$ has probability $(1-r_X)(1-r_Y)$. The idea of the proof is to split this into three events $\overline{Z} = \Delta_1$, $\overline{Z} = \Delta_2$, and $\overline{Z} = \Delta_3$ (where $\overline{Z} = \Delta_i$ can also occur if $Z \neq [\Delta,\Delta]$) such that $I(X;Y|\overline{Z} = \Delta_i) = 0$ for $i = 1,2,3$. More precisely, the random variable $\overline{Z}$ will be defined such that $\overline{Z} = \Delta_1$ is possible not only if $Z = [\Delta,\Delta]$, but also if $Z = [0,1]$ and $Z = [1,0]$, whereas $\overline{Z} = \Delta_2$ is also possible if $Z = [0,\Delta]$ and $Z = [\Delta,0]$, and finally $\overline{Z} = \Delta_3$ if $Z = [1,\Delta]$ and $Z = [\Delta,1]$. We determine the maximal possible probability of $Z = [\Delta,\Delta]$ which allows that this event can completely be splitted.

We define the random variable $\overline{Z}$ as follows, by giving the joint distribution with $Z$:

$$\begin{aligned}
P_{\overline{Z}Z}(\Delta_1,[\Delta,\Delta]) &= \mu \cdot \frac{r_X r_Y \alpha}{1-2\alpha} \\
P_{\overline{Z}Z}(\Delta_1,[0,1]) &= P_{\overline{Z}Z}(\Delta_1,[1,0]) \;=\; \mu \cdot P_Z([0,1]) \;=\; \mu \cdot P_Z([1,0]) \;=\; \mu \cdot \frac{r_X r_Y \alpha}{2} \\
P_{\overline{Z}Z}(\Delta_2,[\Delta,\Delta]) &= P_{\overline{Z}Z}(\Delta_3,[\Delta,\Delta]) \;=\; \mu \cdot r_X(1-r_Y)\left(\frac{1-\alpha}{\sqrt{1-2\alpha}} - 1\right) \\
P_{\overline{Z}Z}(\Delta_2,[\Delta,0]) &= P_{\overline{Z}Z}(\Delta_2,[0,\Delta]) \;=\; P_{\overline{Z}Z}(\Delta_3,[\Delta,1]) \;=\; P_{\overline{Z}Z}(\Delta_3,[1,\Delta]) \;=\; \ldots \\
\ldots &= \mu \cdot P_Z([0,\Delta]) \;=\; \mu \cdot \frac{r_X(1-r_Y)}{2}
\end{aligned}$$

and $\overline{Z} = Z$ otherwise. The parameter $0 \leq \mu \leq 1$ is such that

$$\sum_{i=1}^{3} P_{\overline{Z}Z}(\Delta_i,[\Delta,\Delta]) = P_Z([\Delta,\Delta]) \ .$$

Note that $\mu > 1$ is not possible. It is easy to see that the random variable $\overline{Z}$ can be obtained by sending $Z$ over a channel specified by some conditional probability distribution $P_{\overline{Z}|Z}$. We show that $I(X;Y|\overline{Z} = \Delta_i) = 0$ for $i = 1,2,3$. For $i = 1$ this follows from

$$P_{XY\overline{Z}}(0,0,\Delta_1) = P_{XY\overline{Z}}(1,1,\Delta_1) = \mu \cdot \frac{r_X r_Y \alpha(1-\alpha)}{2(1-2\alpha)}$$

and

$$P_{XY\overline{Z}}(0,1,\Delta_1) = P_{XY\overline{Z}}(1,0,\Delta_1) = \mu \cdot \frac{r_X r_Y \alpha}{1-2\alpha} \cdot \frac{\alpha}{2} + \mu \cdot \frac{r_X r_Y \alpha}{2} = \mu \cdot \frac{r_X r_Y \alpha(1-\alpha)}{2(1-2\alpha)} \ .$$

For $i = 2$ and $i = 3$ one can easily verify that

$$P_{XY\overline{Z}}(0, 0, \Delta_i) \cdot P_{XY\overline{Z}}(1, 1, \Delta_i) = P_{XY\overline{Z}}(0, 1, \Delta_i) \cdot P_{XY\overline{Z}}(1, 0, \Delta_i)$$

holds, and that this implies that $X$ and $Y$ are statistically independent, given that $\overline{Z} = \Delta_i$. If $\overline{z} \notin \{\Delta_1, \Delta_2, \Delta_3\}$ then $I(X;Y|\overline{Z} = \overline{z}) = 0$ obviously holds, and we conclude $I(X;Y|\overline{Z}) = 0$ and $I(X;Y{\downarrow}Z) = 0$.

The maximal probability $P_Z([\Delta, \Delta])$ such that the event $Z = [\Delta, \Delta]$ can be completely splitted into $\overline{Z} = \Delta_i$ as above is the sum of the probabilities $P_{\overline{Z}Z}(\Delta_i, [\Delta, \Delta])$ $(i = 1, 2, 3)$ with $\mu = 1$. Thus the described construction of $\overline{Z}$ works if

$$\frac{r_X r_Y \alpha}{1 - 2\alpha} + 2r_X(1 - r_Y) \cdot \left( \frac{1 - \alpha}{\sqrt{1 - 2\alpha}} \right) \geq P_Z([\Delta, \Delta]) = (1 - r_X)(1 - r_Y) \ , \tag{30}$$

and this is equivalent to (17). □

Note that the condition given in the lemma is sufficient, but not necessary for $I(X;Y{\downarrow}Z) = 0$. For certain $r_X$, $r_Y$, and $\alpha$, a more complicated cascade, where $Z = [0, \Delta]$ and $Z = [\Delta, 0]$ are *not* transformed symmetrically to $\overline{Z} = \Delta_2$, leads to a slightly better bound.