1

Unconditionally Secure Key Agreement and the Intrinsic Conditional Information

Ueli Maurer, Senior Member, IEEE, and Stefan Wolf

The authors are with the Department of Computer Science, Swiss Federal Institute of Technology (ETH Zürich), CH-8092 Zürich, Switzerland. E-mail: {maurer,wolf}@inf.ethz.ch .

The second author was supported by the Swiss National Science Foundation (SNF), grant no. 20–42105.94.

Some of the results of this paper have been presented at ASIACRYPT '96, Kyongju, Korea, 1996 [16] and at ISIT '97, Ulm, Germany, 1997 [15].

Abstract

This paper is concerned with secret-key agreement by public discussion. Assume that 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 between X and Y when given Z, denoted by $I(X;Y{\downarrow}Z)$, which is an upper bound on $S(X;Y||Z)$. The special scenarios are analyzed 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 two scenarios in which Z is generated by sending X and Y over erasure channels. In the first two scenarios it can be shown that the secret key rate is strictly positive if and only if $I(X;Y{\downarrow}Z)$ is strictly positive. For the third 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.

I. Introduction

Perfectly secure key agreement has been studied recently by several authors [19], [6], [13], [2], [9], [7], [16]. Two possible approaches are based on quantum cryptography (e.g., see [2]) 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 [13] and in [16] it has been illustrated how such noise 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. Noise in communication channels has also been shown useful in other respects. In [5] and [4] for instance it was shown how to realize various cryptographic primitives, such as bit commitment or oblivious transfer, based on a noisy channel.

It is a classical cryptographic problem to transmit a message M from a sender (referred to as Alice) to a receiver (Bob) over an insecure communication channel such that an adversary (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 [17], 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 adversary 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 [13] this pessimism has been relativized by pointing out that Shannon's apparently innocent assumption that, except for the secret key, the opponent 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 adversary has some minimal uncertainty about the signal received by the legitimate receivers.

Wyner [19] and subsequently Csiszár and Körner [6] considered a scenario in which the opponent 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 [13] that this assumption is not needed 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 [11] (as a strengthened version of the definition given in [13]) as follows. We assume in the following that the distribution P_{XYZ} is publicly known.

Definition 1: The *secret key rate of X and Y with respect to Z*, denoted by $S(X;Y||Z)$, is the maximum rate at which Alice and Bob can agree on a secret key S in such a way that the amount of information that Eve obtains about S is arbitrarily small. In other words, it is the maximal R such that for every $\varepsilon > 0$ 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, compute the same key S with probability at least $1 - \varepsilon$, satisfying

$$I(S; VZ^N) \leq \varepsilon \ , \tag{1}$$

$$\frac{1}{N} H(S) \geq R - \varepsilon \ , \tag{2}$$

and[1]

$$H(S) \geq \log |\mathcal{S}| - \varepsilon \ . \tag{3}$$

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

Remark. Note that this definition corresponds to the *strong* secret key rate as introduced in [11]. In contrast to all earlier definitions of a rate made in the context of secret transmissions (e.g., of the secrecy capacity in Wyner's [19] and Csiszár and Körner's [6] models), not only the *rate* at which Eve obtains information about the secret key, but the *total amount* of information she learns about this key, must be arbitrarily small. However, it will be shown in a final version of [11] that the secret key rates with respect to the weaker and stronger definitions are equal. Hence we can restrict ourselves to the stronger, more satisfactory definition.

Remark. The problem of secret-key agreement has also been studied for the case where the channel connecting Alice and Bob is not authentic, i.e., the adversary is also able to modify or insert messages. It has been shown in [10], [14], [18] that secret-key agreement can even be possible in this case (if the distribution P_{XYZ} satisfies certain properties).

The following lower bound on $S(X;Y||Z)$ is proved in [11] (and first in [13] for the weaker definition), and follows from a result by Csiszár and Körner [6].

Theorem 1: For all distributions P_{XYZ} we have

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

It has been first shown by an example in [13] that the secret key rate $S(X;Y||Z)$ can be strictly positive even when both $I(X;Z) > I(X;Y)$ and $I(Y;Z) > I(Y;X)$ hold.

[1] Throughout the paper, all logarithms are to the base 2.

We give a brief outline of the rest of this paper. In Section 2 we define a new conditional information measure and show that this measure gives an improved upper bound on the secret key rate. In Section 3 we formulate some fundamental properties of the secret key rate. Sections 4 and 5 address the problem whether secret-key agreement is always possible when this new upper bound is strictly positive. We consider this for the cases where X, Y, and Z are generated by sending a binary random variable over independent channels (Scenario 1 in Section 4), and where Z is generated by sending X and Y over erasure channels (Scenarios 2 and 3 in Section 5). For Scenarios 1 and 2 it is shown that secret-key agreement is possible if the intrinsic conditional information is positive. For a generalized version of Scenario 2, in which Eve obtains both Alice's and Bob's information with a certain probability $1-\delta$, the new information measure is shown to be closely related to δ and to a new, natural quantity measuring the deviation of Alice's and Bob's information from statistically independent information. For Scenario 3 finally, we show that a new protocol is more powerful than the previously known protocols.

II. The Intrinsic Conditional Mutual Information

A. *Motivation and Definition*

The following upper bound on the secret key rate was proved in [13]:

$$S(X;Y||Z) \leq \min\{I(X;Y), I(X;Y|Z)\} \ . \tag{4}$$

Trying to reduce the quantity $I(X;Y|Z)$ in this 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 $\overline{Z}$. Obviously,

$$S(X;Y||Z) \leq S(X;Y||\overline{Z}) \leq I(X;Y|\overline{Z}) \tag{5}$$

holds for every such $\overline{Z}$. A similar bound also appeared in [1]. Inequality (5) motivates the following definition of the intrinsic conditional mutual information between X and Y when given Z, which is the infimum of $I(X;Y|\overline{Z})$, taken over all discrete 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 be-*

tween X and Y when given Z, denoted by $I(X;Y{\downarrow}Z)$, is

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

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

The intrinsic conditional information satisfies the following inequalities:

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

where $\overline{Z}$ is generated by sending Z over an arbitrary channel.

Theorem 2: For arbitrary random variables X, Y, and Z, we have

$$S(X;Y||Z) \leq I(X;Y{\downarrow}Z) \ . \tag{6}$$

Proof: Bound (6) follows from the definition of the intrinsic information and from inequality (5). ■

Theorem 2 implies in particular that secret-key agreement can be possible only if $I(X;Y{\downarrow}Z) > 0$. The following simple example shows that the intrinsic conditional information can be equal to 0, and secret-key agreement is hence impossible, even when both $I(X;Y|Z) > 0$ and $I(X;Y) > 0$. Thus the bound (6) is a strict improvement of bound (4). 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 $I(X;Y) = 3/2$ 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 conditional 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 (possibly reduced by giving Z), but *not* the *additional* information between X and Y brought in by Z.

B. A Graphical Representation

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 Figure 1). Note that the union of all inner regions corresponds to $H(XY)$.

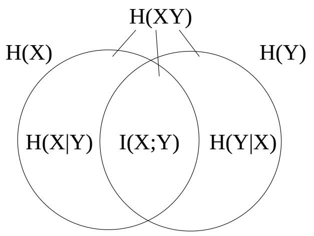

Fig. 1. Two random variables

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 Figure 2).

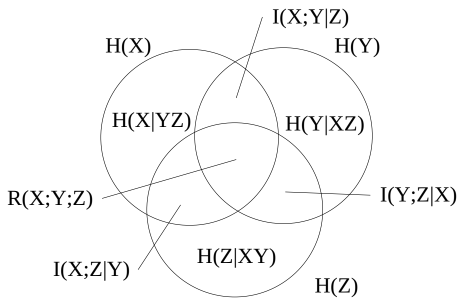

Fig. 2. Three random variables

Note that if $I(X;Y) < I(X;Y|Z)$, such a representation is also possible, but some regions are negative. For example when X and Y are independent symmetric bits, and $Z = X \oplus Y$, then $I(X;Y) = I(X;Z) = I(Y;Z) = 0$, but $I(XY;Z) = 1$. For a systematic treatment of such representations, see [20].

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 $P_{\overline{Z}|Z}$. Note that $I(X;\overline{Z}) \leq I(X;Z)$ and $I(Y;\overline{Z}) \leq I(Y;Z)$ hold for such random variables $\overline{Z}$. The particular $\overline{Z}$ which minimizes $I(X;Y|\overline{Z})$ fulfills $R(X;Y;\overline{Z}) \geq 0$. (This means that there are no negative regions in the graphical representation, although this may be the case for X, Y, and Z). The quantity $I(X;Y{\downarrow}Z)$ can be directly associated with one of the regions (see Figure 3). The random variable $\overline{Z}$ is the one that maximally reduces the size of this region.

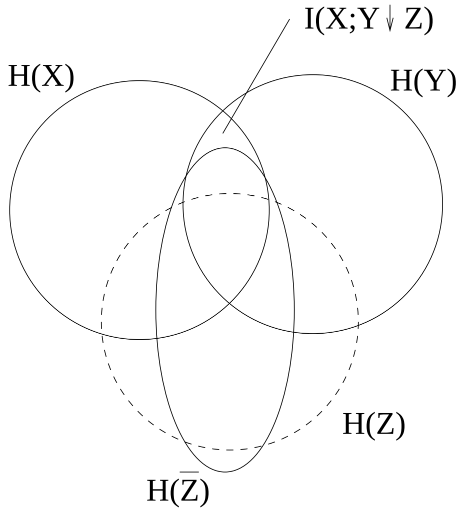

Fig. 3. Visualization of $I(X;Y{\downarrow}Z)$

III. Secret-Key Agreement with General Random Variables

As stated above, the intrinsic conditional mutual information $I(X;Y{\downarrow}Z)$ defined above gives a new upper bound on 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.

Conjecture 1: Let P_{XYZ} be such that $I(X;Y{\downarrow}Z) > 0$. Then $S(X;Y||Z) > 0$.

An even stronger conjecture would be that $S(X;Y||Z) = I(X;Y \downarrow Z)$ holds for every distribution P_{XYZ}. In the following sections we prove the validity of Conjecture 1 for several special scenarios. It is a fundamental open problem to prove or disprove the conjecture for the general case.

As a preparation for the analysis in the following sections we prove some important and basic properties of the secret key rate. The three lemmas below are very intuitive and follow quite directly from the definition of the secret key rate. Lemma 3 states that Alice and Bob cannot increase the secret key rate by ignoring certain realizations of the random variables X and Y. We say that Alice and Bob obtain new random variables by *restriction of the ranges* if they discard realizations that do not lie in certain subsets $\hat{\mathcal{X}}$ and $\hat{\mathcal{Y}}$ of $\mathcal{X}$ and $\mathcal{Y}$. Lemma 4 states that processing X and Y cannot increase the secret key rate.

Lemma 3: 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 Ω is the event that $X \in \hat{\mathcal{X}}$ and $Y \in \hat{\mathcal{Y}}$, and if $P(\Omega)$ is its probability, then the joint distribution of $\hat{X}$ and $\hat{Y}$ with Z is defined as follows:

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

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

$$S(X;Y||Z) \geq P(\Omega) \cdot S(\hat{X};\hat{Y}||Z) \ . \tag{7}$$

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

Proof: The secret key rate $S(X;Y||Z)$ is the maximum key-generation rate, taken over all possible protocols between Alice and Bob. One possible strategy is to restrict the ranges of their random variables. With probability $P(\Omega)$, they both receive random variables $\hat{X}$ and $\hat{Y}$, respectively, and inequality (7) follows. ■

Lemma 4: 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

$$S(\overline{X};\overline{Y}||Z) \leq S(X;Y||Z) \ .$$

Proof: As in the proof of Lemma 3, the statement follows because it is one of the possible strategies for Alice and Bob to send X and Y over two channels, and because the secret key rate is the maximum key generation rate taken over all possible protocols. ■

Lemma 5 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 5: Let X, Y, Z, and U be arbitrary random variables. Then

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

Proof: Obviously, only condition (1) in the definition of the secret key rate is affected. Because $I(S;V[Z,U]^N) \leq \varepsilon$ implies $I(S;VZ^N) \leq \varepsilon$, the lemma follows. ■

Theorem 6 is an immediate consequence of the three Lemmas 3, 4, and 5.

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 $P_{\overline{Y}|\hat{Y}}$.

Theorem 6: 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 $S(X;Y||Z) > 0$.

IV. Noisy Versions of a Binary Signal

The first special scenario, which we analyze completely in this section, is defined as follows.

Scenario 1. Let R be an arbitrary binary random variable, and let X, Y, and Z be arbitrary discrete random variables, generated by sending R over independent channels C_A, C_B, and C_E, i.e.,

$$P_{XYZ|R} = P_{X|R} \cdot P_{Y|R} \cdot P_{Z|R} \ . \tag{8}$$

In other words, X, Y, and Z are statistically independent when given R. The following

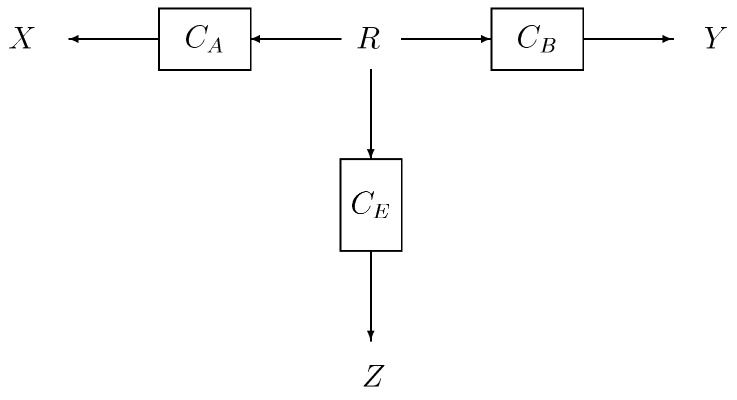

Fig. 4. Scenario 1.

is a different but equivalent characterization of Scenario 1. There exist $0 \leq \lambda \leq 1$ and probability distributions $P_X^{(1)}$, $P_X^{(2)}$, $P_Y^{(1)}$, $P_Y^{(2)}$, $P_Z^{(1)}$, and $P_Z^{(2)}$ such that

$$P_{XYZ} = \lambda \cdot P_X^{(1)} \cdot P_Y^{(1)} \cdot P_Z^{(1)} + (1-\lambda) \cdot P_X^{(2)} \cdot P_Y^{(2)} \cdot P_Z^{(2)},$$

i.e., P_{XYZ} is a weighted sum of two different distributions over the set $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ that both correspond to independent random variables with ranges $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$. The results of this section hold for all distributions with this property.

The main result of this section is the following theorem, which characterizes completely the cases for which $S(X;Y||Z) > 0$ in Scenario 1, i.e., for which secret-key agreement is possible in principle, and which implies that Conjecture 1 is true in this case.

Theorem 7: In Scenario 1, the following conditions are equivalent.

(A) $I(X;Y|Z) > 0$,

(B) $S(X;Y||Z) > 0$,

(C) $I(X;Y{\downarrow}Z) > 0$.

The proof that *(A)* implies *(B)* is subdivided into several steps stated below as lemmas. We begin with the special case where R is a symmetric binary random variable and all three channels are binary symmetric. This special result is not necessary for the proof of Theorem 7, but we show it in order to present the protocol and some estimates that will be useful later. In Appendix A we prove a result similar to Theorem 7 for *continuous* random variables X, Y, and Z generated from independent binary-input channels.

A. Binary Symmetric Channels

Let us first consider the following special case of Scenario 1. Let $P_R(0) = P_R(1) = 1/2$ and consider three binary symmetric channels C_A, C_B, and C_E with bit error probabilities α, β, and ε, respectively, i.e., we have

$$P_{X|R}(0,0) = 1-\alpha\ , \quad P_{Y|R}(0,0) = 1-\beta\ , \quad \text{and} \quad P_{Z|R}(0,0) = 1-\varepsilon\ ,$$

where $0 \leq \alpha < 1/2$, $0 \leq \beta < 1/2$, and $0 < \varepsilon \leq 1/2$. We can assume here that $\alpha = \beta$, i.e., that Alice's and Bob's channels are identical. If for example $\alpha < \beta$, Alice can cascade her channel with another binary symmetric channel with error probability $(\beta-\alpha)/(1-2\alpha)$ to obtain error probability β. (In this particular case of binary symmetric channels it is not even necessary to assume $\alpha = \beta$. The statement of Lemma 8 also holds if $\alpha \neq \beta$ when the party with the greater error probability is the sender and the other party is the receiver in Protocol A described below.)

Scenario 1.1. The distribution P_{RXYZ} is defined by the fact that all the random variables are binary and symmetric, and by

$$P_{X|R}(0,0) = P_{Y|R}(0,0) = 1-\alpha \ \text{ and } \ P_{Z|R}(0,0) = 1-\varepsilon\ ,$$

where $0 \leq \alpha < 1/2$ and $0 < \varepsilon \leq 1/2$.

In Scenario 1.1, Alice can send a randomly chosen bit C to Bob by the following protocol, which was already described in [13].

Protocol A. Let N be fixed. Alice 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 exactly 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 the only two codewords $[0,0,\ldots,0]$ and $[1,1,\ldots,1]$.

It is obvious that Eve's optimal strategy for guessing C 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. Note that although it is not the adversary's ultimate goal to guess the bits C sent by Alice, Lemma 10 below leads, from Eve's error probability when guessing these bits, to a lower bound on the secret key rate.

Protocol A is computationally efficient, but it is not efficient in terms of the size of the generated secret key. There exist variants of the protocol, using parity checks instead of repeat codes, that are much more efficient in terms of the achievable key generation rate [12].

We show first that for all possible choices of α and ε, in particular even if Eve's channel is superior to both Alice's and Bob's channel, Bob's error probability β_N about the bit sent by Alice decreases, for $N \longrightarrow \infty$, asymptotically faster than Eve's error probability γ_N when she uses the optimal strategy for guessing this bit. (Note that γ_N is an *average* error probability, and that for a particular realization, Eve's error probability will typically be smaller or greater than γ_N.)

Lemma 8: If Protocol A is used in Scenario 1.1, there exist real-valued positive constants b and c with $b < c$ such that $\beta_N \leq b^N$ and $\gamma_N \geq c^N$ for sufficiently large N.

For the proof of Lemma 8, we need the following fact on binomial coefficients.

Lemma 9: For sufficiently large N, we have

$$\binom{N}{N/2} \geq \frac{1}{\sqrt{2\pi N}} \cdot 2^N \ .$$

Proof: Stirling's formula (see for example [8]) states that $n!/((n/e)^n \cdot \sqrt{2\pi n}) \to 1$ for $n \to \infty$, and thus we have for sufficiently large even N

$$\binom{N}{N/2} = \frac{N!}{((N/2)!)^2} \geq \frac{1}{2} \cdot \frac{N^N \cdot \sqrt{2\pi N} \cdot e^N}{e^N \cdot (N/2)^N \cdot \pi N} = \frac{1}{\sqrt{2\pi N}} \cdot 2^N \ ,$$

and this concludes the proof. ■

Proof of Lemma 8: Let α_{rs} ($r, s \in \{0,1\}$) be the probability that the single bit 0 sent by Alice is received by Bob as r and by Eve as s. Then

$$\begin{aligned} \alpha_{00} &= (1-\alpha)^2(1-\varepsilon) + \alpha^2\varepsilon \ , \\ \alpha_{01} &= (1-\alpha)^2\varepsilon + \alpha^2(1-\varepsilon) \ , \\ \alpha_{10} &= \alpha_{11} = \alpha(1-\alpha) \ . \end{aligned}$$

Let $p_{a,N}$ be the probability that Bob accepts the message sent by Alice. If we assume (without loss of generality) that N is even, then

$$\beta_N = \frac{1}{p_{a,N}} \cdot (\alpha_{10} + \alpha_{11})^N = \frac{1}{p_{a,N}} \cdot (2\alpha - 2\alpha^2)^N \tag{9}$$

$$\gamma_N \geq \frac{1}{2} \cdot \frac{1}{p_{a,N}} \cdot \binom{N}{N/2} \alpha_{00}^{N/2} \alpha_{01}^{N/2} \; . \tag{10}$$

The last expression is half of the probability that Bob receives the correct codeword and that Eve receives the same number of 0's and 1's, given that Bob accepts. Note that (10) gives a lower bound on Eve's average error probability when guessing C for all possible strategies because in this symmetric case, Eve obtains no information about the bit C, and half of the guesses will be incorrect. From Lemma 9 we conclude that

$$\gamma_N \geq \frac{1}{2} \cdot \frac{1}{p_{a,N}} \cdot \frac{1}{\sqrt{2\pi N}} \cdot 2^N \cdot \sqrt{\alpha_{00}\alpha_{01}}^N = \frac{C}{\sqrt{N}} \cdot \frac{\left(2\sqrt{\alpha_{00}\alpha_{01}}\right)^N}{p_{a,N}}$$

for some constant C, and for sufficiently large N. For $0 < \varepsilon \leq 1/2$ we have

$$\sqrt{\alpha_{00}\alpha_{01}} = \sqrt{(1 - 2\alpha + \alpha^2 - \varepsilon + 2\alpha\varepsilon)(\alpha^2 - 2\alpha\varepsilon + \varepsilon)} > \alpha - \alpha^2. \tag{11}$$

For $\varepsilon = 0$ equality holds in (11), and for $\varepsilon > 0$ the greater factor of the product under the square root is decreased by the same value by which the smaller factor is increased. Hence the square root of this product is greater than $\alpha - \alpha^2$. (For $\varepsilon = 1/2$ the factors are equal, and the left hand side of (11) is maximal, as expected.) Because

$$(1 - 2\alpha + 2\alpha^2)^N \leq p_{a,N} = (1 - 2\alpha + 2\alpha^2)^N + (2\alpha - 2\alpha^2)^N < 2 \cdot (1 - 2\alpha + 2\alpha^2)^N \; , \tag{12}$$

we conclude that $\beta_N \leq b^N$ and $\gamma_N \geq c^N$ for sufficiently large N, $b = (2\alpha - 2\alpha^2)/(1 - 2\alpha + 2\alpha^2)$, and $c = 2\sqrt{\alpha_{00}\alpha_{01}}/(1 - 2\alpha + 2\alpha^2) - \delta$ (where δ can be made arbitrarily small for sufficiently large N). From the above, we conclude that $c > b$ for sufficiently small δ. ■

The fact that Eve has a greater error probability than Bob when guessing C does not automatically imply that Eve has a greater uncertainty about this bit in terms of Shannon entropy, and hence that $S(X;Y||Z) > 0$. The next lemma together with Lemma 8 nevertheless implies that the secret key rate is positive in Scenario 1.1.

Lemma 10: 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

(where $X^N = [X_1, \ldots, X_N]$) and C (and possibly some random bits) such that with some probability $p_{a,N} > 0$, Bob (who knows M and Y^N) publicly 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 γ_N is at least c^N for some $c > b$ and for sufficiently large N, then $S(X;Y||Z) > 0$.

Proof: According to Theorem 1 is 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

$$I(\hat{X};\hat{Y}) - I(\hat{X};\hat{Z}) > 0 \tag{13}$$

with $\hat{Z} = [Z^N, U]$, where U 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 (13) holds for sufficiently large N. If Bob accepts then

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

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

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

where $p_{E,\hat{z}}$ is the probability of guessing C incorrectly with the optimal strategy given that $\hat{Z} = \hat{z}$. Note that $p_{E,\hat{z}} \leq 1/2$, hence $h(p_{E,\hat{z}}) \geq p_{E,\hat{z}}$, for all $\hat{z}$. Given that Bob publicly rejects, we have $H(\hat{X}|\hat{Y}) = H(\hat{X}|\hat{Z}) = H(\hat{X}|U) = 0$. From $p_{a,N} > 0$ we conclude that $I(\hat{X};\hat{Y}) - I(\hat{X};\hat{Z}) > 0$. ∎

B. General Binary-Input Channels and the Proof of Theorem 7

First we show that the above results hold even when Eve knows R *precisely* with a certain probability smaller than 1. This is the case if Z is generated from R by a binary *erasure* channel instead of a binary symmetric channel, i.e., if Z is either equal to a special erasure symbol Δ, or else $Z = R$.

Scenario 1.2. Let R, X, and Y be as in Scenario 1.1, but let Z be generated from R by a (possibly asymmetric) binary erasure channel (with erasure symbol Δ) C_E^*, independent of the pair (C_A,C_B), and with transition probabilities $P_{Z|R}(\Delta,0) = \delta > 0$, $P_{Z|R}(0,0) = 1-\delta$, $P_{Z|R}(\Delta,1) = \delta' > 0$, and $P_{Z|R}(1,1) = 1-\delta'$.

Lemma 11: If Protocol A is used in Scenario 1.2, there exist real-valued positive constants b and c with $b < c$ such that $\beta_N \leq b^N$ and $\gamma_N \geq c^N$ for sufficiently large N.

Proof: We show first that we can assume without loss of generality that C_E^* is symmetric. Let $\delta < \delta'$, and let an oracle be given that tells Eve the correct bit R with probability $(\delta' - \delta)/\delta'$ if $R = 1$ and $Z = \Delta$. According to Lemma 5, the additional information U provided by this oracle cannot increase the secret key rate. The random variable Z, together with the oracle, is equivalent to a random variable generated from R by a symmetric binary erasure channel with erasure probability δ, and which is independent of the pair (C_A,C_B).

If $\delta = 1$, the lemma is trivial. Let $\delta < 1$, and let $0 < \rho < \min\{\delta, 1-\delta\}$. For sufficiently large N, the probability that the number of bits (out of N bits) known to Eve is even and lies between $(1-\delta-\rho)N$ and $(1-\delta+\rho)N$ is at least $1/3$. We can assume without loss of generality that N and $(1-\delta-\rho)N$ are even integers. (Otherwise, ρ can be chosen slightly smaller in order to fulfill this.) We give a lower bound on Eve's average error probability γ_N about the bit sent by Alice, given that Bob accepts. As in the proof of Lemma 8 we obtain a lower bound on Eve's error probability γ_N by taking a (small) part of all positive terms adding up to γ_N. We have

$$\begin{aligned}
\gamma_N \quad &\geq \quad \frac{1}{2} \cdot \frac{(1-2\alpha+2\alpha^2)^N}{p_{a,N}} \cdot \frac{1}{3} \cdot \binom{(1-\delta-\rho)N}{(1-\delta-\rho)N/2} \\
&\quad \cdot \left[\frac{(1-\alpha)^2}{(1-\alpha)^2+\alpha^2}\right]^{(1-\delta+\rho)N/2} \left[\frac{\alpha^2}{(1-\alpha)^2+\alpha^2}\right]^{(1-\delta+\rho)N/2} \\
&\geq \quad \frac{1}{2} \cdot \frac{1}{p_{a,N}} \cdot \frac{1}{3} \cdot \frac{1}{\sqrt{2\pi(1-\delta-\rho)N}} \cdot \left[(1-2\alpha+2\alpha^2)^{\delta-\rho} 2^{1-\delta-\rho} (\alpha-\alpha^2)^{1-\delta+\rho}\right]^N
\end{aligned}$$

for sufficiently large N. Here we have made use of Lemma 9. The first expression is $1/2$ times a lower bound on the probability that Bob receives the correct codeword, that Eve knows an even number of bits which lies between $(1-\delta-\rho)N$ and $(1-\delta+\rho)N$, and that she receives the same number of 0's and 1's in her reliable bits, given that Bob accepts. In this case, Eve obtains no information about the bit sent by Alice. The expressions $(1-\alpha)^2/((1-\alpha)^2+\alpha^2)$ and $\alpha^2/((1-\alpha)^2+\alpha^2)$ are the probabilities that $R=X$ and $R\neq X$, respectively, given that $X=Y$. Bob's error probability, given that he accepts, is $\beta_N=(2\alpha-2\alpha^2)^N/p_{a,N}$. For sufficiently small (positive) ρ we have

$$(1-2\alpha+2\alpha^2)^{\delta-\rho}2^{1-\delta-\rho}(\alpha-\alpha^2)^{1-\delta+\rho}>2\alpha-2\alpha^2$$

because $\delta>0$ and $1-2\alpha+2\alpha^2>2\alpha-2\alpha^2$. The lemma is proved because of inequality (12). ■

We now consider the general Scenario 1. The following lemma states equivalent characterizations of the condition $I(X;Y|Z)>0$.

Lemma 12: In Scenario 1, the following three conditions are equivalent.

(i) $I(X;Y|Z)>0$.

(ii) $I(X;R)>0$, $I(Y;R)>0$, and $H(R|Z)>0$.

(iii) There exist $x,x'\in\mathcal{X}$ such that

$$P_{X|R}(x,0)>P_{X|R}(x,1)\quad\text{and}\quad P_{X|R}(x',0)<P_{X|R}(x',1)\ , \tag{14}$$

there exist $y,y'\in\mathcal{Y}$ such that

$$P_{Y|R}(y,0)>P_{Y|R}(y,1)\quad\text{and}\quad P_{Y|R}(y',0)<P_{Y|R}(y',1)\ , \tag{15}$$

and there exists $z\in\mathcal{Z}$ such that

$$P_Z(z)>0\quad\text{and}\quad 0<P_{R|Z}(0,z)<1\ . \tag{16}$$

Proof: First we give an alternative characterization of the independence of the three channels, i.e., of $P_{XYZ|R}=P_{X|R}\cdot P_{Y|R}\cdot P_{Z|R}$. (We sometimes omit all the arguments of the probability distribution functions. In this case the statements hold for all possible choices of arguments. For example, $P_{X|Y}=P_X$ stands for $P_{X|Y}(x,y)=P_X(x)$ for all $x\in\mathcal{X}$ and $y\in\mathcal{Y}$.) From

$$P_{YZ|R}=\sum_{x\in\mathcal{X}}P_{XYZ|R}=\sum_{x\in\mathcal{X}}P_{X|R}\cdot P_{Y|R}\cdot P_{Z|R}=P_{Y|R}\cdot P_{Z|R}$$

and

$$P_R \cdot P_{YZ|R} \cdot P_{X|YZR} = P_{XYZR} = P_R \cdot P_{X|R} \cdot P_{Y|R} \cdot P_{Z|R}$$

we conclude that $P_{X|YZR} = P_{X|R}$ and, analogously, that $P_{Y|XZR} = P_{Y|R}$ and $P_{Z|XYR} = P_{Z|R}$.

(i) implies (ii). Let $I(X;Y|Z) > 0$. Assume $I(X;R) = 0$. Then $P_{X|YZR} = P_{X|R} = P_X$, and X is also independent of YZ (and hence of Z). Thus

$$I(X;Y|Z) = H(X|Z) - H(X|YZ) = H(X) - H(X) = 0 \ ,$$

which is a contradiction. We conclude that $I(X;R) > 0$ and by a symmetric argument that $I(Y;R) > 0$. Finally assume $H(R|Z) = 0$. Then

$$\begin{aligned} I(X;Y|Z) &= H(X|Z) + \underbrace{H(R|XZ)}_{0} - H(X|YZ) - \underbrace{H(R|XYZ)}_{0} \\ &= H(XR|Z) - H(XR|YZ) \\ &= \underbrace{H(R|Z)}_{0} + H(X|RZ) - \underbrace{H(R|YZ)}_{0} - H(X|RYZ) \\ &= H(X|R) - H(X|R) = 0 \ , \end{aligned}$$

which is a contradiction. Hence $H(R|Z) > 0$.

(ii) implies (iii). Let $I(X;R) > 0$, that is X and R are not statistically independent, which implies that there exists $\overline{x}$ such that $P_{X|R}(\overline{x},0) \neq P_{X|R}(\overline{x},1)$, i.e., such that one of the inequalities of (14) holds. Because

$$\sum_{x \in \mathcal{X}} P_{X|R}(x,0) = \sum_{x \in \mathcal{X}} P_{X|R}(x,1) = 1 \ ,$$

there must as well exist an element of $\mathcal{X}$ satisfying the other inequality of (14). Similarly we conclude the existence of appropriate y and y' from $I(Y;R) > 0$. Finally, $P_{R|Z}(0,z) \in \{0,1\}$ for all $z \in \mathcal{Z}$ with $P_Z(z) > 0$ would imply that $H(R|Z) = 0$. Hence (16) holds for some $z \in \mathcal{Z}$.

(iii) implies (i). Let x, x', y, y', and z be as in *(iii)*. It suffices to prove that $I(X;Y|Z =$

$z) > 0$ because $P_Z(z) > 0$. This is equivalent to the statement that X and Y are not statistically independent, given $Z = z$. We show that

$$P_{X|YZ}(x, y, z) > P_{X|YZ}(x, y', z) \ . \tag{17}$$

For both $\overline{y} = y$ and $\overline{y} = y'$ we have

$$P_{X|YZ}(x, \overline{y}, z) = P_{X|R=0}(x) \cdot P_{R|YZ}(0, \overline{y}, z) + P_{X|R=1}(x) \cdot P_{R|YZ}(1, \overline{y}, z) \ .$$

Because $P_{X|R=0}(x) > P_{X|R=1}(x)$, in order to prove (17), we have to show

$$P_{R|YZ}(0, y, z) > P_{R|YZ}(0, y', z) \ , \tag{18}$$

and because of $P_{R|YZ} = P_{Y|R} \cdot P_{RZ}/(P_{Y|Z} \cdot P_Z)$, inequality (18) is equivalent to

$$\frac{P_{Y|R}(y, 0)}{P_{Y|Z}(y, z)} > \frac{P_{Y|R}(y', 0)}{P_{Y|Z}(y', z)} \ ,$$

which follows from

$$\begin{aligned}
& P_{Y|R=0}(y) \cdot [P_{Y|R=0}(y') \cdot P_{R|Z=z}(0) + P_{Y|R=1}(y') \cdot P_{R|Z=z}(1)] \\
> \quad & P_{Y|R=0}(y) \cdot P_{Y|R=0}(y') \\
> \quad & [P_{Y|R=0}(y) \cdot P_{R|Z=z}(0) + P_{Y|R=1}(y) \cdot P_{R|Z=z}(1)] \cdot P_{Y|R=0}(y') \ .
\end{aligned} \tag{19}$$

Both inequalities in (19) follow from the fact that $0 < P_{R|Z=z}(0) < 1$, and because of (15). ■

We are now ready to prove Theorem 7.

Proof of Theorem 7: Clearly, *(B)* implies *(C)* by Theorem 2, and *(C)* implies *(A)* by the definition of the intrinsic information. We show that *(A)* implies *(B)*. Given that $I(X;Y|Z) > 0$, we construct, from R, X, Y, and Z, random variables $\tilde{R}$, $\tilde{X}$, $\tilde{Y}$, and U with the following properties (see also Figure 5):

1. $\tilde{X}$ and $\tilde{Y}$ are generated from X and Y, respectively, with positive probability.
2. $\tilde{R}$ is binary and symmetric, and $\tilde{X}$ and $\tilde{Y}$ can be interpreted as being generated by sending $\tilde{R}$ over two independent binary symmetric channels with identical error probability $\alpha < 1/2$.

3. $\tilde{Z} := [Z, U]$ contains exactly the same information about $\tilde{R}$ as a random variable generated by sending $\tilde{R}$ over a binary erasure channel (which is independent of the channels from (3)) with positive erasure probabilities $\delta > 0$ and $\delta' > 0$.

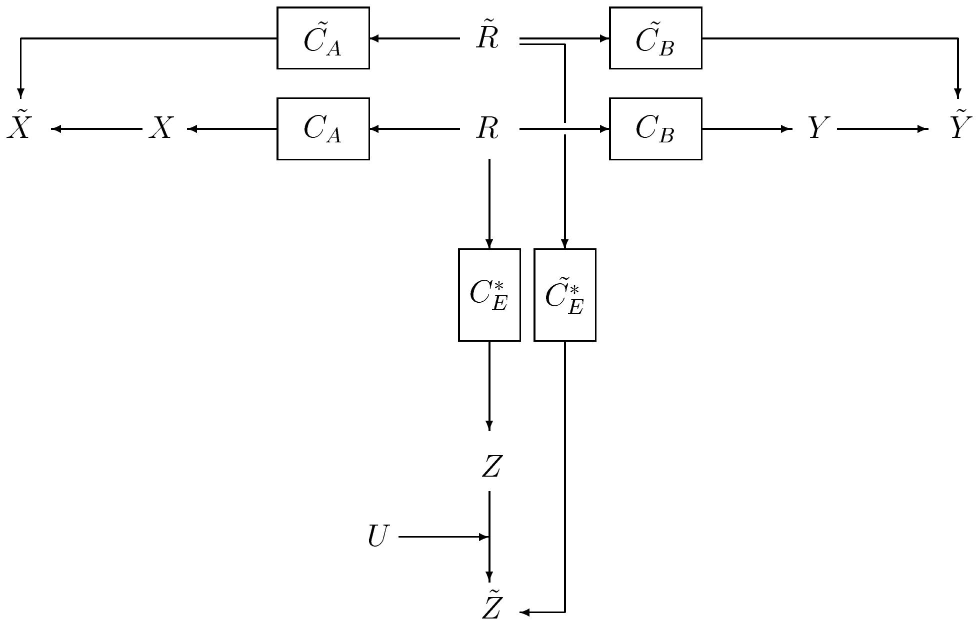

Fig. 5. The random variables in the proof of Theorem 7

For such random variables $\tilde{X}$, $\tilde{Y}$, and U, we have by Lemma 11 that $S(\tilde{X};\tilde{Y}||[Z,U]) > 0$, and with Theorem 6 we conclude that $S(X;Y||Z) > 0$. Hence it remains to show that suitable random variables $\tilde{R}$, $\tilde{X}$, $\tilde{Y}$, and U exist.

According to Lemma 12, there exist $x, x' \in \mathcal{X}$ and $y, y' \in \mathcal{Y}$ such that (14) and (15) hold. Let $\tilde{X}$ and $\tilde{Y}$ be obtained from X and Y as follows. First, the ranges of X and Y are restricted to $\{x, x'\}$ and $\{y, y'\}$, respectively, and secondly, the resulting random variables X' and Y' (which correspond to a new random experiment) are made symmetric. This is done by sending X' over the following channel to obtain $\tilde{X}$ (we assume $P_X(x) \geq P_X(x')$ without loss of generality):

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

$$P_{\tilde{X}|X'}(1,x') = 1 \ , \quad P_{\tilde{X}|X'}(0,x') = 0 \ .$$

In an analogous way, $\tilde{Y}$ and $\tilde{R}$ are obtained from Y' and R, respectively.

According to Lemma 12 there exists $z \in \mathcal{Z}$ such that $P_Z(z) > 0$ and $0 < P_{R|Z}(0,z) < 1$. Let the random variable U be defined as follows. If $Z \neq z$, let $U = \tilde{R}$, and if $Z = z$, let $U = \Delta$. Intuitively, the information U can be thought as being provided by an oracle that tells Eve the correct $\tilde{R}$ if $Z \neq z$. Such an oracle can only decrease Eve's average error probability and, according to Lemma 5, the secret key rate.

It remains to show that $\tilde{R}$, $\tilde{X}$, $\tilde{Y}$, and U have all the stated properties. The properties 1 and 3 are satisfied by definition of the random variables. It remains to prove 2. First, it is clear that $\tilde{R}$, $\tilde{X}$, and $\tilde{Y}$ are binary and symmetric. For the rest, we consider the case $P_R(0) \geq 1/2$ and $P_{X'}(x) \geq 1/2$. The other cases are analogous. We have to show $P_{\tilde{X}\tilde{R}}(0,0) > P_{\tilde{X}}(0) \cdot P_{\tilde{R}}(0) = 1/4$, which is sufficient because $\tilde{R}$ and $\tilde{X}$ are symmetric binary random variables.

$$
\begin{aligned}
P_{\tilde{X}\tilde{R}}(0,0) &= P_{\tilde{X}\tilde{R}X'R}(0,0,x,0) \\
&= P_R(0) \cdot P_{\tilde{R}|R}(0,0) \cdot P_{X'|R}(x,0) \cdot P_{\tilde{X}|X'}(0,x) \\
&= \frac{1}{4} \cdot \frac{P_{X'|R}(x,0)}{P_{X'}(x)} > \frac{1}{4}
\end{aligned}
$$

because $P_{X'|R}(x,0) > P_{X'|R}(x,1)$ and $0 < P_R(0) < 1$. An analogous result can be proved for $\tilde{Y}$. As in the proof of Lemma 8, the error probabilities of the two channels can be made identical, and we have proved property 2. The theorem now follows from Lemma 11, Theorem 6, and Lemma 10. Note that in this application of Lemma 10 the event that Bob accepts means that Alice and Bob both accept a sufficiently large number N of consecutive realizations of X and Y (if Alice does not accept, she sends $M =$"reject" over the public channel), and that Bob accepts the received message sent by Alice. (Of course this would be a very wasteful and inefficient way of generating a secret key in practice. For example it is not necessary that the N realizations of X and Y accepted by Alice and Bob are consecutive.) This concludes the proof. ∎

Remark. The condition that R is a *binary* random variable is crucial in Theorem 7. To see this, consider the following example: Let R be uniformly distributed over the set $\mathcal{R} := \{r_{00}, r_{01}, r_{10}, r_{11}\}$, and let X, Y, and Z be binary random variables, generated from R by the following independent channels (let δ be the Kronecker symbol, i.e., $\delta_{ij} = 1$ if

$i = j$, and otherwise $\delta_{ij} = 0$):

$$P_{X|R}(x, r_{ij}) = \delta_{xi} \ , \quad P_{Y|R}(y, r_{ij}) = \delta_{yj} \ , \quad P_{Z|R}(z, r_{ij}) = \delta_{z,i \oplus j} \ .$$

Note that for all $r \in \mathcal{R}$, $Z = X \oplus Y$, that is $I(X;Y|Z) = 1$. On the other hand $I(X;Y) = 0$, and hence $S(X;Y||Z) = 0$.

In fact, any distribution P_{XYZ} can be seen as generated by sending a random variable R over three independent channels for some R with $|\mathcal{R}| \leq |\mathcal{X}| \cdot |\mathcal{Y}| \cdot |\mathcal{Z}|$. Such a random variable R can be defined as follows. Let $\mathcal{R} := \{r_{xyz} \mid (x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}\}$ and $P_{X|R}(\overline{x}, r_{xyz}) = \delta_{\overline{x}x}$, $P_{Y|R}(\overline{y}, r_{xyz}) = \delta_{\overline{y}y}$, and $P_{Z|R}(\overline{z}, r_{xyz}) = \delta_{\overline{z}z}$.

V. Towards the General Case: Protocol A Is not Optimal

In this section we assume that X and Y are completely general random variables, and that 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 it is less difficult to analyze than the completely general situation. Additionally, more general situations can be reduced, by the methods of Theorem 6, to such a scenario (with respect to the question whether secret-key agreement is possible).

We approach the general situation by studying two extremal cases. For Scenario 2 below, the statement of Conjecture 1 is shown to be true, whereas for Scenario 3, this problem remains open. Also for Scenario 3, we prove that Protocol A is not optimal. A new protocol is shown to be strictly stronger with respect to the possibility or impossibility of secret-key agreement.

Scenario 2. The random variables X and Y are binary and distributed according to

$$P_{XY}(0,0) = P_{XY}(1,1) = \frac{1-\alpha}{2} \ , \quad P_{XY}(0,1) = P_{XY}(1,0) = \frac{\alpha}{2} \tag{20}$$

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 3. The random variables X and Y are distributed according to (20), and $Z = [Z_X, Z_Y]$, where Z_X and Z_Y are generated by sending X and Y over two independent erasure channels with positive erasure probabilities r_X and r_Y, respectively.

The two scenarios are illustrated in Figure 6.

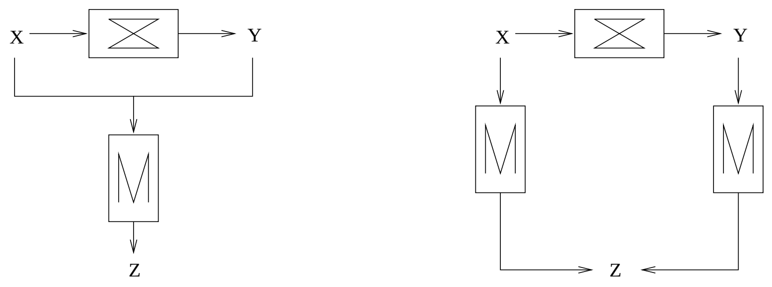

Fig. 6. Scenarios 2 and 3. The boxes stand for binary symmetric and for erasure channels, respectively. (Note that the erasure channel in Scenario 2 is not binary.)

Prior to the analysis of Scenarios 2 and 3 we show that under a condition which appears to be satisfied most likely if X, Y, and Z are 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 Scenarios 2 or 3. Theorem 6 then implies that $S(X;Y||Z) > 0$ if $S(\overline{X};\overline{Y}||[Z,U]) > 0$. The proof of the following lemma is related to one of the arguments in the proof of Theorem 7.

Lemma 13: 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 (20) 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

$$P_{X|Y}(x,y) > P_X(x) > P_{X|Y}(x,y') \tag{21}$$

and

$$P_{Y|X}(y,x) > P_Y(y) > P_{Y|X}(y,x') \ . \tag{22}$$

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 $P_{\hat{X}}(x) \geq 1/2$):

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

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

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 (20) 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 (21) and of (22). ■

If, for $\overline{X}$ and $\overline{Y}$ obtained as in Lemma 13, 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, i.e., the pair $[Z,U]$ can be interpreted as being generated by sending $[\overline{X}, \overline{Y}]$ over an erasure channel with positive erasure probability.

We conclude that very general situations can be reduced to Scenario 2, in which Z is obtained by sending $[X,Y]$ over an erasure channel. In an analogous way, general distributions can be reduced to Scenario 3. However, it appears to be difficult to decide in general which reduction leads to the strongest results.

A. Analysis of Scenario 2

Scenario 2 has been defined above as the symmetric situation where X and Y are distributed according to (20) with $\alpha < 1/2$, and where Z is obtained by sending $[X,Y]$ over an erasure channel with erasure probability $1-r$. We derive a condition for when Protocol A (with parameter N) allows secret-key agreement. Bob's conditional error probability when guessing the bit sent by Alice, given that he accepts, is

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

where $p_{a,N} = \alpha^N + (1-\alpha)^N$ 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 Δ. In the latter case her error probability is $1/2$, independently of her strategy. Hence Eve's error probability, given that Bob accepts, is

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

Using Lemma 10, we conclude that Protocol A works and allows the generation of a secret key if

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

Theorem 14 shows that Protocol A is optimal in Scenario 2 in the sense that if *some* protocol allows secret-key agreement in principle, then Protocol A also does. Another consequence of Theorem 14 is that Conjecture 1 is true in Scenario 2.

Theorem 14: In Scenario 2, the following four conditions are equivalent.

(i) $I(X;Y{\downarrow}Z) > 0$,

(ii) $1 - r > \alpha/(1-\alpha)$,

(iii) Protocol A allows secret-key agreement,

(iv) $S(X;Y||Z) > 0$.

Proof: It remains to show that *(i)* implies *(ii)*, i.e., that $I(X;Y \downarrow Z) = 0$ holds whenever $1 - r \leq \alpha/(1-\alpha)$. Let the random variable $\overline{Z}$ 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

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

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., $I(X;Y|\overline{Z} = \overline{\Delta}) = 0$. ■

B. Erasure Probability, Deviation from Independence, and Intrinsic Information

In this section we show that a property of the intrinsic information which is similar to the implication from *(i)* to *(ii)* in Theorem 14 can be proved also in the case of general random variables X and Y. Namely, we show that $I(X;Y{\downarrow}Z)=0$ if Eve knows X and Y precisely with some positive probability, and if the joint 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 joint 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 joint 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} we have

$$0 \le d_{ind}(P_{XY}) \le 1 , \tag{23}$$

with equality on the left hand side of (23) if and only if the random variables X and Y are independent, and with equality on the right hand side of (23) if and only if there exist x and y such that $P_{XY}(x,y) = 0$ although $P_X(x) \cdot P_Y(y) \neq 0$. Furthermore,

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

This can be seen by taking the uniform distribution for Q_{XY}. When X and Y are distributed according to (20), then

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

Theorem 15 implies that secret-key agreement is impossible under the surprisingly simple and intuitive condition that the probability that Eve reliably knows X and Y equals or exceeds $d_{ind}(P_{XY})$.

Theorem 15: Let X and Y be arbitrary random variables with joint 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

$$(1-r)\cdot P_{XY}(x,y) \leq \lambda \cdot Q_{XY}(x,y) \leq P_{XY}(x,y) \tag{24}$$

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 (24), $\overline{Z}$ 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

$$P_{XY\overline{Z}}(x,y,\overline{\Delta}) = \lambda \cdot Q_{XY}(x,y)\ . \tag{25}$$

This can be done for all pairs (x,y), and (25) implies $P_{XY|\overline{Z}=\overline{\Delta}} = Q_{XY}$, i.e., that X and Y are independent when given $\overline{Z}=\overline{\Delta}$. Hence

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

and the theorem is proved. ■

C. Analysis of Scenario 3.

In this section we analyze Scenario 3. Let α be the probability that $X \neq Y$, and let r_X and r_Y be the probabilities that Eve does *not* receive the erasure symbol from her (independent) channels. We assume here that $r_Y \geq r_X$. For fixed α and r_Y, we prove three different upper bounds on r_X with the property that secret-key agreement is possible if r_X is smaller than at least one of these bounds. Moreover, a new protocol is presented that is applicable for a larger class of distributions P_{XYZ} than Protocol A, hence proving that Protocol A is not optimal for Scenario 3.

The first upper bound on r_X comes from a rather straight-forward argument. According to Theorem 1 the secret key rate is positive if $I(X;Y) > I(X;Z)$. This condition is equivalent to

$$H(X|Y) < H(X|Z)\ , \tag{26}$$

where

$$
\begin{aligned}
H(X|Y) &= h(\alpha) \\
H(X|Z) &= (1-r_X)(1-r_Y) + (1-r_X)r_Y h(\alpha) \ .
\end{aligned}
$$

Lemma 16: In Scenario 3, $S(X;Y||Z)$ is strictly positive if

$$
r_X < 1 - \frac{h(\alpha)}{1 - r_Y + r_Y h(\alpha)} \ . \tag{27}
$$

If Lemma 16 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 α^N. On the other hand, the probability p_{01} that Bob accepts and receives the correct bit, and that Eve guesses 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, independently 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 10 we can conclude that Protocol A works if and only if

$$
\alpha < (1-\alpha)(1-r_X)(1-r_Y) \ , \tag{28}
$$

and the following lemma is proved.

Lemma 17: In Scenario 3, Protocol A allows secret-key agreement, and $S(X;Y||Z)$ is positive, if

$$
r_X < 1 - \frac{\alpha}{(1-\alpha)(1-r_Y)} \ . \tag{29}
$$

Note that this bound is strictly positive only if $1 - r_Y > \alpha/(1-\alpha)$. This is the same condition as in Theorem 14 of the previous section.

We remark that each of the expressions in (27) and (29) can be greater than the other. If r_Y is constant and $\alpha \to 0$, the expression of (29) is greater, whereas if $r_Y = \alpha/(1-\alpha)$, the expression of (29) equals 0, and the expression of (27) is greater than 0 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 effect that Alice's and Bob's bits become more reliable 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 in Scenario 3 the following protocol is superior to Protocol A.

Protocol B. Bob randomly chooses a bit C and 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 $[1, 1, \ldots, 1]$.

The analysis of the protocol shows that it is advantageous for Alice and Bob when Bob, and not Alice, is the sender of the bit in Protocol B if $r_Y \geq r_X$. Note that Protocol B corresponds to Protocol A for the choice $t = 1$. Protocol B is, as Protocol A, efficient in terms of computation but wasteful with respect to the achievable rate of generated secret key. An efficiency improvement similar to the parity-check version of Protocol A [12] exists also for Protocol B.

The analysis of this protocol in Scenario 3 is quite technical, and is given in Appendix B, where Theorem 18 is proved. It gives an upper bound on r_X when given α and r_Y. We only mention here the surprising fact that t must typically be chosen only slightly greater than $1/2$ (whereas it is obvious that the choice $t = 1/2$ is completely useless).

Theorem 18: Protocol B allows secret-key agreement in Scenario 3, and $S(X;Y||Z)$ is positive, if

$$r_X < \frac{2\left(1-\frac{\alpha}{1-\alpha}\right)^2(1-r_Y)}{5-4r_Y} \tag{30}$$

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

$$r_X < (1-r_Y)\left(1-\frac{\alpha}{1-\alpha}-\frac{1-r_Y}{2}-\frac{1}{8}\right) \tag{31}$$

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

Theorem 18 shows that in Scenario 3, Protocol B is strictly better than Protocol A, which is therefore not optimal. It is easy to see that the upper bounds of (30) and (31) are greater than the bounds given by (27) and (29) in many cases. We consider two examples.

If r_Y is constant and $\alpha \to 1/2$, then the bound given in (27) tends to 0 much faster than (30) (which applies in this situation). The bound of (29) is even negative. On the other hand, if $\alpha = 3/7$, and $r_Y \to 1$, then (27) is smaller than (31) (which applies here). The bound (29) is negative again.

Note that the bounds (30) and (31) are not tight. In particular, the bounds from an optimal analysis of Protocol B must be greater than the bound from Protocol A because Protocol A is a special case of Protocol B. However, an exact analysis of Protocol B appears to be difficult.

Finally, we give a pessimistic bound on r_X for Scenario 3. As in the previous section we derive a condition here for the fact that $I(X;Y{\downarrow}Z) = 0$.

Theorem 19: In Scenario 3, $I(X;Y{\downarrow}Z) = 0$ if

$$r_X \geq \frac{(1-r_Y)(1-2\alpha)}{r_Y\alpha + 2(1-r_Y)((1-\alpha)\sqrt{1-2\alpha}-(1-2\alpha)) + (1-r_Y)(1-2\alpha)}\ . \tag{32}$$

The proof of Theorem 19 is given in Appendix C. Of course the bound on r_X given in (32) is greater than the bounds (27), (29), and (30) (or (31), respectively) for all possible choices of α and r_Y.

VI. Concluding Remarks

We have investigated the problem of generating a provably secure key by public discussion from correlated information. Steps have been taken towards characterizing under what conditions on this information such secret-key agreement is possible in principle. In particular, we have introduced a new information measure which turned out to provide such a characterization in many situations. However, it is not clear whether this is true in general. For Scenario 3 discussed above, the resulting (sufficient but not necessary) conditions for $I(X;Y\downarrow Z)=0$ and for the presented protocols for secret-key agreement to be successful are not exactly complementary. (This is true although both the optimistic and pessimistic bounds of Theorems 18 and 19 can be slightly improved by a better but more complicated analysis.) We suggest as an open problem to derive *necessary and sufficient* conditions for either $I(X;Y\downarrow Z)=0$ and $S(X;Y||Z)>0$, and to decide whether Conjecture 1 also holds in Scenario 3, and in general.

Acknowledgments

The authors thank Andrin Bachmann, Christian Cachin, and Marten van Dijk for interesting discussions on the subject of this paper, and two anonymous referees for their valuable comments on the presentation of the results.

References

[1] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography - Part I: secret sharing," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1121–1132, 1993.

[2] 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.

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

[4] C. Crépeau, "Efficient cryptographic protocols based on noisy channels," in *Advances in Cryptology - EUROCRYPT '97*, Lecture Notes in Computer Science, Springer-Verlag, vol. 1233, pp. 306–317, 1997.

[5] C. Crépeau and J. Kilian, "Achieving oblivious transfer using weakened security assumptions," in *29th Symposium on Foundations of Computer Science*, pp. 42–52, IEEE, 1988.

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

[7] M. van Dijk, "Secret key sharing and secret key generation," Ph. D. Thesis, Technische Universiteit Eindhoven, 1997.

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

[9] 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.

[10] U. M. Maurer, "Information-theoretically secure secret-key agreement by NOT authenticated public discussion," in *Advances in Cryptology - EUROCRYPT '97*, Lecture Notes in Computer Science, Springer-Verlag, vol. 1233, pp. 209–225, 1997.

[11] U. M. Maurer, "The strong secret key rate of discrete random triples," in *Communication and Cryptography – Two Sides of One Tapestry*, Kluwer Academic Publishers, pp. 271–285, 1994.

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

[13] 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.

[14] U. M. Maurer and S. Wolf, "Privacy amplification secure against active adversaries," in *Advances in Cryptology - CRYPTO '97*, Lecture Notes in Computer Science, Springer-Verlag, vol. 1294, pp. 307–321, 1996.

[15] U. M. Maurer and S. Wolf, "The intrinsic conditional mutual information and perfect secrecy," in *Proc. of the 1997 IEEE Symp. on Information Theory*, Ulm, Germany, 1997.

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

[17] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, Oct. 1949.

[18] S. Wolf, "Strong security against active attacks in information-theoretic secret-key agreement," to appear in *Advances in Cryptology - ASIACRYPT '98*, Lecture Notes in Computer Science, Springer-Verlag, 1998.

[19] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[20] R. W. Yeoung, "A new outlook on Shannon's information measures," *IEEE Transactions on Information Theory*, vol. 37, no. 3, pp. 466–474, 1991.

Appendix A: Continuous Random Variables from Independent Binary-Input Channels

Here we show that the result of Theorem 7 also holds when the random variables that are generated from R are not discrete. For example, this is the case if 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 [3]):

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

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 IV we assume that X, Y, and Z are generated by sending a binary random variable R over independent channels, i.e.,

$$f_{XYZ|R} = f_{X|R} \cdot f_{Y|R} \cdot f_{Z|R} \ , \tag{33}$$

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

Theorem 20: 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 $f_{XY|Z}$. Assume that (33) 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:

Fact 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

$$P_{\overline{X}|R}(0,0) > P_{\overline{X}|R}(1,0) \tag{34}$$

and

$$P_{\overline{X}|R}(0,1) < P_{\overline{X}|R}(1,1) \tag{35}$$

(as well as the corresponding inequalities when replacing $\overline{X}$ by $\overline{Y}$) hold.
Fact 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 Theorem 7 show that Facts 1 and 2 together imply $S(X;Y||Z) > 0$.

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

$$\mathrm{Prob}_X[f_{X|R=0}(x) \neq f_{X|R=1}(x)] > 0 \ . \tag{36}$$

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 (36) 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

$$P_{X|R=0}(A_0) > P_{X|R=1}(A_0) \tag{37}$$

and

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

(where $P_{X|R=0}(A_0)$ stands for $\int_{A_0} f_{X|R=0}\,dx$). Inequality (37) holds because if $P_{X|R=0}(A_0) = P_{X|R=1}(A_0)$, 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 (36). 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 (34) and (35) hold, as well as the corresponding inequalities for $\overline{Y}$.

Proof of Fact 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 D with $\mu(D) > 0$ (where μ denotes the Lebesgue-measure of $\mathbf{R}$) and

$$I(X;Y|Z=z) > 0 \quad \text{for all} \quad z \in D\ . \tag{38}$$

Because of (38) we have both $f_{R|Z=z}(0) > 0$ and $f_{R|Z=z}(1) > 0$ for all $z \in D$. (If for example $f_{R|Z=z}(0) = 0$, 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 D_n be the (measurable) set of all z in D such that $f_{R|Z=z}(0) \geq P_R(0)/n$ and $f_{R|Z=z}(1) \geq P_R(1)/n$. Then $D = \cup D_n$, and $\mu(D) > 0$ implies

$$0 < \mu(D) = \mu(\cup D_n) \leq \sum_n \mu(D_n) .$$

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

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

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

if $z \in D_{n_0}$ 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 $\mu(D_{n_0})/n_0 > 0$. ■

Appendix B: Analysis of Protocol B in Scenario 3

Let the protocol parameter t be fixed, and let

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

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

$$\beta_N = \frac{\alpha^{tN}(1-\alpha)^{(1-t)N}}{(1-\alpha)^{tN}\alpha^{(1-t)N} + \alpha^{tN}(1-\alpha)^{(1-t)N}} \leq \left(\frac{\alpha}{1-\alpha}\right)^{N/(2K)} . \tag{39}$$

Eve's conditional error probability γ_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 on γ_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.

$$\gamma_N \geq \frac{1}{2} \cdot \max_{0\leq s\leq (1-t)} \left\{ \binom{tN}{sN} (r_Y)^{sN} (1-r_Y)^{(t-s)N} \binom{(1-t)N}{sN} (r_Y)^{sN} (1-r_Y)^{(1-t-s)N} \right\} \cdot (1-r_X)^N . \tag{40}$$

The next lemma gives a simpler lower bound that can be derived from the bound (40) by determining its asymptotic behavior.

Lemma 21: The lower bound (40) implies that

$$\gamma_N^{2K/N} \geq 1 - \frac{1}{4K} - \frac{1}{16(1-r_Y)K} - 2Kr_X \tag{41}$$

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 [8]) 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 (40) can be replaced by the corresponding expressions, and a straightforward computation leads to the following asymptotic behavior of the lower bound on γ_N.

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

for sufficiently large N. ■

The bound (41) in the above lemma holds for all K that correspond to a protocol parameter t which satisfies $r_Y/2 \leq 1-t$. This condition is equivalent to

$$\frac{1}{2K} \leq 1 - r_Y \ . \tag{42}$$

The idea of the proof of Theorem 30 is to find the best choice for K (i.e., the best choice of t in Protocol B) with respect to the fixed parameters α and r_Y, and such that (42) 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 18: According to (39) and (41), Protocol B (with parameter t) works for secret-key agreement if

$$\gamma_N^{2K/N} \geq 1 - \frac{1}{4K} - \frac{1}{16(1-r_Y)K} - 2Kr_X > \frac{\alpha}{1-\alpha} \geq \beta_N^{2K/N} \ , \tag{43}$$

and if the condition (42) also holds. The reason is that (43) implies that Eve's error probability about the bit sent by Bob is asymptotically greater than Alice's error probability for $N \to \infty$. Lemma 10 states that this is sufficient for the possibility of secret-key agreement by public discussion. Let $\delta := 1 - \alpha/(1-\alpha)$. Then (43) is satisfied if

$$r_X < \frac{\delta}{2K} - \frac{1}{8K^2}\left(1 + \frac{1}{4(1-r_Y)}\right) \ . \tag{44}$$

This bound depends on K, and from (44) 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 (42). It is easy to see that the expression on the right of (44) 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 δ 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). Choosing $K = K_0$ is compatible with (42) if $\delta \geq 5/4 - r_Y$. Then the condition (44) is

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

If $\delta > 5/4 - r_Y$, the condition (42) is not fulfilled for $K = K_0$. For $K = K_0' := 1/(2-2r_Y)$ (the smallest choice for K that satisfies (42)) the right hand side of (44) equals

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

This proves Theorem 18. ■

Note that the main objective of the above analysis of Protocol B is to show that it leads to a strict improvement of Protocol A, rather than to characterize the performance of the protocol completely. In particular, the bounds of Theorem 18 are not tight by two reasons. First, it is not necessary to choose t such that $r_Y/2$ is a possible choice for s, as done in the proof of Lemma 21. Secondly, 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 appears to be difficult to determine Eve's optimal strategy of guessing the bit, and hence to compute the *exact* error probability of her guess. Note that with an optimal analysis, Protocol B would clearly turn out to be at least as good as Protocol A in *any* situation, because Protocol A is a special case of Protocol B and corresponds to the choice $t = 1$. It is finally conceivable that the above results can be improved when a block protocol is used in which both Alice and Bob (and not only Bob) have a block that is not composed by N times the same bit. However, such a protocol appears to be much more difficult to analyze.

Appendix C: Proof of Theorem 19

Proof: We show that if (32) 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 $\overline{Z} = \Delta_3$ also 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 split.

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

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

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

and $\overline{Z}=Z$ otherwise. The parameter $0\le\mu\le 1$ 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, which 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 split 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

$$
\frac{r_X r_Y\alpha}{1-2\alpha}+2r_X(1-r_Y)\cdot\left(\frac{1-\alpha}{\sqrt{1-2\alpha}}-1\right)\ge P_Z([\Delta,\Delta])=(1-r_X)(1-r_Y)\ , \tag{45}
$$

and this is equivalent to (32). ■

Remark. Note that the condition given in the lemma is sufficient, but not necessary for $I(X;Y\downarrow Z)=0$. If $r_X\neq r_Y$, a better bound can be achieved when $Z=[0,\Delta]$ and

$Z = [\Delta, 0]$ (as well as $Z = [1, \Delta]$ and $Z = [\Delta, 1]$) are not transformed symmetrically to $\overline{Z} = \Delta_2$ ($\overline{Z} = \Delta_3$), but each with the maximal possible probability, i.e., $r_X(1 - r_Y)/2$ and $(1 - r_X)r_Y/2$, respectively. The condition (19) for $I(X;Y{\downarrow}Z) = 0$ can then be replaced by the better, but more complicated condition

$$\frac{r_X r_Y \alpha}{1 - 2\alpha} - r_X(1 - r_Y) - r_Y(1 - r_X) + \sqrt{T} \geq (1 - r_X)(1 - r_Y) \ ,$$

where

$$T := r_X^2(1 - r_Y)^2 + r_Y^2(1 - r_X)^2 + \left(2 + \frac{4\alpha^2}{1 - 2\alpha}\right) r_X(1 - r_X)r_Y(1 - r_Y) \ .$$