Information-Theoretic Secret-Key Agreement: The Asymptotically Tight Relation Between the Secret-Key Rate and the Channel Quality Ratio

Daniel Jost[1][0000−0002−6562−9665], Ueli Maurer[1], and João L. Ribeiro[2⋆][0000−0002−9870−0501]

[1] Department of Computer Science, ETH Zurich, 8092 Zurich, Switzerland.
{dajost, maurer}@inf.ethz.ch

[2] Department of Computing, Imperial College London, SW7 2AZ London, UK.
j.lourenco-ribeiro17@imperial.ac.uk

Abstract. Information-theoretic secret-key agreement between two parties Alice and Bob is a well-studied problem that is provably impossible in a plain model with public (authenticated) communication, but is known to be possible in a model where the parties also have access to some correlated randomness. One particular type of such correlated randomness is the so-called satellite setting, where uniform random bits (e.g., sent by a satellite) are received by the parties and the adversary Eve over inherently noisy channels. The antenna size determines the error probability, and the antenna is the adversary's limiting resource much as computing power is the limiting resource in traditional complexity-based security. The natural assumption about the adversary is that her antenna is at most Q times larger than both Alice's and Bob's antenna, where, to be realistic, Q can be very large.

The goal of this paper is to characterize the secret-key rate per transmitted bit in terms of Q. Traditional results in this so-called satellite setting are phrased in terms of the error probabilities ϵ_A, ϵ_B, and ϵ_E, of the binary symmetric channels through which the parties receive the bits and, quite surprisingly, the secret-key rate has been shown to be strictly positive unless Eve's channel is perfect ($\epsilon_E = 0$) or either Alice's or Bob's channel output is independent of the transmitted bit (i.e., $\epsilon_A = 0.5$ or $\epsilon_B = 0.5$). However, the best proven lower bound, if interpreted in terms of the channel quality ratio Q, is only exponentially small in Q. The main result of this paper is that the secret-key rate decreases asymptotically only like $1/Q^2$ if the per-bit signal energy, affecting the quality of all channels, is treated as a system parameter that can be optimized. Moreover, this bound is tight if Alice and Bob have the same antenna sizes.

Motivated by considering a fixed sending signal power, in which case the per-bit energy is inversely proportional to the bit-rate, we also propose a definition of the secret-key rate per second (rather than per transmitted bit) and prove that it decreases asymptotically only like $1/Q$.

⋆ Part of the work was performed while at ETH Zurich.

1 Introduction

1.1 Motivation for Information-theoretic Security

In cryptography, one generally considers two types of security of cryptographic schemes. *Unconditional* or *information-theoretic* security means that not even an adversary with unbounded computing power can cause a violation of the security property, whereas *computational* security means that the violation of the security property is impossible for an adversary with (suitably) bounded computing power, but is usually possible for a computationally unbounded adversary. Information-theoretic security was first defined and considered in Shannon's ground-breaking paper [22].

While for the most part cryptographic research is focused on computational security, actually the state of the art in complexity theory is that no cryptographic scheme has been proven to be computationally secure for a general and realistic model of computation. Instead, the term "provable security" is often used for schemes for which a reduction from a commonly agreed conjectured hard problem (such as factoring large integers) is known: Any adversary breaking the cryptographic scheme could be transformed (by the reduction), with reasonable efficiency loss, into an algorithm solving the hard problem with noticeable probability. Therefore, under the assumption that the problem is indeed hard, the scheme is secure.

In summary, there are two main advantages of information-theoretic security:

- Information-theoretic security is stronger because, compared to computational security, the security holds against a larger class of adversaries.
- The security proof does not require an unproven computational assumption.

1.2 Circumventing Impossibility Results

Unfortunately, information-theoretic security is in many settings unachievable, often provably so, at least for practical settings. For instance, Shannon's famous impossibility result [22] states that perfectly secure encryption is impossible unless the secret key has at least as much entropy as the message. This result is often quoted as showing that information-theoretic security is not practical since exchanging a fresh truly random key for every message is generally completely impractical.

The significance of such an impossibility result depends on the generality of the conditions underlying the impossibility proof. For example, Shannon's impossibility result was stated (and proven) only under the restriction that the communication between sender and receiver is one-way. That this result also holds in the more realistic setting with interactive communication between sender and receiver has been proven by Maurer only in 1993 [11]. It is therefore possible that a careful re-examination of impossibility results allows to circumvent them by a slight change of the model, where such a change should be as realistic as possible and should not destroy the practicality of schemes proven secure in the model.

A prominent such modification is quantum key distribution (QKD), where one assumes that the honest parties can exchange quantum information and thereby achieves perfect security. Given that being able to exchange quantum information is a very strong assumption for many practical scenarios, however, classical settings are still of great interest. One such model, proposed by Maurer [15] and investigated by many researchers in different contexts, is the so-called bounded-storage model. Here one assumes that the adversary's memory resources are bounded, but no assumption about the adversary's computing power is needed. Unfortunately, it seems very hard to argue that schemes proven secure in this model are practical for a reasonable bound on the adversary's memory capacity.

Other notable earlier attempts include the works of Wyner [25] and Csiszár and Körner [4], where all parties are connected by noisy channels (and only one-way communication between the two honest parties is allowed), and the work of Ozarow and Wyner [19], where the adversary is allowed to observe a bounded subset of the message's encoding. In these models, perfectly secure encryption is possible only when the adversary is at a disadvantage compared to the honest parties, which is rarely the case in practice.

A more promising approach in the context of secret-key agreement is the so-called *secret-key agreement by public discussion* model proposed by Maurer [16,11]. In this model, two parties Alice and Bob wish to agree on a secret key by communicating over a public authenticated channel perfectly accessible to the adversary Eve. In this setting, without further assumptions, key agreement is provably impossible. However, by a slight modification of the model, namely by considering a setting where Alice, Bob, and Eve have access to correlated random variables X, Y, and Z, respectively, with joint probability distribution P_{XYZ}, secret-key agreement becomes possible, even if X and Y are almost not correlated and even if Z is strongly correlated with both X and Y.

Often one considers a setting where the experiment generating X, Y, and Z is repeated many times (independently), and one then considers the *secret-key rate*, the maximal rate (per realization of the random experiment) at which Alice and Bob can generate secret-key bits. Surprisingly, in this model, secret-key agreement (and thus perfectly secure encryption) is also possible in many cases where Eve starts with an advantage over Alice and Bob.

1.3 The Satellite Setting

A setting of particular interest is the so-called satellite setting: A satellite (or for instance a deep-space radio source) broadcasts a sequence of uniformly random bits that Alice, Bob, and Eve receive via antennas of different sizes.

In order to achieve a meaningfully large secret-key rate in this setting, one has to assume that the adversary's resources are bounded. While in computationally secure cryptography the bounded resource is the computing power, in the satellite model the natural bounded resource of the adversary is her antenna quality, that closely corresponds to the antenna size. Given that for most practical settings the honest parties' antenna sizes are more or less fixed, we specify

in the following, for simplicity, this bound on Eve's antenna size as the maximal ratio Q between Eve's antenna size and the size of the smaller one of either Alice's or Bob's antennas. Analogously to the computational setting where the ratio between the adversary's and the honest parties' computing power must be assumed to be quite large, this antenna size ratio Q can be very large as well, in realistic settings. If the honest parties use for instance mobile phones, then it is very well imaginable that Q is in the order of magnitude of a million.

The satellite setting is modeled as a sequence of uniform random bits being generated and Alice, Bob, and Eve receiving them over independent binary symmetric channels with error probabilities ϵ_A, ϵ_B, and ϵ_E, respectively. Traditionally, the secret-key rate in the satellite model has then been specified in terms of the error probabilities ϵ_A, ϵ_B, and ϵ_E, respectively, capturing the fact that the antenna sizes clearly affect those error probabilities. However, it is natural to consider the signal strength of the satellite, i.e., the amount of energy it uses to broadcast each bit, as a design parameter we can control, implying that the error probabilities are no longer a priori fixed. Moreover, this highlights an interesting trade-off, as increasing the energy per bit means that the error probabilities of Alice, Bob, and Eve all decrease simultaneously, which is at the same time advantageous (Alice and Bob getting more information) and disadvantageous (Eve getting more information). As a consequence, the essential question in the satellite setting is: What is the best secret-key rate for given antenna sizes of the honest parties if we are willing to assume an upper bound on Eve's antenna size, but consider the signal strength as a design parameter to maximize over?

1.4 Contributions

Quite surprisingly, it has been shown by Maurer and Wolf [11,14] that in the satellite model secret-key agreement is possible even if Eve's channel is almost perfect, i.e., if ϵ_E is arbitrarily close to 0 but not exactly 0, and if Alice's and Bob's channels have arbitrarily high error probability but still some information (i.e., ϵ_A and ϵ_B are close to 0.5 but not exactly 0.5). However, the lower bound for the secret-key rate obtained via the original repeater-code protocol in [11], when interpreted in terms of the ratio Q, is only exponentially small in Q. In contrast, the secret-key ratio as a function of Q has already been briefly considered by Maurer and Gander [6], who conjectured based on numerical results that the rate of the parity-check protocol (introduced in [16]) asymptotically decreases like $1/Q^2$, for a setting where Alice's and Bob's antennas are assumed to be of equal size.

As our main technical contribution, we prove that both the rate of the parity-check protocol and the optimal secret-key rate are indeed inversely proportional to Q^2 in Section 4. This matches the numerical results and the conjecture by Gander and Maurer. We point out that the lower bound on the secret-key rate is proved by showing that the parity-check protocol, which is an explicit and simple protocol, achieves this rate in the given setting, rather than providing a pure existence proof of a protocol achieving this rate.

In the full version [9], we also generalize the secret-key rate as a function of the antenna ratio Q to the case where Alice and Bob can have antennas of different sizes, by specifying well motivated and relevant quantities for both lower and upper bounds.

In addition, we consider the setting where the power consumption of the satellite is bounded; for instance, by the size of its solar panels. Nevertheless, we can adjust the energy used to broadcast each bit by adjusting the bit-rate, i.e., the number of bits broadcast per second, while maintaining a fixed power consumption. Hence, the energy used to broadcast each bit is inversely proportional to the bit-rate. This motivates the study of the secret-key rate per second rather than the secret-key rate per bit. In order to investigate the secret-key rate per second, we introduce a novel quantity that approximates it in Section 5. We then show that this quantity decreases inversely proportional to Q, rather than Q^2, which makes a significant difference, since Q must be assumed to be very large.

1.5 A Note on the Practicality of the Satellite Setting

While the satellite setting attempts to mimic a real-world scenario, it also abstracts away many practical issues which affect its immediate applicability. For instance, the satellite setting encodes some basic assumptions on the adversary that might not necessarily hold in practice, such as the assumption that Eve will quantize the signal she receives. Moreover, the setting basically assumes a passive adversary, by assuming that the adversary can neither influence the bits the honest parties receive from the satellite, nor tamper with their communication. While the former restriction could be translated into some sort of physical assumption, the authenticated communication is something that can easily be obtained in a separate step. We can allow Alice and Bob to start with a small shared secret-key, which they can then use to authenticate the channel with information-theoretic security [23]. In this case, the goal of a protocol is to amplify a short initial secret-key into a very long secret-key, like in quantum key distribution.

As a consequence, even if one could imagine proving stronger results that hold if the channels can be to a certain degree dependent, or consider a setting where the adversary tries to get an advantage by considering the actual analog signal she receives, we nevertheless believe that proving theoretical results in our setting is meaningful. Showing that the secret-key rate under a channel quality constraint is reasonably large, and that the rate of a simple protocol asymptotically behaves like the secret-key rate in this setting can be seen as a step towards showing that the satellite setting is practical. In short, we feel that the problem studied in this paper is one of the most relevant and natural scientific problems extractable from the general setting.

1.6 Related Work

There have been considerable efforts to find good approximations for the secret-key rate, both in the satellite setting and for more general probability distributions, and also for settings with more than three parties.

The first bounds on the secret-key rate were proved by Maurer [11,13], and by Ahlswede and Csiszár [1], who studied the secret-key rate when only one-way communication from Alice to Bob is allowed. Later, Maurer and Wolf [12] and Renner, Skripsky, and Wolf [20] introduced improved upper bounds for general distributions, called the *intrinsic mutual information* and the *reduced intrinsic mutual information*, respectively. Csiszár and Narayan [5] extended the study of the secret-key rate to settings with more than three parties, and exhibited connections between information-theoretic secret-key agreement and the problem of *communication for omniscience*. Then, Gohari and Anantharam [7] showcased new lower and upper bounds on the secret-key rate for an arbitrary number of parties, which in particular are strict improvements over the previously known bounds for our setting.

There has been some recent interest in the secret-key rate in the finite blocklength setting, where the number of available realizations (X, Y, Z) is bounded. Tyagi and Watanabe [24] showcase a connection between the secret-key rate in this setting and binary hypothesis testing, and use it to obtain an upper bound on the secret-key rate for a bounded number of realizations. Later, Hayashi, Tyagi, and Watanabe [8] used this connection to better understand how the gap between the secret-key rate in the finite blocklength and asymptotic settings decreases as the number of available realizations increases, for certain probability distributions.

For the satellite setting, there exist better lower bounds on the secret-key rate due to the study of several *advantage distillation protocols*. The first such protocol, called the *repeater-code protocol*, was introduced and studied by Maurer [16,11]. An improved version of this protocol, called the *parity-check protocol*, was studied by Gander and Maurer [16,6]. Later, Liu, Van Tilborg, and Van Dijk [10] proposed another protocol that seems to outperform the parity-check protocol. However, the rate achieved by the proposed protocol was only numerically computed in a simulation where Eve follows a certain fixed strategy, which is not known to be optimal. Furthermore, finding a clean expression for the rate of this protocol that can be analyzed (as is done for the rate of the parity-check protocol) appears infeasible, and so it is very difficult to extract tangible rate lower bounds, even when assuming that the proposed strategy for Eve is optimal.

The scenario where Alice, Bob, and Eve receive the random bits in the satellite setting through Gaussian channels, instead of binary symmetric channels, was first considered by Maurer and Wolf [12]. Later, Naito et al. [18] showed that Alice and Bob can extract more secret-key rate in the Gaussian scenario than in the BSC scenario, as they are able to make use of soft-decoding.

2 Preliminaries

2.1 Notation

We denote random variables by uppercase letters such as X, Y, and Z. We may denote sequences of random variables $X_1, X_2, \ldots, X_N$ as X^N. We say that $X_1, X_2, \ldots, X_N$ are i.i.d. if all the X_i are independent random variables and they all have the same distribution. Most sets are denoted by uppercase calligraphic letters such as $\mathcal{S}$. The set of real numbers is denoted by $\mathbb{R}$ and for a natural number $n \in \mathbb{N}$, $[n]$ denotes the set $\{1, \ldots, n\}$. Given a set $\mathcal{S}$, the size of $\mathcal{S}$ is denoted by $|\mathcal{S}|$. For a string $x \in \{0,1\}^*$, $|x|$ denotes the length of x. The (Hamming) weight of a string $x \in \{0,1\}^*$ is defined as $w(x) \coloneqq |\{i : x_i = 1\}|$, where x_i is the i-th entry of x. We denote the logarithm to the base 2 by log and the natural logarithm by ln. The closed interval in $\mathbb{R}$ between two real numbers a and b is denoted by $[a, b]$.

Given an event A, we denote the probability that A happens by $\Pr[A]$, which is the sum of the probabilities of all outcomes in event A. Given two events A and B, the probability that A and B happen simultaneously is denoted by $\Pr[A, B]$. The conditional probability of A given B, provided $\Pr[B] > 0$, is $\Pr[A|B] \coloneqq \frac{\Pr[A,B]}{\Pr[B]}$.

The probability distribution of a finite random variable X is denoted by P_X, and so $P_X(x)$ denotes the probability that X takes the value x. Given an event A, $P_{X|A}$ denotes the conditional probability distribution of X conditioned on A. For two finite random variables X and Y, $P_{X|Y}(\cdot, y)$ denotes the probability distribution of X conditioned on the event $Y = y$.

2.2 Information Theory

Throughout this paper we will make use of some fundamental concepts from information theory. We briefly define the required notions in this section; a more detailed exposition of this field can be found in [3].

Fix a finite random variable X with range $\mathcal{X}$. The *entropy of* X, denoted by $H(X)$, is defined as

$$H(X) \coloneqq -\sum_{x \in \mathcal{X}} P_X(x) \log P_X(x).$$

Intuitively, the entropy measures the uncertainty about a given random variable. In fact, a finite random variable X with range $\mathcal{X}$ satisfies $0 \leq H(X) \leq \log|\mathcal{X}|$ with equality in the lower bound if and only if $P_X(x) = 1$ for some $x \in \mathcal{X}$, and with equality in the upper bound if and only if X is uniform over $\mathcal{X}$. We call

$$h(p) \coloneqq -p \log(p) - (1-p) \log(1-p)$$

the *binary entropy function* and note that for a binary random variable X with $P_X(1) = p$ we have that $H(X) = h(p)$.

Given two finite random variables X and Y with ranges $\mathcal{X}$ and $\mathcal{Y}$, respectively, we define the *conditional entropy of X given Y*, denoted by $H(X|Y)$, as

$$H(X|Y) \coloneqq \sum_{y \in \mathcal{Y}} P_Y(y) H(X|Y = y).$$

Given an event A, $H(X|Y, A)$ is defined as

$$H(X|Y, A) \coloneqq \sum_{y \in \mathcal{Y}} P_{Y|A}(y) H(X|Y = y, A).$$

We define the *mutual information between X and Y*, denoted by $I(X;Y)$, as

$$I(X;Y) \coloneqq H(X) - H(X|Y).$$

Intuitively, the mutual information measures how independent two random variables are, and we have $I(X;Y) = 0$ if and only if X and Y are independent. Given an event A, $I(X;Y|A)$ is defined as

$$I(X;Y|A) \coloneqq H(X|A) - H(X|Y, A).$$

Finally, if additionally Z is a finite random variable with range $\mathcal{Z}$, the *conditional mutual information between X and Y given Z*, denoted by $I(X;Y|Z)$, is defined as

$$I(X;Y|Z) \coloneqq \sum_{z \in \mathcal{Z}} P_Z(z) I(X;Y|Z = z).$$

We will be dealing with a simple instance of a *discrete memoryless channel*. A discrete memoryless channel with input X and output W is characterized by a conditional probability distribution $P_{W|X}$. The term *memoryless* stems from the fact that the channel's output depends only on the current input, and so is independent of previous channel utilizations. The *binary symmetric channel with error probability* ϵ is the discrete memoryless channel with input $X \in \{0, 1\}$ and conditional probability distribution such that $P_{W|X}(b, b) = 1 - \epsilon$ and $P_{W|X}(1 - b, b) = \epsilon$ for $b \in \{0, 1\}$. Intuitively, the binary symmetric channel receives a bit as input and flips it with a certain error probability.

The *capacity* is a fundamental quantity associated to every channel. Informally, the capacity of a channel is the optimal rate at which one can communicate through the channel while ensuring that the decoding error probability goes to zero as the number of channel uses increases. Shannon [21] proved that the capacity of a channel $P_{W|X}$ is given by $\max_{P_X} I(X;W)$. In particular, it is easily shown that the capacity of the binary symmetric channel with error probability ϵ is $1 - h(\epsilon)$, where h is the binary entropy function.

3 Secret-Key Agreement by Public Discussion

In the following section, we revisit the basic models of information-theoretically secure secret-key agreement on which we will build in Sections 4 and 5.

3.1 The Source Model and the Secret-Key Rate

We study information-theoretic secret-key agreement, in which Alice and Bob want to agree on a shared secret-key, about which Eve has (almost) no information. To circumvent the trivial impossibility results, we consider the model introduced by Maurer [16,11], called *secret-key agreement by public discussion from common information.* In this model, we assume that in addition to a bidirectional authenticated noiseless channel, which Eve can listen in to but not tamper with, the parties also share some form of correlated randomness. More specifically, we will look at the setting where the correlated randomness of Alice, Bob, and Eve consists of several independent and identically distributed realizations of discrete random variables X, Y, and Z, respectively, distributed according to some joint probability distribution P_{XYZ}.

Remark 1. As already mentioned, the assumption that an authenticated channel exists between Alice and Bob is not a significant drawback in the model. We can allow Alice and Bob to start with a small shared secret-key, which they can then use to authenticate the channel with information-theoretic security [23]. In this case, the goal of a protocol is to amplify a short initial secret-key into a very long secret-key, analogous to quantum key distribution.

In this setting, the main quantity of interest is the maximal rate (per number of realizations of X, Y, and Z received) at which Alice and Bob can generate secret-key bits, about which Eve has almost no information, as a function of the probability distribution P_{XYZ}. We first define what we mean by a secret-key agreement protocol. The following definition is taken from [17], and we show in the full version [9] that it is actually a composable definition, and hence the obtained key can be securely used in any context.

Definition 1. *Given a finite probability distribution P_{XYZ}, an (N, R, ϵ)-secret-key agreement protocol for P_{XYZ} is an interactive protocol for Alice and Bob, who receive $X^N = (X_1, \ldots, X_N)$ and $Y^N = (Y_1, \ldots, Y_N)$, respecitvely, as input. Then they generate a communication transcript $C^M = (C_1, \ldots, C_M)$ (where M is also a random variable) by sending messages over authenticated channels in an alternating manner. After the interaction is finished, Alice and Bob produce outputs S_A and S_B over the finite range $\mathcal{S}$, respectively.*

We require that if for $i \in [N]$,[3] the random variables (X_i, Y_i, Z_i) are i.i.d. according to P_{XYZ}, then the following properties must hold:

1. $H(S_A) \geq N(R - \epsilon)$;
2. $H(S_A) \geq \log|\mathcal{S}| - \epsilon$;
3. $\Pr[S_A = S_B] \geq 1 - \epsilon$;
4. $I(S_A; Z^N C^M) \leq \epsilon$.

[3] We denote by $[n]$ the set $\{1, 2, \ldots, n\}$, see Section 2 for an exhaustive introduction on the notation we use.

Intuitively, property 1 in Definition 1 states that, on average, Alice and Bob extract at least $R - \epsilon$ secret bits per realization of (X, Y, Z), i.e., the rate is at least $R - \epsilon$. Property 2 enforces that S_A is almost uniform over $\mathcal{S}$, property 3 implies that S_A and S_B should coincide with high probability, and property 4 means that Eve's information, which consists of Z^N and the transcript C^M, gives almost no information about the secret keys S_A and S_B. We are now ready to define the secret-key rate.

Definition 2. *Given a finite probability distribution P_{XYZ}, the* secret-key rate for P_{XYZ} *(abbreviated as the* secret-key rate *when the context is clear), denoted by $S(X;Y\|Z)$, is the supremum of all real numbers R such that for all $\epsilon > 0$ and large enough N there exists an (N, R, ϵ)-secret-key agreement protocol for P_{XYZ}.*

The secret-key rate was first studied by Maurer [16,11], while Csiszár and Körner [1] studied the *one-way* secret-key rate, where only one-way communication from Alice to Bob is allowed.

The following theorem states basic bounds for the secret-key rate. The lower bound was proved by Maurer [11,13] and Csiszár and Körner [1], while the upper bound was proved by Maurer [11].

Lemma 1 ([11, Theorem 2] and [13, Theorem 4]). *For all finite probability distributions P_{XYZ}, we have*

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

Note that our definition of the secret-key rate corresponds to the so-called strong secret-key rate, which Maurer and Wolf [17] have proven to be equivalent to the weak one initially considered in the lower bounds.

3.2 A Special Case: The Satellite Setting

Our focus will lie on the secret-key rate of a conceptually simple, but realistic and interesting, class of distributions P_{XYZ}, named the *satellite setting*.

Fix real numbers $\epsilon_A, \epsilon_B, \epsilon_E \in [0, 1/2]$ and consider the following experiment:

1. Sample a bit $R \in \{0, 1\}$ uniformly at random;
2. Send R to Alice, Bob, and Eve through independent binary symmetric channels with error probabilities ϵ_A, ϵ_B, and ϵ_E, respectively. The random variables X, Y, and Z are the output of these three channels.

This class of distributions was introduced by Maurer [16,11]. The satellite setting earned its name because a realistic implementation of such a scenario would consist of having a satellite orbiting the Earth which broadcasts random bits. On the ground, Alice, Bob, and Eve would be in possession of their own antennas, which they can use to listen to the satellite broadcasts. The quality of a party's antenna would then dictate how reliably they receive the random bits from the satellite. For instance, a better antenna leads to a smaller error probability.

An additional surprising benefit of this model is that secret-key agreement is possible whenever it is not trivially impossible, as stated in the following theorem of Maurer and Wolf [11,14].

Theorem 1 ([14, Theorem 2, adapted]). *We have $S(X;Y\|Z) > 0$ if and only if $\epsilon_E > 0$ and $\epsilon_A, \epsilon_B < 1/2$.*

This stands in stark contrast to the well-known fact that secret-key agreement with one-way communication from Alice to Bob (in the sense of [1]) is impossible whenever Eve's antenna is better than both Alice's and Bob's antennas, i.e., whenever $\epsilon_E < \epsilon_A$ and $\epsilon_E < \epsilon_B$.

While Theorem 1 assures that the secret-key rate is positive in all non-trivial settings, computing (or even approximating) it has proven to be a surprisingly difficult problem for most parameters ϵ_A, ϵ_B, and ϵ_E.

3.3 Advantage Distillation Protocols

In the following section, we present some required background to understand the proofs in Sections 4 and 5, and in particular we introduce the parity-check protocol that we use to lower bound the secret-key rate.

The parity-check protocol is an example of a so-called *advantage-distillation protocol*, which is a type of protocol introduced in [11,14] to prove Theorem 1 in the satellite setting.

Definition 3. *Let P_{XYZ} denote a finite probability distribution. An* advantage-distillation protocol for P_{XYZ} *is then an interactive protocol for Alice and Bob, who receive $X^N = (X_1, \ldots, X_N)$ and $Y^N = (Y_1, \ldots, Y_N)$, respectively, as input for some N. Then they generate a communication transcript $C^M = (C_1, \ldots, C_M)$ by sending messages over authenticated channels in an alternating manner. Afterwards, Alice and Bob produce outputs $\hat{X}$ and $\hat{Y}$, respectively.*

For all large enough N, if the random variables (X_i, Y_i, Z_i) are i.i.d. according to P_{XYZ}, we require that

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

where $\hat{Z} = (Z^N, C^M)$ denotes Eve's total information at the end of the protocol.

Intuitively, Bob ends up with more information about Alice than Eve does, and so the protocol "distills" an advantage for Alice and Bob over Eve.

Note that such an advantage-distillation protocol itself is not a secret-key agreement protocol according to Definition 1, as it neither guarantees that Alice and Bob output the same key, nor guarantees that Eve has arbitrary small information about Alice's output. However, for any probability distribution P_{XYZ} and advantage-distillation protocol, we can consider the induced probability distribution $P_{\hat{X}\hat{Y}\hat{Z}}$ from running the protocol on N i.i.d. realizations of P_{XYZ}, and then simply apply a secret-key agreement protocol for this distribution. Along this line, we can then also introduce the secret-key rate of an advantage-distillation protocol.

Definition 4. *Given a finite probability distribution P_{XYZ} and an advantage-distillation protocol, the* secret-key rate of the advantage-distillation protocol for P_{XYZ} *is the supremum of all real numbers R such that for all $\epsilon > 0$ and large enough N there exists a secret-key agreement protocol for $P_{\hat{X}\hat{Y}\hat{Z}}$, such that the composed protocol is an (N, R, ϵ)-secret-key agreement protocol for P_{XYZ}.*

The existence of an advantage-distillation protocol implies $S(X;Y\|Z) > 0$, since we have

$$S(X;Y\|Z) \geq \frac{S(\hat{X};\hat{Y}\|\hat{Z})}{N} \geq \frac{I(\hat{X};\hat{Y}) - I(\hat{X};\hat{Z})}{N} > 0,$$

where the second inequality follows from Lemma 1.

The Repeater-Code Protocol The first advantage distillation protocol was the *repeater-code protocol* [16,11]. It works as follows:

1. Alice samples $R \in \{0,1\}$ uniformly at random and sends $R \oplus X^N = (R \oplus X_1, \ldots, R \oplus X_N)$ to Bob over the authenticated channel;
2. Bob computes $R \oplus X^N \oplus Y^N = (R \oplus X_1 \oplus Y_1, \ldots, R \oplus X_N \oplus Y_N)$ and sets $A = 1$ if $R \oplus X^N \oplus Y^N = 0^N$ or $R \oplus X^N \oplus Y^N = 1^N$. Otherwise, Bob sets $A = 0$. Then, Bob sends A to Alice through the authenticated channel;
3. If $A = 1$, then Alice sets $\hat{X} = R$ and Bob sets $\hat{Y} = R \oplus X_1 \oplus Y_1$. Otherwise, if $A = 0$, then Alice and Bob set $\hat{X} = \hat{Y} = \perp$.

Maurer and Wolf [14] proved that, in the satellite setting, for all triples $(\epsilon_A, \epsilon_B, \epsilon_E)$ with $\epsilon_A < 1/2$, $\epsilon_B < 1/2$, and $\epsilon_E > 0$ and for N large enough we have

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

where $\hat{Z} \coloneqq (Z^N, R \oplus X^N, A)$ denotes Eve's total information.

While the repeater-code protocol is good enough to prove that secret-key agreement is possible in the satellite setting, it guarantees only a very small lower bound on the secret-key rate, especially when ϵ_A and ϵ_B are much larger than ϵ_E. This issue motivated the search for better advantage distillation protocols in the satellite setting.

The Parity-Check Protocol Gander and Maurer [16,6] studied an improved protocol, called the *parity-check protocol*. The parity-check protocol with ℓ rounds works as follows:

1. Alice and Bob start with initially empty strings U_A and U_B, respectively;
2. Alice and Bob divide X^N and Y^N into pairs (X_{2i-1}, X_{2i}) and (Y_{2i-1}, Y_{2i}), respectively, for $i = 1, \ldots, \lfloor N/2 \rfloor$;
3. For each i, Alice sends $X_{2i-1} \oplus X_{2i}$ to Bob via the authenticated channel;
4. Bob sets $A_i = 1$ if $X_{2i-1} \oplus X_{2i} = Y_{2i-1} \oplus Y_{2i}$. Otherwise, Bob sets $A_i = 0$. Then, he sends A_i to Alice;

5. If $A_i = 1$, Alice adds X_{2i-1} to her string U_A and Bob adds Y_{2i-1} to his string U_B, and they discard X_{2i} and Y_{2i}, respectively (i.e., these bits are not added to U_A and U_B, respectively). If $A_i = 0$, Alice and Bob discard the bits (X_{2i-1}, X_{2i}) and (Y_{2i-1}, Y_{2i}), respectively;
6. If $\ell = 1$, then Alice and Bob stop the protocol. Alice sets $\hat{X} = U_A$ and Bob sets $\hat{Y} = U_B$;
7. If $\ell > 1$ and $|U_A| \geq 2^{\ell-1}$, Alice and Bob run the parity-check protocol with $\ell - 1$ rounds on the strings U_A and U_B. Otherwise, if $|U_A| < 2^{\ell-1}$, then Alice and Bob set $\hat{X} = \perp$ and $\hat{Y} = \perp$, respectively.

If $\hat{X}$ and $\hat{Y}$ are the outputs of the parity-check protocol with ℓ rounds, then each pair of bits $(\hat{X}_i, \hat{Y}_i)$ behaves like the output of a successful run of the repeater-code protocol with $N \coloneqq 2^\ell$. Furthermore, all pairs $(\hat{X}_i, \hat{Y}_i)$ are identically distributed and independent of each other.

Again, consider the satellite setting and assume, without loss of generality, that $\epsilon_A \geq \epsilon_B$. Analogous to [6], let us now introduce a couple of useful quantities in the setting of running the parity-check protocol.

Definition 5. *Consider the satellite setting with error probabilities ϵ_A, ϵ_B, and ϵ_E respectively. Let (X, Y, Z) be distributed according to the thereby induced distribution P_{XYZ}. Then we define*

$$\beta \coloneqq \Pr[X \neq Y] = \epsilon_A(1 - \epsilon_B) + (1 - \epsilon_A)\epsilon_B$$

and for $r, s \in \{0, 1\}$

$$\alpha_{rs} \coloneqq \Pr[X \oplus Y = r, X \oplus Z = s],$$

which satisfy

$$\begin{aligned}
\alpha_{00} &= \epsilon_A\epsilon_B\epsilon_E + (1 - \epsilon_A)(1 - \epsilon_B)(1 - \epsilon_E) \\
\alpha_{01} &= \epsilon_A\epsilon_B(1 - \epsilon_E) + (1 - \epsilon_A)(1 - \epsilon_B)\epsilon_E \\
\alpha_{10} &= \epsilon_A(1 - \epsilon_B)\epsilon_E + (1 - \epsilon_A)\epsilon_B(1 - \epsilon_E) \\
\alpha_{11} &= \epsilon_A(1 - \epsilon_B)(1 - \epsilon_E) + (1 - \epsilon_A)\epsilon_B\epsilon_E.
\end{aligned}$$

Moreover, considering L independent draws from P_{XYZ}, and let

$$\beta_L \coloneqq \Pr[X^L \oplus Y^L = 1^L | X^L \oplus Y^L \in \{0^L, 1^L\}] = \frac{\beta^L}{\beta^L + (1 - \beta)^L},$$

and $p_{L,w}$ denote the probability that $X^L \oplus Y^L \in \{0^L, 1^L\}$ and $X^L \oplus Z^L$ is a specific codeword of Hamming weight w, i.e.,

$$p_{L,w} \coloneqq \alpha_{00}^{L-w}\alpha_{01}^{w} + \alpha_{10}^{L-w}\alpha_{11}^{w}.$$

Using those quantities, we can now express the secret-key rate of the parity-check protocol.

Theorem 2 (rephrased form [6]). *Let $R(\ell, \epsilon_A, \epsilon_B, \epsilon_E)$ denote the secret-key rate of the parity-check protocol when using ℓ rounds, and Alice, Bob, and Eve having error probabilities ϵ_A, ϵ_B, and ϵ_E, respectively. We then have*

$$R(\ell, \epsilon_A, \epsilon_B, \epsilon_E) \geq 2^{-\ell}\, \Phi(2^\ell, \epsilon_A, \epsilon_B, \epsilon_E) \prod_{i=0}^{\ell-1} \left(\beta_{2^i}^2 + (1-\beta_{2^i})^2\right),$$

where

$$\Phi(L, \epsilon_A, \epsilon_B, \epsilon_E) \coloneqq \sum_{w=0}^{L} \binom{L}{w} \frac{p_{L,w}}{\beta^L + (1-\beta)^L} h\left(\frac{p_{L,w}}{p_{L,w} + p_{L,L-w}}\right) - h(\beta_L),$$

and β, β_L, and $p_{L,w}$ are according to Definition 5.

The intuition behind Theorem 2 is the following: Suppose there are N_i bits left after i rounds of the parity-check protocol. These N_i bits are partitioned into $\lfloor N_i/2 \rfloor$ pairs (if N_i is even, Alice and Bob discard a bit), and, in round $i+1$, Alice and Bob keep a bit from a given pair with probability $\beta_{2^i}^2 + (1-\beta_{2^i})^2$. Therefore, we have

$$\mathrm{E}[N_{i+1} \mid N_i \text{ bits after } i \text{ rounds}] \approx \frac{\beta_{2^i}^2 + (1-\beta_{2^i})^2}{2} \cdot N_i,$$

where N_{i+1} is the random variable denoting the number of bits after $i+1$ rounds of the parity-check protocol.

The lower bound on the secret-key rate obtained through the parity-check protocol is, for most choices of error probabilities in the satellite setting, much better than the lower bound given by the repeater-code protocol. Note that the parity-check protocol consists of the iterative application of the repeater-code protocol with length 2 to pairs of bits of X^N and Y^N. This protocol can be further improved in a natural way for some interesting choices of error probabilities in the satellite setting by modifying the length of the repeater-code protocol that is applied iteratively, and reutilizing discarded bits from failed runs of the repeater-code protocol which are "almost" successful. We do not expand on this, since the original parity-check protocol suffices for our needs.

4 The Secret-Key Rate Under a Fixed Channel Quality Ratio

4.1 Modeling a Fixed Channel Quality Ratio

In this section, we formally define the main quantity used in this work. Recall that we want to consider a setting where we assume that the antenna sizes of the honest parties are fixed, but where the energy the satellite uses to send a bit is a design parameter that we can adjust in order to achieve an optimal secret-key rate. To obtain a meaningful lower bound on the secret-key rate in this setup,

however, we need to make an assumption about Eve's capabilities, which in the satellite setting correspond to her antenna size. In order to simplify the model, we moreover do not consider the actual antenna sizes, but the ratio between Eve's antenna size and Alice's and Bob's. Therefore, in the following we want to assume that Eve's antenna is exactly Q times larger than both Alice's and Bob's antennas. For ease of exposition, we will also assume that Alice and Bob have antennas of the same size. This was the setting considered by Gander and Maurer [6]. In the full version [9], we analyze the general setting where Alice's and Bob's antennas may differ in size, and also where Eve's antenna is only known to be *at most*, or *at least*, Q times larger than Alice's and Bob's, instead of exactly Q times larger.

To model the antenna size ratio, we choose the ratio of the channel capacities, which reflect the qualities of the respective channels. Recall that the satellite model with BSC's is a simplification of the more realistic analog model with Additive White Gaussian Noise (AWGN) channels (if the channel input is X, then the output is $X + Z$, where Z is distributed according to a normal distribution with mean zero and variance N, where N is also called the noise power). It is well-known that the capacity (in bits per second) of an AWGN channel is given by $C_{\mathsf{AWGN}} = B \log(1 + S/N)$, where B is the bandwidth (in the spectrum), S is the signal power, and N is the noise power (see [3, Chapter 9]). The signal power is proportional to the total antenna surface, independently of whether the antenna consists of several independent small antennas or one large one. In the low-signal regime, i.e., if $S/N \ll 1$, we have that C is essentially proportional to S (for fixed noise power N), and hence to the antenna size too. In short, in such a regime the channel capacity is essentially proportional to the product of the surface of the receiver's antenna and the energy used to transmit the bits. Hence, when considering two of the antennas, the ratio of their capacity stays approximately constant when adjusting the energy that is used to transmit each bit and, therefore, this ratio is a good approximation of the ratio of the antenna sizes.

While this justification is based on the AWGN model, we assume that it essentially carries over to the BSC model of the satellite setting. Observe that the satellite setting using BSC's can be interpreted in a natural way as a version of the satellite setting with AWGN channels where Alice, Bob, and Eve quantize the signals they receive.

This leads us to the following definition of the channel quality ratio between two binary symmetric channels.

Definition 6. *The* channel quality ratio between the BSC with error probability α and the BSC with error probability γ, *denoted* $\rho(\alpha, \gamma)$, *is defined as*

$$\rho(\alpha, \gamma) \coloneqq \frac{1 - h(\gamma)}{1 - h(\alpha)}.$$

Assume a fixed antenna size ratio Q between Eve's and Alice's antennas, and hence between Eve's and Bob's antennas, since Alice and Bob are assumed to have antennas of the same size. Considering the energy spent per bit as a design

parameter then corresponds to freely choosing $\alpha = \epsilon_A = \epsilon_B$ and $\gamma = \epsilon_E$ under the constraint that $\rho(\alpha, \gamma) = Q$. This leads to the following definition, where the supremum corresponds to choosing the energy per bit in an optimal manner.

Definition 7. *The* secret-key rate for an adversary with an exactly Q times better channel*, denoted by $S(Q)$, is defined as*

$$S(Q) \coloneqq \sup_{\substack{\alpha,\gamma \\ \rho(\alpha,\gamma)=Q}} S(\alpha, \alpha, \gamma).$$

In the following sections, we will give an exact characterization (up to a multiplicative constant) of the asymptotic behavior of $S(Q)$ when Q increases. In particular, we settle the conjecture of Gander and Maurer [6] in the affirmative.

4.2 A Lower Bound on $S(Q)$

Our first main result is that $S(Q)$ decreases at most inversely proportional to Q^2. We omit or shorten most proofs of intermediate results in this section. Detailed proofs can be found in the full version [9].

Theorem 3. *There exist a constant $c > 0$ such that*

$$\frac{c}{Q^2} \leq S(Q)$$

for all $Q \geq 1$.

To prove this result, we actually show that the parity-check protocol [16] (c.f. Section 3.3) achieves this rate, which was first conjectured to be true by Gander and Maurer [6], based on numerical evidence.

Definition 8. *The* secret-key rate of the parity-check protocol for an adversary with an exactly Q times better channel*, denoted by $R(Q)$, is defined as*

$$R(Q) \coloneqq \sup_{\substack{\ell,\alpha,\gamma \\ \rho(\alpha,\gamma)=Q}} R(\ell, \alpha, \alpha, \gamma),$$

where $R(\ell, \epsilon_A, \epsilon_B, \epsilon_E)$ denotes the rate per random bit achieved by the parity-check protocol using ℓ rounds when Alice, Bob, and Eve have error probabilities ϵ_A, ϵ_B, and ϵ_E, respectively.

Since the secret-key rate $S(\epsilon_A, \epsilon_B, \epsilon_E)$ is defined as the secret-key rate of the best possible protocol, we trivially get the following lower bound.

Lemma 2. *Let $Q \geq 1$. Then, we have $R(Q) \leq S(Q)$.*

We now proceed by proving that there exists a constant $c > 0$ such that $\frac{c}{Q^2} \leq R(Q)$ for all $Q \geq 1$, which will eventually conclude the proof. In order to prove such a lower bound, we need to lower bound the supremum in the definition of $R(Q)$. We achieve this by carefully choosing a sequence of triples

$(\ell_k, \alpha_k, \gamma_k)$ such that $R\left(\frac{1-h(\gamma_k)}{1-h(\alpha_k)}\right)$ does not decrease too quickly when compared to $\frac{1-h(\gamma_k)}{1-h(\alpha_k)}$. Namely, in the first step we will show that

$$R\left(\frac{1-h(\gamma_k)}{1-h(\alpha_k)}\right) \geq \frac{c_1}{k^4}$$

for some constant $c > 0$, and then in a second step use that $\frac{1-h(\gamma_k)}{1-h(\alpha_k)}$ increases like k^2, in order to derive the desired result.

Lower bounding the secret-key rate of the parity-check protocol with concrete parameters. In this section we show that for $\ell_k = 2\log(k)$ rounds, in the satellite setting with $\epsilon_A = \epsilon_B = \alpha_k = 1/2 - 1/k$, and $\epsilon_E = \gamma_k = 2/5$, the secret-key rate of the parity-check protocol $R(\ell_k, \alpha_k, \alpha_k, \gamma_k)$ decreases inversely proportional to k^4. For simplicity, we drop the subscript k in most terms from now on.

Before deriving the actual lower bound on $R(\ell, \alpha, \alpha, \gamma)$, we introduce an auxiliary quantity and prove some properties about it. Recall the definition of α_{rs} for $r, s \in \{0, 1\}$ and $p_{L,w}$ from Definition 5 in Section 3.3. In the following, let

$$p'_{L,w} \coloneqq \alpha_{00}^{L-w}\alpha_{01}^{w}. \tag{1}$$

We now present a few lemmas about $p_{L,w}$, $p'_{L,w}$, and their relation.

Lemma 3. *Let $\alpha = \epsilon_A = \epsilon_B = 1/2 - 1/k$. Then we have*

$$p_{L,w} = \alpha_{00}^{L-w}\alpha_{01}^{w} + (\alpha(1-\alpha))^L = p'_{L,w} + (\alpha(1-\alpha))^L > p'_{L,w}.$$

Lemma 4. *Let $p'_{L,w}$ as defined in (1). Then $p'_{L,w}$ is equal to the probability that $X^L \oplus Z^L$ is a particular codeword of weight w and $X^L = Y^L$, i.e. for any $c \in \{0,1\}^L$ with $w(c) = w$, where $w(c)$ denotes the Hamming weight of c, we have*

$$\Pr[X^L \oplus Z^L = c, X^L = Y^L] = p'_{L,w}.$$

Lemma 5. *We have*

$$h\left(\frac{p_{L,w}}{p_{L,w} + p_{L,L-w}}\right) \geq h\left(\frac{p'_{L,w}}{p'_{L,w} + p'_{L,L-w}}\right)$$

for all L and w.

Lemma 6. *Let $0 \leq \delta \leq L/2$. Then*

$$\frac{p'_{L,L/2+\delta}}{p'_{L,L/2-\delta}} = \left(\frac{\alpha_{01}}{\alpha_{00}}\right)^{2\delta}.$$

Lemma 7. *For all $L/2 \geq x \geq y \geq 0$ the following two properties hold*

1. $h\left(\frac{p'_{L,L/2-x}}{p'_{L,L/2-x}+p'_{L,L/2+x}}\right) \leq h\left(\frac{p'_{L,L/2-y}}{p'_{L,L/2-y}+p'_{L,L/2+y}}\right)$
2. $h\left(\frac{p'_{L,L/2-x}}{p'_{L,L/2-x}+p'_{L,L/2+x}}\right) = h\left(\frac{p'_{L,L/2+x}}{p'_{L,L/2+x}+p'_{L,L/2-x}}\right)$.

Next, we lower bound $R(\ell, \alpha, \alpha, \gamma)$, i.e., the rate of the parity-check protocol when using ℓ rounds, Alice and Bob have the same error probability α, and Eve has error probability γ, in a sequence of lemmas.

Lemma 8. *For all $k \in \{2^j : j \in \mathbb{N}\}$, let $\ell_k = 2\log(k)$, $\alpha_k = 1/2 - 1/k$, and $\gamma_k = 2/5$. We then have*

$$R(\ell_k, \alpha_k, \alpha_k, \gamma_k) \geq \frac{1}{k^4}\Phi(k^2, \alpha_k, \alpha_k, \gamma_k),$$

where Φ is defined as in Theorem 2.

Lemma 9. *For $k \in \{2^j : j \in \mathbb{N}\}$, let $\ell_k = 2\log(k)$, $\alpha_k = 1/2 - 1/k$, and $\gamma_k = 2/5$. Then there exists a positive constant $c > 0$ such that*

$$\Phi(k^2, \alpha_k, \alpha_k, \gamma_k) \geq c$$

for large enough $k \in \{2^j : j \in \mathbb{N}\}$, where Φ is defined as in Theorem 2.

Proof. We present a sketch of the proof. The complete proof can be found in the full version [9]. First, it holds that

$$\lim_{k\to\infty} h\left(\beta_{k^2}\right) = h\left(\lim_{k\to\infty} \frac{1}{1+(1+8/k^2)^{k^2}}\right) = h\left(\frac{1}{1+e^8}\right) < 5\cdot 10^{-3}. \tag{2}$$

Furthermore, using Lemmas 3 and 5 to 7 it can be seen that

$$\begin{aligned}\sum_{w=0}^{k^2} \binom{k^2}{w} \frac{p_{k^2,w}}{\beta^{k^2}+(1-\beta)^{k^2}} \cdot h\left(\frac{p_{k^2,w}}{p_{k^2,w}+p_{k^2,k^2-w}}\right) \\ \geq \frac{1}{2} \sum_{w=k^2(1/2-2/k)}^{k^2(1/2+2/k)} \binom{k^2}{w} \frac{p'_{k^2,w}}{(1-\beta)^{k^2}} \cdot h\left(\frac{1}{1+\left(\frac{\alpha_{01}}{\alpha_{00}}\right)^{4k}}\right)\end{aligned} \tag{3}$$

for large enough k.

In order to lower bound the binary entropy term in (3), we can use the definition of α_{rs} (recall Definition 5) to show that

$$\lim_{k\to\infty} h\left(\frac{1}{1+\left(\frac{\alpha_{01}}{\alpha_{00}}\right)^{4k}}\right) = h\left(\frac{1}{1+e^{-32/5}}\right) > 1.7\cdot 10^{-2}. \tag{4}$$

We now define $W \coloneqq (w(X^{k^2} \oplus Z^{k^2}) \mid X^{k^2} = Y^{k^2})$ as the random variable denoting the Hamming weight of $X^{k^2} \oplus Z^{k^2}$ conditioned on $X^{k^2} = Y^{k^2}$, i.e., W is defined in the modified random experiment obtained by conditioning on $X^{k^2} = Y^{k^2}$. Through Lemma 4, we have

$$\sum_{w=k^2(1/2-2/k)}^{k^2(1/2+2/k)} \binom{k^2}{w} \frac{p'_{k^2,w}}{(1-\beta)^{k^2}} = \Pr[|W - k^2/2| \leq 2k]. \tag{5}$$

It suffices now to find a suitable lower bound for $\Pr[|W - k^2/2| \leq 2k]$. In order to do that, we will apply Chebyshev's inequality. It can be shown that

$$\frac{k^2}{2} - 2k \leq \mathrm{E}[W] - k \leq \mathrm{E}[W] + k \leq \frac{k^2}{2} + 2k,$$

and hence

$$\Pr[|W - k^2/2| \leq 2k] \geq \Pr[|W - \mathrm{E}[W]| \leq k] \geq 1 - \frac{\mathrm{Var}[W]}{k^2} \geq \frac{3}{4}, \tag{6}$$

where the second inequality follows from Chebyshev's inequality, and the third inequality follows from the fact that $\mathrm{Var}[W] \leq k^2/4$.

Combining (3), (4), (5), and (6) yields

$$\sum_{w=0}^{k^2} \binom{k^2}{w} \frac{p_{k^2,w}}{\beta^{k^2} + (1-\beta)^{k^2}} \cdot h\left(\frac{p_{k^2,w}}{p_{k^2,w} + p_{k^2,k^2-w}}\right)$$
$$> \frac{1}{2} \cdot \frac{3}{4} \cdot 1.7 \cdot 10^{-2} > 5 \cdot 10^{-3} > h(\beta_{k^2})$$

for large enough $k \in \{2^j : j \in \mathbb{N}\}$, which concludes the proof. □

Combining Lemmas 8 and 9 yields the main result of this subsection.

Lemma 10. *For all $k \in \{2^j : j \in \mathbb{N}\}$, let $\ell_k = 2\log(k)$, $\alpha_k = 1/2 - 1/k$, and $\gamma_k = 2/5$. Then there exists a constant $c > 0$ such that we have*

$$R(\ell_k, \alpha_k, \alpha_k, \gamma_k) \geq \frac{c}{k^4}$$

for large enough $k \in \{2^j : j \in \mathbb{N}\}$.

Deriving a lower bound in Q. It now remains to show that Lemma 10 actually implies the desired lower bound in Q. We can prove this by using the fact that $\frac{1-h(\gamma_k)}{1-h(\alpha_k)}$ increases like k^2, and then substituting this term by Q.

Lemma 11. *For all $k \in \{2^j : j \in \mathbb{N}\}$, let $\ell_k = 2\log(k)$, $\alpha_k = 1/2 - 1/k$, and $\gamma_k = 2/5$. We then have*

$$R\left(\frac{k^2}{200}\right) \geq R(\ell_k, \alpha_k, \alpha_k, \gamma_k).$$

We are now ready to prove Theorem 3 by substituting $k^2/200$ in place of Q.

Proof (Theorem 3). Follows by combining Lemmas 10 and 11, and extending the result to all $Q \geq 1$. □

4.3 An Upper Bound on $S(Q)$

As a second main result, we show that $S(Q)$ decreases at least inversely proportional to Q^2.

Theorem 4. *We have*

$$S(Q) \leq \frac{4\ln(2)^2}{Q^2} < \frac{2}{Q^2}$$

for all $Q \geq 1$.

Before we can prove Theorem 4, we need the following auxiliary result.

Lemma 12 ([2, Theorem 2.2]). *If* $p = 1/2 - \epsilon$, *we have*

$$\frac{2\epsilon^2}{\ln(2)} \leq 1 - h(p) \leq 4\epsilon^2.$$

We now proceed by showing two lemmas that we will reuse later.

Lemma 13. *Let* $Q \geq 1$, $\alpha, \gamma \in [0, 1/2]$ *such that* $\frac{1-h(\gamma)}{1-h(\alpha)} = Q$, *and* $\delta \coloneqq 1/2 - \alpha$. *We then have*

$$S(\alpha, \alpha, \gamma) \leq 16\delta^4.$$

Proof. Note that

$$S(\alpha, \alpha, \gamma) \leq I(X;Y) = 1 - h(\beta),$$

where X and Y are Alice's and Bob's random variables in the satellite setting with $\epsilon_A = \epsilon_B = \alpha$, and, as before, $\beta \coloneqq \Pr[X \neq Y] = 2\alpha(1-\alpha)$. Since $\beta = 2\alpha(1-\alpha) = 1/2 - 2\delta^2$, using $\epsilon \coloneqq 2\delta^2$, it follows by Lemma 12 that

$$1 - h(\beta) \leq 16\delta^4,$$

concluding the proof. □

It remains to bound δ^4 by a function of Q.

Lemma 14. *Let* $Q \geq 1$, $\alpha, \gamma \in [0, 1/2]$ *such that* $\frac{1-h(\gamma)}{1-h(\alpha)} = Q$, *and* $\delta \coloneqq 1/2 - \alpha$. *We then have*

$$2\delta^2 \leq \frac{\ln(2)}{Q}.$$

Proof. Using Lemma 12 we obtain

$$\frac{2\delta^2}{\ln(2)} \leq 1 - h(\alpha) = \frac{1 - h(\gamma)}{Q} \leq \frac{1}{Q}.$$

□

We are now ready to conclude the overall proof of Theorem 4.

Proof (Theorem 4). Combining Lemmas 13 and 14 yields

$$S(Q) = \sup_{\substack{\alpha,\gamma \\ \rho(\alpha,\gamma)=Q}} S(\alpha, \alpha, \gamma) \leq 16\delta^4 \leq \frac{4\ln(2)^2}{Q^2} < \frac{2}{Q^2}$$

for all $Q \geq 1$. □

4.4 Combining all bounds

Recall that, by Definition 7, $S(Q)$ denotes the secret-key rate in the setting where Alice's and Bob's channels are identical (i.e., $\epsilon_A = \epsilon_B$ always), and Eve's channel is exactly Q times better than both of Alice's and Bob's. Moreover, by Definition 8, $R(Q)$ denotes the secret-key rate of the parity-check protocol in the same setting. In Sections 4.2 and 4.3, we have overall proven the following bounds on $S(Q)$ and $R(Q)$:

$$\frac{c}{Q^2} \leq R(Q) \leq S(Q) \leq \frac{2}{Q^2}$$

for some $c > 0$. Thus, in this setting we have determined the secret-key rate $S(Q)$ up to a multiplicative constant, and on the way proved the conjecture by Gander and Maurer.

Corollary 1. *We have $S(Q) = \Theta(1/Q^2)$. Moreover, the parity-check protocol from [6] achieves rate $\Omega(1/Q^2)$ in this setting.*

5 The Secret-Key Rate per Second Under a Fixed Channel Quality Ratio

In this section, we consider the scenario where the power consumption of the satellite is bounded; for instance, due to the size of its solar panels. Nevertheless, we can adjust the energy used to broadcast each bit by adjusting the bit-rate, i.e., the number of bits broadcast per second, while maintaining a fixed power consumption. In this setting, the natural quantity to optimize for is clearly the secret-key rate per second, rather than the secret-key rate per random bit.

5.1 Defining the Secret-Key Rate per Second

When defining the secret-key rate per second in the satellite model, there is one inherent issue: the abstraction using BSC's instead of AWGN channels actually abstracted away any notion of time. Hence, to nevertheless devise a quantity that can serve as a heuristic of the secret-key rate per second, expressed as a function of the error probabilities, we once again consider the AWGN setting. In contrast to the capacity of the BSC, which is measured as the number of bits that can be reliably transmitted per bit sent, the capacity of an AWGN channel is measured in bits that can be reliably transmitted per second.

As mentioned in Section 4.1, the capacity of an AWGN channel is $C_{\mathsf{AWGN}} = B\log(1 + S/N)$, where B is the bandwidth (in the spectrum), S is the signal power, and N is the noise power. Importantly, the capacity of the AWGN channel is a physical property of the channel that is not influenced by the way we encode and decode. A BSC can be seen as an AWGN channel where all parties perform hard decoding, i.e., measure the signal over a given interval in time and output a 1 if the average value in this time is above a certain threshold and 0

otherwise. Hence, we can also look at the capacity per bit in the AWGN model by normalizing by the "bit-rate", meaning the number of bits the parties output per second when applying their hard decoding. If we now double this bit-rate, the capacity per second has to remain constant, hence the capacity per bit must decrease by a factor of two. Therefore, the capacity per bit is inversely proportional to the bit-rate. Moreover, this capacity per bit of the AWGN channel roughly corresponds to the capacity of the BSC, as long as hard decoding is not too far from an optimal encoding scheme, which is the case in a regime with small signal to noise ratio. Thus, for a BSC with significant error probabilities, the capacity is inversely proportional to the bit-rate. This implies that, asymptotically, the secret-key rate per second, which is equal to the secret-key rate per bit times the bit-rate, behaves like the secret-key rate per bit divided by the capacity of the binary symmetric channel.

As a consequence, we can define the secret-key rate per second by dividing the secret-key rate per bit by the capacity of the honest parties' channel, which we assume to have larger error probabilities than Eve's channel, and hence deliver the better approximation.

Definition 9. *The* secret-key rate per second for an adversary with an exactly Q times better channel*, denoted by $S^*(Q)$, is defined as*

$$S^*(Q) \coloneqq \sup_{\substack{\alpha,\gamma \\ \rho(\alpha,\gamma)=Q}} \frac{S(\alpha,\alpha,\gamma)}{1-h(\alpha)}.$$

where $S(\epsilon_A,\epsilon_B,\epsilon_E)$ is the secret-key rate of the satellite setting with error probabilities ϵ_A, ϵ_B, and ϵ_E for Alice, Bob, and Eve, respectively.

5.2 Bounds on the Secret-Key Rate per Second

In this section, we establish the exact asymptotic behavior of $S^*(Q)$ as a function of Q, up to a multiplicative constant. For the lower bound, we will, analogously to Section 4.2, make use of the fact that the secret-key rate achieved by the parity-check protocol is a lower bound of the secret-key rate. Therefore, we also introduce the secret-key rate per second of the parity-check protocol.

Definition 10. *The* secret-key rate per second of the parity-check protocol for an adversary with an exactly Q times better channel*, denoted by $R^*(Q)$, is defined as*

$$R^*(Q) \coloneqq \sup_{\substack{\ell,\alpha,\gamma \\ \rho(\alpha,\gamma)=Q}} \frac{R(\ell,\alpha,\alpha,\gamma)}{1-h(\alpha)}.$$

where $R(\ell,\epsilon_A,\epsilon_B,\epsilon_E)$ denotes the rate per random bit achieved by the parity-check protocol using ℓ rounds.

We then obtain the following asymptotically exact characterization of the secret-key rate per second.

Theorem 5. *There exist constants* $c_1, c_2 > 0$ *such that*

$$\frac{c_1}{Q} \leq R^*(Q) \leq S^*(Q) \leq \frac{c_2}{Q}$$

for all $Q \geq 1$.

6 Conclusions and Open Problems

In this paper we investigated the secret-key rate in the satellite setting with the additional property that the satellite can freely choose the energy spent when transmitting a bit. In order to study this setting, we assumed there is a "quality ratio" Q between Eve's and the honest parties' antennas, which is an intrinsic property of the system that must stay fixed over all possible choices for the satellite. We model this quality ratio as the ratio of the capacities of the BSC's associated to Eve and the honest parties. Therefore, in our model, the extra degree of freedom for the satellite means that he can choose the error probabilities for Eve and the honest parties as long as the BSC's induced by them have capacity ratio Q. This setting was briefly considered for the first time by Gander and Maurer [6].

We motivated and introduced the quantity $S(Q)$ as a secret-key rate measure for the modified satellite setting just described. While even approximating the secret-key rate of the original satellite setting appears to be very complex, we are actually able to show that $S(Q) = \Theta(1/Q^2)$ when Q grows. This proves a conjecture of Gander and Maurer [6]. The mild decrease of the secret-key rate as a function of Q, coupled with the fact that our lower bound is obtained by considering a simple, explicit advantage distillation protocol, can be interpreted as a first step towards showing that information-theoretic secret-key agreement may be more practical than what is usually believed. We also propose a heuristic definition of the secret-key rate per second, instead of "per random bit", and show that this quantity behaves like $\Theta(1/Q)$. In the full version of this paper [9], we generalize our results to the more general setting where Alice's and Bob's antennas may have different sizes, and furthermore one does not know the exact ratio between Eve's and the honest parties' antennas – only whether it is at most, or at least, some value.

In terms of future work, we envision several main problems. First, one should extend our analysis to settings where Alice and Bob have antennas of vastly different sizes, addressing the typical client-server scenarios. Second, there is a need for a better model of the secret-key rate per second, which should be built on an abstraction level that does not abstract away time, thereby allowing one to verify our conjecture that the secret-key rate per second behaves like $1/Q$ in practice. Finally, and most importantly, one should address the issues that still prevent the satellite model from being used in practice, for instance by studying the secret-key rate in a similar setting to ours when the adversary does not quantize her analog signal, or by investigating the potential effect of an active adversary jamming the signal.

References

1. Ahlswede, R., Csiszár, I.: Common randomness in information theory and cryptography. I. Secret sharing. IEEE Transactions on Information Theory 39(4), 1121–1132 (1993)
2. Calabro, C.: The Exponential Complexity of Satisfiability Problems. Ph.D. thesis, University of California, San Diego (2009)
3. Cover, T., Thomas, J.: Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing). Wiley-Interscience (2006)
4. Csiszár, I., Körner, J.: Broadcast channels with confidential messages. IEEE Transactions on Information Theory 24(3), 339–348 (1978)
5. Csiszár, I., Narayan, P.: Secrecy capacities for multiple terminals. IEEE Transactions on Information Theory 50(12), 3047–3061 (2004)
6. Gander, M.J., Maurer, U.M.: On the secret-key rate of binary random variables. In: Proceedings of the 1994 IEEE International Symposium on Information Theory (ISIT 1994). p. 351. IEEE (1994)
7. Gohari, A.A., Anantharam, V.: Information-Theoretic Key Agreement of Multiple Terminals: Part I. IEEE Transactions on Information Theory 56(8), 3973–3996 (2010)
8. Hayashi, M., Tyagi, H., Watanabe, S.: Secret key agreement: General capacity and second-order asymptotics. IEEE Transactions on Information Theory 62(7), 3796–3810 (2016)
9. Jost, D., Maurer, U., Ribeiro, J.L.: Information-theoretic secret-key agreement: The asymptotically tight relation between the secret-key rate and the channel quality ratio. Cryptology ePrint Archive, Report 2017/1130 (2017), `https://eprint.iacr.org/2017/1130`
10. Liu, S., Van Tilborg, H.C.A., Van Dijk, M.: A Practical Protocol for Advantage Distillation and Information Reconciliation. Designs, Codes and Cryptography 30(1), 39–62 (2003), `https://doi.org/10.1023/A:1024755209150`
11. Maurer, U.M.: Secret key agreement by public discussion from common information. IEEE Transactions on Information Theory 39(3), 733–742 (1993)
12. Maurer, U.M., Wolf, S.: Unconditionally secure key agreement and the intrinsic conditional information. IEEE Transactions on Information Theory 45(2), 499–514 (1999)
13. Maurer, U.: The strong secret key rate of discrete random triples. In: Blahut, R.E., Costello, D.J., Maurer, U., Mittelholzer, T. (eds.) Communications and Cryptography. The Springer International Series in Engineering and Computer Science (Communications and Information Theory), vol. 276, pp. 271–285. Springer, Boston, MA (1994)
14. Maurer, U., Wolf, S.: Towards Characterizing When Information-Theoretic Secret Key Agreement Is Possible. In: Kim, K., Matsumoto, T. (eds.) Advances in Cryptology – ASIACRYPT 1996. Lecture Notes in Computer Science, vol. 1163, pp. 196–209. Springer, Berlin, Heidelberg (1996)
15. Maurer, U.M.: Conditionally-perfect secrecy and a provably-secure randomized cipher. Journal of Cryptology 5(1), 53–66 (1992), `https://doi.org/10.1007/BF00191321`
16. Maurer, U.M.: Protocols for Secret Key Agreement by Public Discussion Based on Common Information. In: Brickell, E.F. (ed.) Advances in Cryptology – CRYPTO 1992. Lecture Notes in Computer Science, vol. 740, pp. 461–470. Springer, Berlin, Heidelberg (1992)

17. Maurer, U.M., Wolf, S.: Information-Theoretic Key Agreement: From Weak to Strong Secrecy for Free. In: Preneel, B. (ed.) Advances in Cryptology — EUROCRYPT 2000. Lecture Notes in Computer Science, vol. 1807, pp. 351–368. Springer, Berlin, Heidelberg (2000)
18. Naito, M., Watanabe, S., Matsumoto, R., Uyematsu, T.: Secret key agreement by reliability information of signals in Gaussian Maurer's Model. In: Proceedings of the 2008 IEEE International Symposium on Information Theory (ISIT 2008). pp. 727–731. IEEE (2008)
19. Ozarow, L.H., Wyner, A.D.: Wire-Tap Channel II. In: Beth, T., Cot, N., Ingemarsson, I. (eds.) Advances in Cryptology – EUROCRYPT 1984. Lecture Notes in Computer Science, vol. 209, pp. 33–50. Springer, Berlin, Heidelberg (1984)
20. Renner, R., Skripsky, J., Wolf, S.: A new measure for conditional mutual information and its properties. In: Proceedings of the 2003 IEEE International Symposium on Information Theory (ISIT 2003). p. 259. IEEE (2003)
21. Shannon, C.: A mathematical theory of communication. Bell System Technical Journal 27(3), 379–423 (1948)
22. Shannon, C.: Communication theory of secrecy systems. Bell System Technical Journal 28(4), 656–715 (1949)
23. Stinson, D.: Universal hashing and authentication codes. Designs, Codes and Cryptography 4(3), 369–380 (1994)
24. Tyagi, H., Watanabe, S.: A bound for multiparty secret key agreement and implications for a problem of secure computing. In: Nguyen, P.Q., Oswald, E. (eds.) Advances in Cryptology – EUROCRYPT 2014. Lecture Notes in Computer Science, vol. 8441, pp. 369–386. Springer, Berlin, Heidelberg (2014)
25. Wyner, A.D.: The wire-tap channel. Bell System Technical Journal 54(8), 1355–1387 (1975)