Strong Security Against Active Attacks in Information-Theoretic Secret-Key Agreement

Stefan Wolf

Department of Computer Science
Swiss Federal Institute of Technology (ETH Zürich)
CH-8092 Zürich, Switzerland
E-mail address: wolf@inf.ethz.ch

Abstract. The problem of unconditionally secure key agreement, in particular privacy amplification, by communication over an insecure and not even authentic channel, is investigated. The previous definitions of such protocols were weak in the sense that it was only required that after the communication not both parties falsely believe that the key agreement was successful. In such a protocol however it is possible that Eve deceives one of the legitimate partners, i.e., makes him accept the outcome of the protocol although no secret key has been generated. In this paper we introduce the notion of strong protocols which protect each of the parties simultaneously and, in contrast to previous pessimism, it is shown that such protocols exist. For the important special case of privacy amplification, a strong protocol is presented that is based on a new, interactive way of message authentication with an only partially secret key. The use of feedback in such authentication allows to reduce the size of the authenticator, hence of the additional information about the key leaked to the adversary, without increasing the success probability of an active attack. Finally, it is shown that in the scenario where the parties and the adversary have access to repeated realizations of a random experiment, previously derived criteria for the possibility of secret-key agreement against active opponents hold for the new, strong definition of robustness against active attacks rather than for the earlier definition.

Keywords. Secret-key agreement, privacy amplification, authentication, unconditional secrecy, information theory.

1 Introduction

1.1 Provably Secure Key Agreement

The security of presently used cryptosystems, for instance of all public-key cryptographic protocols, is based on unproven assumptions on the hardness of certain computational problems such as the discrete logarithm problem or the integer factoring problem. The fact that all these schemes face the risk of being broken by progress in the theory of efficient algorithms motivates the search for systems whose security can be rigorously proved. In particular, protocols for the generation of a provably secure key have attracted much attention in the past few years.

In [5] for instance, a general model for secret-key agreement by public communication over an authentic channel was described. Here, two parties Alice and Bob who want to generate a secret key have access to random variables X and Y, respectively, whereas the adversary Eve knows a random variable Z. The three random variables X, Y, and Z are distributed according to some distribution P_{XYZ}.

Generally, a protocol for secret-key agreement in this scenario is often described as consisting of three phases. In the first phase, called *advantage distillation*, Alice and Bob use their advantage over Eve offered by the authenticity of the public channel, to generate an advantage over Eve in terms of their knowledge about each other's information. During the second phase, *information reconciliation*, Alice and Bob agree on a mutual string S by using error-correction techniques, and in the third phase, *privacy amplification*, the partial secret S is transformed into a shorter, highly secret string S'. Bennett *et. al.* [1] have shown that the length of S' can be nearly $H_2(S|Z = z)$, the Rényi entropy of S when given Eve's complete knowledge $Z = z$ about S.

Privacy amplification, which was first introduced by Bennett *et. al.* [2], can alternatively be seen as a *special case* of secret-key agreement from common information, namely the case where Alice and Bob have identical information, i.e., where P_{XYZ} has the property that $\text{Prob}[X = Y] = 1$. Another important special class of distributions P_{XYZ} in the secret-key agreement scenario is where X, Y, and Z consist of many independent realizations of the same random experiment [5].

1.2 Strong Security Against Active Opponents

Secret-key agreement has also been studied when dropping the condition that the channel connecting Alice and Bob is authentic [4],[6]. However, it is clear that such key agreement can only be possible if Alice and Bob already have some kind of advantage over Eve initially, and if this advantage implies that Eve cannot successfully impersonate Bob towards Alice, or vice versa. The conditions on a protocol for such key agreement have been defined as follows. After the phase of insecure communication, both Alice and Bob either accept or reject the outcome and compute a string when accepting. It was demanded that if the adversary is passive only, then both parties accept and agree on a mutual highly secure string. If the adversary is active on the other hand, then with high probability at least one of the parties must reject (or the secret-key agreement must have been successful).

Unfortunately, this definition is not completely satisfactory. Since it is only required that one of the parties rejects in case of an active attack, it is not excluded that the other party is deceived by Eve, i.e., accepts although secret-key agreement was not successful. On the other hand, it is impossible to achieve that always both Alice and Bob reject in case of an active attack. Eve can always leave Alice and Bob in opposite states by blocking certain messages, as Theorem 2 shows.

However, we propose how nearly as powerful protocols, called *strong* protocols, can be defined which are not impossible to achieve. For a strong protocol it is required that, with high probability, either both Alice and Bob reject, or the secret-key agreement is successful. It is not required that both Alice and Bob accept in the latter case, but that they both compute a mutual secure key. It seems that this is the strongest possible security one can achieve against active attackers, and that such protocols are what one actually has in mind when speaking about security against active adversaries in secret-key agreement. They have the property that no party can be misled by Eve: whenever a party accepts, the key agreement has been successful. The new protocol definition and some impossibility results are given in Section 2. In the subsequent sections we will present strong protocols in the different scenarios mentioned.

For the case of privacy amplification, treated in Section 3, strong protocols are more difficult to obtain than the weaker protocols of [6], and it is shown that strong protocols necessarily are more complicated. A new way of authenticating messages must be used which is interactive rather than one-way. The crucial point is that the authenticator of a message can be much shorter, leaking less information about the partly secret string, but maintaining security even against adversaries having partial knowledge about the key.

The scenario where the parties' (and the adversary's) information consists of repeated realizations of the same random experiment is treated in Section 4. It is shown that the criteria given in [4] for the existence (in this scenario) or inexistence (in the general scenario) of protocols secure against active opponents are not correct for the protocol definition of [4], but that these (or closely related) criteria characterize the existence of *strong* protocols in this scenario. Correcting these earlier results, we show that a (weak) protocol exists if and only if Eve can either not simulate the random variable X, using Z, in such a way that someone knowing Y cannot distinguish between X and Eve's simulation, or vice versa. In [4] it was stated that a protocol exists if *both* X and Y are not simulatable by Eve this way. By modifying the protocols of [4], we show that the last condition perfectly characterizes the existence of strong protocols.

2 Secret-Key Agreement by Communication over an Insecure and Non-Authentic Channel

2.1 Definition of Weak and Strong Protocols

Definition 1. Assume that two parties Alice and Bob both know discrete random variables X and Y, respectively, and that an adversary Eve knows a random variable Z, where the joint distribution of the random variables is P_{XYZ}. In a *protocol for secret-key agreement*, Alice and Bob exchange messages $C_1, C_2, \ldots$ over an insecure channel, where the messages $C_1, C_3, \ldots$ are sent by Alice, and the messages $C_2, C_4, \ldots$ are sent by Bob. Each message C_i depends on the sender's knowledge when sending the message and possibly on some random bits R_i, i.e., $H(C_i|X, C_1 \cdots C_{i-1}, R_i) = 0$ if i is odd and $H(C_i|Y, C_1 \cdots C_{i-1}, R_i) = 0$ if i is even[1]. At the end of the protocol, both Alice and Bob either accept or reject the outcome, and decide whether to compute a string S'_A or S'_B, respectively. If a party accepts, then it always computes a string. However, a party can also decide to compute a string when rejecting the outcome of the protocol. The above decisions and the strings S'_A and S'_B are determined by X or Y, respectively, and by the messages sent and received. The protocol is called a *one-way-transmission protocol* if messages are sent only into one direction. Otherwise, a protocol is called *interactive.*

Let r be an integer, and let $\varepsilon, \delta > 0$. A $(P_{XYZ}, r, \varepsilon, \delta)$*-protocol for secret-key agreement by communication over an insecure and non-authenticated channel* (or simply $(P_{XYZ}, r, \varepsilon, \delta)$-protocol) is a protocol for secret-key agreement with the following properties.

1. Correctness and privacy. If Eve is a passive wire-tapper, then both Alice and Bob accept at the end of the protocol, and secret-key agreement must have been successful. The latter is the event that S'_A and S'_B are r-bit strings satisfying

$$\mathrm{Prob}[S'_A \neq S'_B] \leq \varepsilon \qquad \text{and} \qquad H(S'_A|CZ) \geq r - \varepsilon \ , \tag{1}$$

where H stands for the (Shannon) entropy function, and where $C := (C_1, C_2, \ldots)$ summarizes the entire communication held over the public channel.

2. (Weak) robustness. For every possible strategy of Eve, the probability that either Alice *or* Bob rejects the outcome of the protocol or secret-key agreement has been successful, must be at least $1 - \delta$.

The protocol is called *strong* if condition *2* can be replaced by condition *2'* below. In contrast to this, a protocol satisfying *2* will also be called *weak* in the following.

2'. Strong robustness. For every possible strategy of Eve, the probability that either *both* Alice and Bob reject the outcome of the protocol or secret-key agreement has been successful, must be at least $1 - \delta$. ◦

2.2 Impossibility Results

Of course it is most desirable to use protocols for which Alice and Bob either both accept (and secret-key agreement is successful) or both reject with high probability. However, the following theorem states that such a synchronization cannot be achieved, and makes precise what was already stated in [4].

Theorem 2. *Assume that there exists a strong $(P_{XYZ}, r, \varepsilon, \delta)$-protocol with the modified robustness property that with probability at least $1 - \delta$, either both Alice*

[1] Here, the C_i stand for the messages actually sent and received by the corresponding party (thus possibly modified by the active opponent).

and Bob reject, or both parties accept and secret-key agreement has been successful. Then either suitable strings can be computed even without communication, i.e., there exist two functions f and g, mapping $\mathcal{X}$ and $\mathcal{Y}$ to $\{0,1\}^r$, respectively, such that $S'_A := f(X)$ and $S'_B := g(Y)$ satisfy (1), or $\delta = 1$.

The proof idea is that Eve can always leave Alice and Bob in opposite acceptance states by blocking the channel completely after a certain number of rounds of the protocol. A full proof will be given in the final paper.

Clearly, secret-key agreement secure against active adversaries can only be possible if Alice and Bob have some advantage over Eve in terms of the distribution P_{XYZ}. More precisely, this advantage must be such that Eve cannot generate from Z a random variable $\overline{X}$ which Bob, knowing Y, is unable to distinguish from X (and vice versa). In [4], the following property of a distribution P_{XYZ} was defined.

Definition 3. [4] Let X, Y, and Z be random variables. We say that X *is simulatable by Z with respect to Y* if there exists a conditional distribution $P_{\overline{X}|Z}$ such that $P_{\overline{X}Y} = P_{XY}$. ∘

In the final paper, we will describe a simple criterion for simulatability in terms of the probabilities $P_{XYZ}(x,y,z)$. The following theorem states that a strong $(P_{XYZ}, r, \varepsilon, \delta)$-protocol can only exist if both X and Y are not simulatable by Z with respect to each other. In the scenario in which the parties obtain repeated realizations of the same random experiment, this condition is also sufficient (see Section 4). In contrast to the result of [4], a weak protocol can already exist if Eve can either not simulate X or not simulate Y. The proof of Theorem 4 is given in the full paper.

Theorem 4. *Let X, Y, and Z be random variables with distribution P_{XYZ}. If both X and Y are simulatable by Z with respect to each other, and if $r \cdot (1-\varepsilon) - \varepsilon - h(\varepsilon) > 0$, then there exists no weak $(P_{XYZ}, r, \varepsilon, \delta)$-protocol for any $\delta < 1$. If either X is simulatable by Z with respect to Y (and $r \cdot (1-2\varepsilon) - \varepsilon - h(2\varepsilon) > 0$), or Y is simulatable by Z with respect to X (and $r \cdot (1-\varepsilon) - \varepsilon - h(\varepsilon) > 0$), then there exists no strong $(P_{XYZ}, r, \varepsilon, \delta)$-protocol for any $\delta < 1$.*

3 Privacy Amplification

3.1 Protocol Definition

Privacy amplification, introduced in [2] and generalized in [1], is the technique of transforming a partially secret string into a highly secret but shorter string, and corresponds to the special case of secret-key agreement for which $X = Y =: S$ holds with probability 1. The following definition is a strengthened version of the general definition in Section 2. First, it is required that Alice and Bob end up with the same string with probability 1 if Eve is passive. Moreover, the protocol works for an entire class of distributions P_{XYZ} instead of only one distribution. More precisely, Eve's knowledge about the mutual n-bit string S is limited by assuming that $P_{S|Z=z}$ is, for all $z \in \mathcal{Z}$, contained in some subset $\mathcal{D}$ of all possible distributions over the set $\{0,1\}^n$. Typically, $\mathcal{D}$ will consist of all distributions satisfying a certain condition in terms of the Rényi- or min-entropy. The protocol definition in [6] only covered the special case $\mathcal{D} = \mathcal{D}_{\infty,t} := \{P_X \,|\, H_\infty(X) \geq t\}$ for some t. In this paper we will deal with $\mathcal{D}$'s of the form $\mathcal{D} = \mathcal{D}_{2,t} := \{P_X \,|\, H_2(X) \geq t\}$. However, it is conceivable that protocols exist for which $\mathcal{D}$ can (or must) be defined in an entirely different way.

Definition 5. Assume that Alice and Bob both know a mutual n-bit random variable S, and that the random variable Z summarizes Eve's entire knowledge about S. Let $\mathcal{D}$ be a subset of all probability distributions on the set of n-bit strings, let r be an integer, and let $\varepsilon, \delta > 0$. A *(weak or strong)* $(n, \mathcal{D}, r, \varepsilon, \delta)$*-protocol for privacy amplification by communication over an insecure and non-authentic channel* ($(n, \mathcal{D}, r, \varepsilon, \delta)$-protocol for short) is a protocol for secret-key agreement with the following properties. Assume that $P_{S|Z=z} \in \mathcal{D}$ for all $z \in \mathcal{Z}$.

1. Correctness and privacy. If Eve is a passive wire-tapper receiving $Z = z$, then both Alice and Bob must accept at the end of the protocol, and there must exist an r-bit string S' such that $S' = S'_A = S'_B$ and $H(S'|C, Z = z) \geq r - \varepsilon$.

Finally, the same (weak or strong) *robustness* property as in Definition 1 must hold. ∘

3.2 Entropy Measures, the Effect of Side Information, and Knowledge About Partial Strings

Let us first recall the definitions of some information-theoretic quantities used in this paper.

Definition 6. Let X be a discrete random variable with probability function P_X and range $\mathcal{X}$. The *(Shannon) entropy* $H(X)$ of X is[2] $H(X) := -\mathrm{E}[\log P_X] = -\sum_{x \in \mathcal{X}} P_X(x) \log P_X(x)$. The *Rényi entropy* $H_2(X)$ is defined as $H_2(X) := -\log(\mathrm{E}[P_X]) = -\log(\sum_{x \in \mathcal{X}} P_X(x)^2) =: -\log(P_C(X))$, where $P_C(X)$ is called the *collision probability* of X. The *min-entropy* $H_\infty(X)$ is defined as $H_\infty(X) := -\log(\max_{x \in \mathcal{X}}(P_X(x)))$. ∘

Because of Jensen's inequality, $H(X) \geq H_2(X)$ holds for all X, with equality if and only if X is uniformly distributed in $\mathcal{X}$ or in a subset of $\mathcal{X}$. Furthermore, $H_2(X) \geq H_\infty(X) \geq H_2(X)/2$ holds for all X.

In the remainder of this section we provide some facts necessary for the analysis of the protocols described below. We derive bounds on the amount of knowledge (e.g., of an adversary) in terms of Rényi entropy about a partial string, depending on the amount of knowledge about the entire string. This is done both for the cases where the adversary does (Corollary 9) or does not

[2] All the logarithms in this paper are to the base 2, except ln, which is to the base e.

(Lemma 7) obtain information about the remaining part of the string. In both cases, the result is roughly the intuitive fact that [with high probability] one cannot know [substantially] more about a part than about the whole. In the case where the adversary obtains information about the remaining part of the string, the result follows from a general upper bound on the reduction of the Rényi entropy of a random variable when side information is given (Lemma 8). A statement analogous to Lemma 7 also holds with respect to min-entropy [6].

Lemma 7. *Let $S = (S_1, S_2, \ldots, S_n)$ be a random variable consisting of n binary random variables. For any k-tuple $\underline{i} = (i_1, i_2 \ldots, i_k)$, where $1 \le i_1 < i_2 < \cdots < i_k \le n$, let $S_{\underline{i}}$ be the string $(S_{i_1}, S_{i_2}, \ldots, S_{i_k})$. Then $H_2(S_{\underline{i}}) \ge H_2(S) - (n-k)$.*

Proof. Consider a fixed string $s_{\underline{i}} = (s_{i_1}, \ldots, s_{i_k})$. This particular value of the random variable $S_{\underline{i}}$ corresponds to exactly 2^{n-k} values $s = (s_1, \ldots, s_n)$ of the random variable S. Let $p_1, \ldots, p_{2^{n-k}}$ be the probabilities of these strings (in decreasing order), and let $p_0 := \sum_{i=1}^{2^{n-k}} p_i$. Now, we have

$$\sum_{i=1}^{2^{n-k}} p_i^2 = p_0 \cdot \sum_{i=1}^{2^{n-k}} \left(\frac{p_i}{p_0}\right) \cdot p_i \ge p_0 \cdot \frac{p_0}{2^{n-k}} = \frac{p_0^2}{2^{n-k}} \ .$$

Because this holds for every particular string $s_{\underline{i}}$, we have for the collision probabilities of the random variables S and $S_{\underline{i}}$

$$P_C(S_{\underline{i}}) = \sum_{s_{\underline{i}} \in \{0,1\}^k} P_{S_{\underline{i}}}(s_{\underline{i}})^2 \le 2^{n-k} \cdot \sum_{s \in \{0,1\}^n} P_S(s)^2 = 2^{n-k} \cdot P_C(S) \ .$$

Hence $H_2(S_{\underline{i}}) \ge H_2(S) - (n-k)$, and this concludes the proof. □

Lemma 8 gives an upper bound on the reduction of the Rényi entropy $H_2(P)$ of a random variable P when side-information $[Q, R]$ (consisting of a pair of random variables) is given, where $I(P;R) = 0$. It states that this reduction exceeds $\log|\mathcal{Q}|$ substantially only with small probability in both cases. (Note that it is not trivial that no additional reduction is induced by R if $I(P;R) = 0$. For instance, $I(P;Q) = 0$ and $I(P;R) = 0$ together do *not* imply that $H_2(P|Q = q, R = r) = H_2(P)$, as the example $P = Q \oplus R$ shows.) Lemma 8 can be shown similarly to Theorem 4.17 in [3].

Lemma 8. *Let P, Q, and R be discrete random variables with $I(P;R) = 0$. Then $\text{Prob}_{QR}[H_2(P|Q = q, R = r) \ge H_2(P) - \log|\mathcal{Q}| - s] \ge 1 - 2^{-(s/2-1)}$ for all $s > 2$.*

Corollary 9 is a consequence of Lemma 8. It states that a formally slightly weaker result than that of Lemma 7, concerning the knowledge (in terms of H_2) of a partial string, even holds when the rest of the string is made public.

Corollary 9. *Let S be an n-bit string, and let a partition of S be given into two strings S' and S'' of lengths l and $n-l$, respectively. Let $s > 2$ be a security parameter. Then the probability, taken over s'', that $H_2(S'|S'' = s'') \ge H_2(S) - (n-l) - s$ holds is at least $1 - 2^{-(s/2-1)}$.*

3.3 Interaction Versus One-Way Transmission

The case of *one-way-transmission* protocols for privacy amplification by public discussion over a completely insecure channel was already treated in [6]. In Appendix A of this paper, it is shown that such a protocol can never be strong, and a better analysis of the protocol in [6], called Protocol A, is given.

In Section 3.4 we present a strong, hence necessarily *interactive*, protocol for privacy amplification secure against active opponents. This protocol uses interaction for two reasons. First, feedback is necessary to prevent the sender of the first message from accepting when Eve blocks or modifies the message (Theorem 15). Second, it is advantageous to use *interactive* instead of usual one-way authentication when the adversary has some partial information about the key. Here, a message it not authenticated by the sender, but reconfirmed by the receiver by correctly answering a challenge (which is equal to the message itself). The intuitive reason is that the adversary is in a better position if she can freely choose a modified message to authenticate, instead of having to respond to a given challenge, which is necessary for attacking the interactive way of authentication described below.

Lemma 10 provides a method for interactive authentication with a partially secret key K, with the property that the adversary Eve can only answer challenges d correctly by $f_d(K)$ with substantial probability when she knows at least half of the string K (in terms of H_2). Moreover, the same is even true if Eve, given d, learns some $f_{d'}(K)$ of her choice (where $d' \neq d$). Note that this is what she can actually achieve in a substitution attack. Surprisingly, this holds although the length of d and $f_d(K)$ is only a small fraction of the length of K.

Lemma 10. *Let N and ℓ be integers such that 2ℓ divides N and $2^\ell \geq N/\ell$ holds, and let K be a random variable with range $\mathcal{K} \subseteq GF(2^N)$. Let further for any $d \in GF(2^\ell)$ the function $f_d : \{0,1\}^N \to \{0,1\}^\ell$ be defined as $f_d(x) := \sum_{i=0}^{N/\ell-1} d^i x_i$, where $(x_0, \ldots, x_{N/\ell-1}) \in (GF(2^\ell))^{N/\ell}$ is a representation of $x \in GF(2^N)$ with respect to a fixed basis of $GF(2^N)$ over $GF(2^\ell)$, the computations are carried out in the field $GF(2^\ell)$, and the elements of $GF(2^\ell)$ are represented as ℓ-bit strings with respect to a fixed basis of $GF(2^\ell)$ over $GF(2)$. Assume that there exists a (possibly probabilistic) function $d \mapsto d'$, $GF(2^\ell) \to GF(2^\ell)$, such that $d' \neq d$ holds for all d, and such that given $f_{d'}(K)$, the value $f_d(K)$ can be guessed correctly (with some strategy) with probability at least α, taken over the distribution of K, the random choice of d (according to the uniform distribution), and the coin tosses of the guessing strategy. Then $H_2(K) \leq N/2 + (2N/\ell) \cdot \log(1/\alpha)$ or equivalently, $\alpha \leq 2^{-(\ell/2N)\cdot(H_2(K)-N/2)}$.*

Proof. Note first that we can assume without loss of generality that the function $d'(d)$ and the strategy of guessing $f_d(k)$ from $f_{d'}(k)$ are deterministic, since for every possible strategy there exists a deterministic strategy that is at least as good (a randomized strategy can be seen as a combination of deterministic strategies, of which the optimal one can be chosen). Furthermore, there must exist distinct elements $d_1, \ldots, d_{N/\ell}$ of $GF(2^\ell)$ such that $f_{d_i}(k)$ is guessed correctly from $f_{d'_i}(k)$, where $d'_i := d'(d_i)$, for all $i = 1, \ldots, N/\ell$ with probability at least $\alpha^{N/\ell}$ over k. Let $\mathcal{E}$ ($\subseteq \mathcal{K}$) be this event. We prove that[3] $|\mathcal{E}| \leq \sqrt{|\mathcal{K}|}$.

By cancelling $N/2\ell$ of the pairs (d_i, d'_i) and renumbering the remaining pairs, we can obtain $N/2\ell$ pairs (d_i, d'_i) with the property that $d_i \notin \{d'_1, \ldots, d'_{i-1}\}$ holds for all $i = 1, \ldots, N/2\ell$. (In the worst case, all the pairs (d_i, d'_i) occur twice in inverse orderings. Then, every second pair (d_i, d'_i) must be cancelled.)

The event $\mathcal{E}$ has the property that $f_{d'_1}(k) = f_{d'_1}(k^*)$ implies $f_{d_1}(k) = f_{d_1}(k^*)$ for all $k, k^* \in \mathcal{E}$. Otherwise $f_{d_1}(k)$ could not be guessed correctly from $f_{d'_1}(k)$ for all $k \in \mathcal{E}$. Hence $\mathcal{E}$ must be contained in a set $\mathcal{E}_1$ of the form $\mathcal{E}_1 = \cup\{k : f_{d_1}(k) = b(a) \text{ and } f_{d'_1}(k) = a\}$ for some function $b(a)$, where the union is taken over all $a \in GF(2^\ell)$. Analogously, $\mathcal{E}$ must also be contained in sets $\mathcal{E}_i$, $i = 2, \ldots, N/2\ell$, of the same form (with d_1 and d'_1 replaced by d_i and d'_i, respectively), hence $\mathcal{E} \subseteq \cap_{i=1}^{N/2\ell} \mathcal{E}_i$. We show that the cardinality of the set on the right hand side is $\sqrt{|\mathcal{K}|}$. First, observe that every set of at most $N/\ell\,(\leq 2^\ell)$ functions f_{d_i} is, for pairwise distinct $d_i \in GF(2^\ell)$, linearly independent over $GF(2^\ell)$ (the so-called Vandermonde determinant is nonzero in this case). We define $r_l := |\cap_{i=1}^{l} \mathcal{E}_i|$. From the linear independence of $\{f_{d_1}, f_{d'_1}\}$, we first conclude that $r_1 = 2^{N-\ell}$. Furthermore, the linear independence of $f_{d_{l+1}}$ from the set $\{f_{d_1}, \ldots, f_{d_l}, f_{d'_1}, \ldots, f_{d'_{l+1}}\}$ (because $d_{l+1} \notin \{d_1, \ldots, d_l, d'_1, \ldots, d'_l\}$ according to the choice of the pairs (d_i, d'_i)) implies that $r_{l+1} = r_l/2^\ell$ for $l = 1, \ldots, N/2\ell - 1$. Note that this also holds if $d'_{l+1} = d_i$ or $d'_{l+1} = d'_i$ for some $i < l+1$. We conclude that $|\mathcal{E}| \leq r_{N/2\ell} = 2^{N-(N/2\ell)\cdot\ell} = 2^{N/2} = \sqrt{|\mathcal{K}|}$.

On the other hand, $\text{Prob}[\mathcal{E}] = \sum_{k\in\mathcal{E}} P_K(k) \geq \alpha^{N/\ell}$. In the case where P_K restricted to $\mathcal{E}$ is the uniform distribution (this case maximizes the Rényi entropy) with probability at least $\alpha^{N/\ell}/|\mathcal{E}|$, we have $\sum_{k\in\mathcal{K}} P_K(k)^2 \geq \sum_{k\in\mathcal{E}} P_K(k)^2 \geq |\mathcal{E}| \cdot (\alpha^{2N/\ell}/|\mathcal{E}|^2) \geq \alpha^{2N/\ell}/2^{N/2}$, and the claim follows when the negative logarithm is computed on both sides. □

3.4 A Strong Protocol for Privacy Amplification

The new technique for authentication allows the construction of a strong protocol for privacy amplification. However, the fact that the challenge string d, which must uniquely determine the message, i.e., the specification of the hash function for privacy amplification, is short implies that one cannot use universal hash functions, whose descriptions would be too long (see for example [8] for lower bounds on the cardinality of universal classes). We use so-called *extractors* instead, which are small classes of functions allowing to extract the min-entropy H_∞ of a weak random source into a close-to-uniformly distributed string or equivalently, to transform a partially secret into a highly secret string (see Appendix B).

We are now ready to present and analyze the strong protocol for privacy amplification secure against active adversaries. Let n be a multiple of 3, let $0 < m < 1$ be such that $2mn$ is a divisor of $2n/3$, and let $d := (2n/3)/(mn)$.

[3] Throughout the paper, the cardinality of a set M is denoted by $|M|$.

For an n-bit string S, let S_I, S_{II}, and S_{III} be $(n/3)$-bit strings such that $S = S_I||S_{II}||S_{III}$, where $||$ stands for the concatenation of strings. Let further $f_h(S_I||S_{II})$ be defined as in Lemma 10 for $S_I||S_{II} = (S_0, \ldots, S_{d-1}) \in GF(2^{mn})^d$ (here, the S_i are interpreted as elements of $GF(2^{mn})$ with respect to a fixed representation of $GF(2^{mn})$ over $GF(2)$) and $h \in GF(2^{mn})$. Then Protocol B works as follows. The extractor function E will be specified below. By $a \in_R \mathcal{A}$, we express that a is randomly chosen in the set $\mathcal{A}$ according to the uniform distribution.

Protocol B

Alice		**Bob**
$h \in_R GF(2^{mn})$	$\xrightarrow{\ h\ }$	$a := f_h(S_I\|\|S_{II})$
$S' := E(S_{III}, h)$		$b \in_R GF(2^{mn})$
	$\xleftarrow{\ (a,b)\ }$	
if $a \neq f_h(S_I\|\|S_{II})$: stop		$S' := E(S_{III}, h)$
if $a = f_h(S_I\|\|S_{II})$:		
$c := f_b(S_I\|\|S_{II})$	$\xrightarrow{\ c\ }$	if $c = f_b(S_I\|\|S_{II})$: `accept`
`accept`		if $c \neq f_b(S_I\|\|S_{II})$: `reject`

Theorem 11. *Let $t > 2/3$ be a constant. Then there exist constants m and n_0, and for every $n \geq n_0$ a function E, computable in polynomial time, such that Protocol B is a strong $(n\,, \mathcal{D}_{2,tn}\,, \Omega(n)\,, 2^{-\Omega(n)}\,, 2^{-\Omega(n)})$-protocol for privacy amplification by communication over an insecure and non-authentic channel.*

Note that the assumption on Eve's knowledge about S is exactly the same for Protocol B as for Protocol A. However, the price that has to be paid for strong robustness is that the length of the extracted string is only a constant fraction of the length of the key generated by the weak Protocol A, and that a higher round complexity is required in communication.

Proof of Theorem 11. Let $0 < m < t - 2/3$ be constant, and let $z \in \mathcal{Z}$ be the particular value known to the adversary Eve. Assume first that Eve is only *passive*. We give a lower bound on the min-entropy of the string S_{III} from Eve's point of view and given the entire communication C held over the public channel. Since this communication is, given S_I, S_{II}, and $Z = z$, independent of S_{III}, we have $H_2(S_{III}|C = c, S_I = s_I, S_{II} = s_{II}, Z = z) = H_2(S_{III}|S_I = s_I, S_{II} = s_{II}, Z = z) \geq (t - 2/3)n/2$ with probability at least $1 - 2^{-((t-2/3)n/4-1)}$ according to Lemma 7. (Of course Alice and Bob could publish S_I and S_{II} at the end of the protocol, only helping a possible adversary.) Because $H_\infty(X) \geq H_2(X)/2$ for all X, we conclude that

$$H_\infty(S_{III} \,|\, C = c, S_I = s_I, S_{II} = s_{II}, Z = z) \geq (t - 2/3)n/4 \tag{2}$$

holds with probability at least $1 - 2^{-((t-2/3)n/4-1)}$.

From Corollary 19 in Appendix B, we conclude that there exist n_0 and for all[4] $n \geq n_0$ numbers $w \leq mn$, $r = \Omega(n)$, and a function $E : \{0,1\}^{n/3} \times \{0,1\}^w \to \{0,1\}^r$ (computable in polynomial time) with the following property. Under the condition that T is an $(n/3)$-bit random variable with $H_\infty(T) \geq (t - 2/3)n/4$ and that V is a uniformly distributed w-bit random variable, we have for $R := E(T, V)$ that $H(R|V) \geq r \cdot (1 - 2^{-n/(6(\log(n/3))^3)}) \cdot (1 - 2^{-n/(6(\log(n/3))^3)} - 2^{-r})$. For the choice $P_T = P_{S_{III}|C=c,Z=z}$ and $P_R = P_{S'} = P_{E(S_{III},V)}$ (where V is composed by the first w bits of H in a fixed representation) we obtain, using (2) and $I(H; SZ) = 0$, $H(S'|C, Z = z) \geq r - r \cdot (2^{-n/(6(\log(n/3))^3)+1} + 2^{-r} + 2^{-((t-2/3)n/4-1)}) = r - 2^{-\Omega(n)}$.

We consider the case where Eve is an *active* adversary and give an upper bound on the probability of the event that Alice and Bob do not both reject although secret-key agreement has not been successful. It is obvious that this can only occur if Eve can either guess $f_h(S)$ from some $f_{h'}(S)$ (where $h' \neq h$) or guess $f_b(S)$ correctly, where h and b are randomly chosen. The success probability δ of such an active attack is upper bounded by

$$\delta \leq 2^{-(m/2)(t-2/3)n} + 2^{-((t-2/3-m)n/4-1)} + 2^{-(m/2)(t-2/3-m)n/2} = 2^{-\Omega(n)} \ . \tag{3}$$

To see this, we first conclude from Lemma 7 that $H_2((S_I||S_{II})|Z = z) \geq (t - 1/3)n$. According to Lemma 10 (for $K = S_I||S_{II}$, $N = 2n/3$, and $l = mn$) and Lemma 8, the summands in (3) are upper bounds on the probabilities of guessing $f_h(S)$ from some $f_{h'}(S)$, of the event $\mathcal{E}$ that $H_2((S_I||S_{II})|H = h, A = a, Z = z) < n/3 + (t - 2/3 - m)n/2 \leq H_2((S_I||S_{II})|Z = z) - mn - (t - 2/3 - m)n/2$, and of finding $f_b(S)$ when given $\overline{\mathcal{E}}$, respectively. We conclude that Protocol B is a strong protocol with all the required properties. □

4 Independent Repetitions of a Random Experiment

Another important special case of secret-key agreement protocols is the scenario where the information the parties obtain consists of many independent realizations of the same random experiment (with distribution P_{XYZ}) [5]. For the

[4] We can assume, not changing the basic result, that n is a multiple of 3, and that $2mn$ is an integer dividing $2n/3$. Otherwise, mn can be replaced by $k := \lceil mn \rceil$ in the entire proof, and n can be substituted by the unique multiple of $3k$ in the interval $[n, n + 3k - 1]$. Alice and Bob then add the required number of zeroes to the end of S, not changing the distribution of S.

passive-adversary case, the *secret-key rate* $S(P_{XYZ})$ has been defined in [5] as the maximal rate at which a secret key can be generated. The following definition generalizes this notion to the active-adversary case with respect to weak and strong protocols.

Definition 12. The (weak) secret-key rate against active adversaries, denoted $S_w^*(P_{XYZ})$, is the least upper bound of the set of numbers $R \geq 0$ with the property that for all $\varepsilon, \delta > 0$, and for sufficiently large n, there exists a weak $(P_{XYZ}^n, \lfloor Rn \rfloor, \varepsilon, \delta)$-protocol for secret-key agreement by communication over an insecure and non-authentic channel. Here, P_{XYZ}^n stands for the distribution over $\mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$ that corresponds to n independent realizations of the random experiment with distribution P_{XYZ}. The (strong) rate $S_s^*(P_{XYZ})$ is defined analogously, but it is required that the protocol is strong. ∘

Of course, we have $S_s^*(P_{XYZ}) \leq S_w^*(P_{XYZ}) \leq S(P_{XYZ})$ for all distributions P_{XYZ}. The following theorem expresses $S_w^*(P_{XYZ})$ and $S_s^*(P_{XYZ})$ in terms of $S(P_{XYZ})$ and P_{XYZ}, and corrects the results of [4]. Both S_w^* and S_s^* are equal to either S or 0, depending on whether X or Y (or both) are simulatable by Eve. The proof of Theorem 13 follows the lines of [4], and will be given in a final paper.

Theorem 13. *Let P_{XYZ} be a distribution of the random variables X, Y, and Z such that $S(P_{XYZ}) > 0$. Then $S_w^*(P_{XYZ}) = 0$ if and only if both X and Y are simulatable by Z with respect to each other. Otherwise, $S_w^*(P_{XYZ}) = S(P_{XYZ})$. Furthermore, $S_s^*(P_{XYZ}) = 0$ holds if and only if either X or Y is simulatable by Z (with respect to Y or X, respectively). Otherwise $S_s^*(P_{XYZ}) = S(P_{XYZ})$.*

5 Concluding Remarks

Improving earlier results, and relativizing the previous pessimism, we have shown that unconditionally secure key agreement against active opponents is possible in such a way that both parties are simultaneously protected against an adversary's active attacks. Clearly, this property is what someone would naturally request from such a protocol. In the special case of privacy amplification, interactive (instead of one-way) authentication allows to reduce the adversary's gain of information about the partially secret key by using shorter authenticators, without increasing the success probability of a message-substitution attack even by an adversary with partial knowledge about the key. Finally, we have shown that, in the situation of general random variables as well as in the scenario where the parties have access to repeated realizations of the same random experiment, previously formulated non-simulatability criteria characterize the existence of strong rather than weak protocols.

Acknowledgments

We thank Ueli Maurer and Christian Cachin for many interesting discussions. The is work was supported by the Swiss National Science Foundation (SNF).

References

1. C. H. Bennett, G. Brassard, C. Crépeau, and U. M. Maurer, Generalized privacy amplification, *IEEE Transactions on Information Theory*, Vol. 41, Nr. 6, 1995.
2. C. H. Bennett, G. Brassard, and J.-M. Robert, Privacy amplification by public discussion, *SIAM Journal on Computing*, Vol. 17, pp. 210-229, 1988.
3. C. Cachin, *Entropy measures and unconditional security in cryptography*, Ph. D. Thesis, ETH Zürich, Hartung-Gorre Verlag, Zürich, 1997.
4. U. M. Maurer, Information-theoretically secure secret-key agreement by NOT authenticated public discussion, *Advances in Cryptology - EUROCRYPT '97*, Lecture Notes in Computer Science, Vol. 1233, pp. 209-225, Springer-Verlag, 1997.
5. 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.
6. U. M. Maurer and S. Wolf, Privacy amplification secure against active adversaries, *Advances in Cryptology - CRYPTO '97*, Lecture Notes in Computer Science, Vol. 1294, pp. 307-321, Springer-Verlag, 1996.
7. N. Nisan and D. Zuckerman, Randomness is linear in space, *Journal of Computer and System Sciences*, Vol. 52, No. 1, pp. 43-52, 1996.
8. D. R. Stinson, Universal hashing and authentication codes, *Advances in Cryptology - CRYPTO '91*, Lecture Notes in Computer Science, Vol. 576, pp. 74-85, Springer-Verlag, 1992.

Appendix A. One-Way Privacy Amplification

In [1], the following important theorem on privacy amplification secure against passive adversaries has been proved, which implies that there exist protocols for privacy amplification by authenticated communication which allow to extract a string S' whose length is roughly equal to the Rényi entropy of S, given Eve's knowledge.

Theorem 14. [1] *Let S be a random variable with probability distribution P_S and Rényi entropy $H_2(S)$, and let G be the random variable corresponding to the random choice (with uniform distribution) of a member of a universal class of hash functions mapping $\mathcal{S}$ to r-bit strings, and let $S' = G(S)$. Then $r \geq H(S'|G) \geq H_2(S'|G) \geq r - 2^{r-H_2(S)}/\ln 2$.*

We will apply Theorem 14 to the case where all the probabilities are conditioned on $Z = z$. The function G is chosen from a universal class of hash functions. Generally, a class $\mathcal{H}$ of functions mapping $\mathcal{A}$ to $\mathcal{B}$ is called *universal* if for all $x, y \in \mathcal{A}$, $x \neq y$, $\text{Prob}[h(x) = h(y)] = 1/|\mathcal{B}|$ if h is chosen randomly from $\mathcal{H}$ according to the uniform distribution. An example of such a class of functions mapping l-bit strings to r-bit strings (where $l \geq r$) is the set of functions $h_c(x) = \text{LSB}_r(c \cdot x)$ for all $c \in GF(2^l)$. This class contains 2^l different functions.

Let us now consider non-interactive privacy amplification secure against active opponents. Note first that a one-way-transmission protocol cannot be strong.

Theorem 15. *Assume that a strong $(n, \mathcal{D}_{2,t}, r, \varepsilon, \delta)$-one-way-transmission protocol exists. Then $\varepsilon \geq \min\{r, n-t\}$ or $\delta = 1$.*

The proof of this theorem will be given in the final paper. The following (weak) protocol was described already in [6]. Here, S is an n-bit string, and S_I, S_{II}, and S_{III} are the first, second, and third parts of S of length $n/3$.

Protocol A

Alice		**Bob**
$h \in_R GF(2^{n/3})$		
$a := h \cdot S_I + S_{II}$	$\xrightarrow{(h,a)}$	
`accept`		`accept` if $a = h{\cdot}S_I{+}S_{II}$
$S' := \mathrm{LSB}_r(h \cdot S_{III})$		$S' := \mathrm{LSB}_r(h \cdot S_{III})$

The notation $h \in_R GF(2^{n/3})$ means that h is chosen randomly from $GF(2^{n/3})$ according to the uniform distribution. All the computations are carried out in the field $GF(2^{n/3})$.

Theorem 16. *Let n, s, and t be positive integers such that $n > tn > 2n/3 + s$. Then Protocol A is a weak $(n\,,\, \mathcal{D}_{2,t}\,,\, (t - 2/3)n - s\,,\, \varepsilon\,,\, \delta)$-protocol for privacy amplification by communication over an insecure and non-authentic channel for $\varepsilon = r \cdot 2^{-(s/3-1)} + 2^{-s/3}/\ln 2$ and $\delta = 3 \cdot 2^{-(t-2/3)n/4}$.*

Proof. Let $z \in \mathcal{Z}$ be the particular value known to Eve. We first assume that Eve is a passive wire-tapper. Let $(h, a) = (h, h \cdot S_I + S_{II})$ be the message sent from Alice to Bob, and let $\mathcal{E}$ be the event that $H_2(S_{III}|S_I = s_I, S_{II} = s_{II}, Z = z) \geq (t - 2/3)n - 2s/3$ Then $\mathcal{E}$ has, according to Corollary 9, probability at least $1 - 2^{-(s/3-1)}$. Let $r := (t - 2/3)n - s$, and let $S' := \mathrm{LSB}_r(h \cdot S_{III})$. Theorem 14 now implies that $H(S'|HA, \mathcal{E}, Z = z) \geq H(S'|HAS_IS_{II}, \mathcal{E}, Z = z) = H(S'|HS_IS_{II}, \mathcal{E}, Z = z) \geq r - 2^{-s/3}/\ln 2$. We have used $I(S_{III}; HA|S_IS_{II}, Z = z) = 0$. We conclude $H(S'|HA, Z = z) \geq \mathrm{Prob}[\mathcal{E}] \cdot (r - 2^{-s/3}/\ln 2) \geq r - r \cdot 2^{-(s/3-1)} - 2^{-s/3}/\ln 2 =: r - \varepsilon$.

Let us now consider the case where Eve is an active attacker. We give an upper bound on the probability that Eve can substitute a message (h, a) by a different message (h', a'), $h' \neq h$, without being detected. The crucial argument is that $S_I||S_{II}$ is uniquely determined by $(h, h \cdot S_I + S_{II})$ and $(h', h' \cdot S_I + S_{II})$ if $h \neq h'$. Hence the probability of a successful active attack (which can only be a substitution attack according to the definition of Protocol A, where Alice only accepts after having sent a message) is not greater than the probability of guessing S correctly when given (h, a). From Lemmas 7 and 8 we conclude that $H_2((S_I||S_{II})|H = h, A = a, Z = z) \geq (t - 2/3)n/2$ is true with probability at least $1 - 2^{-((t-2/3)n/4-1)}$. If the inequality holds, then the maximal probability of a single string $s_I||s_{II}$ is at most $2^{-H_2((S_I||S_{II})|H=h,A=a,Z=z)/2} \leq 2^{-(t-2/3)n/4}$. Hence, by the union bound, the success probability of an active attack is upper bounded by $2^{-((t-2/3)n/4-1)} + 2^{-(t-2/3)n/4} = 3 \cdot 2^{-(t-2/3)n/4} =: \delta$. $\square$

Appendix B. Extractors

In this appendix we describe the notion of an extractor and some facts needed for Protocol B. For an introduction into the subject and the precise constructions, see [7] and the references therein. Roughly spoken, an extractor allows to efficiently distill the entire (or a substantial part of) the randomness (in terms of the min-entropy) of some source into (almost) truly random bits, using a small additional number of random bits. Theorem 18 was proven in [7], introducing one particular class of extractors. Corollary 19, which is a consequence of Theorem 18, is the statement we need in the analysis of Protocol B.

Definition 17. [7] A function $E : \{0,1\}^N \times \{0,1\}^w \to \{0,1\}^r$ is called a (δ', ε')-extractor if for any distribution P on $\{0,1\}^N$ with min-entropy $H_\infty(P) \geq \delta' N$, the distance of the distribution of $[V, E(X,V)]$ to the uniform distribution over $\{0,1\}^{w+r}$ is at most ε' when choosing X according to P and V according to the uniform distribution over $\{0,1\}^w$. The distance between two distributions P and P' on a set $\mathcal{X}$ is defined as $d(P,P') := (\sum_{x\in\mathcal{X}} |P(x) - P'(x)|)/2$. $\circ$

Theorem 18. [7] *For any parameters $\delta' = \delta'(N)$ and $\varepsilon' = \varepsilon'(N)$ with $1/N \leq \delta' \leq 1/2$ and $2^{-\delta' N} \leq \varepsilon' \leq 1/N$, there exists a (δ', ε')-extractor $E : \{0,1\}^N \times \{0,1\}^w \to \{0,1\}^r$, where $w = O(\log(1/\varepsilon') \cdot (\log N)^2 \cdot (\log(1/\delta'))/\delta')$ and $r = \Omega(\delta'^2 N / \log(1/\delta'))$, and where E is computable in polynomial time.*

Corollary 19. *Let $\delta', m \in (0,1)$ be constants. Then there exists N_0 and for all $N \geq N_0$ a function E, computable in polynomial time, $E : \{0,1\}^N \times \{0,1\}^w \to \{0,1\}^r$, where $w \leq mN$ and $r = \Omega(N)$, such that if T is an N-bit random variable with $H_\infty(T) > \delta' N$, then $H(E(T,V)|V) \geq r \cdot (1 - 2^{-N/(2(\log N)^3)}) \cdot (1 - 2^{-N/(2(\log N)^3)} - 2^{-r})$ for uniformly distributed V.*

Proof. Let $\varepsilon'(N) := 2^{-N/(\log N)^3}$. Then there exists N_0 such that for all $N \geq N_0$ we have $\varepsilon' \geq 2^{-\delta' N}$, and a (δ', ε')-extractor E, mapping $\{0,1\}^{N+w}$ to $\{0,1\}^r$, where $w \leq mN$ (note that $w = O(N/\log N)$ for this choice of ε' and for constant δ') and $r = \Omega(N)$. By definition, this means that for a uniformly distributed random variable V and if $H_\infty(T) \geq \delta' N$, the distance of the distribution of $[V, E(T,V)]$ to the uniform distribution U_{w+r} over $\{0,1\}^{w+r}$ is at most $\varepsilon' = 2^{-N/(\log N)^3}$. Because $d([V, E(T,V)], U_{w+r}) = \mathrm{E}_V[d(E(T,V), U_r)] \leq \varepsilon'$ for uniformly distributed V, the distance of the distribution of $E(T,v)$ to the uniform distribution U_r (over $\{0,1\}^r$) is at most $\sqrt{\varepsilon'}$ with probability at least $1-\sqrt{\varepsilon'}$ over v, i.e., $P_V[d(E(T,V), U_r) \leq 2^{-N/(2(\log N)^3)}] \geq 1 - 2^{-N/(2(\log N)^3)}$. The corollary now follows from $H(Z) \geq k(1 - d(U_k, P_Z) - 2^{-k})$, which is true for every random variable Z with $\mathcal{Z} \subseteq \{0,1\}^k$ [6]. $\square$