PAPER

Secret Key Agreement by Soft-Decision of Signals in Gaussian Maurer's Model*

Masashi NAITO†a), *Nonmember*, Shun WATANABE††b), Ryutaroh MATSUMOTO††c), *and* Tomohiko UYEMATSU††d), *Members*

SUMMARY We consider the problem of secret key agreement in Gaussian Maurer's Model. In Gaussian Maurer's model, legitimate receivers, Alice and Bob, and a wire-tapper, Eve, receive signals randomly generated by a satellite through three independent memoryless Gaussian channels respectively. Then Alice and Bob generate a common secret key from their received signals. In this model, we propose a protocol for generating a common secret key by using the result of soft-decision of Alice and Bob's received signals. Then, we calculate a lower bound on the secret key rate in our proposed protocol. As a result of comparison with the protocol that only uses hard-decision, we found that the higher rate is obtained by using our protocol.

key words: *advantage distillation, AWGN, information theoretic security, key agreement, privacy amplification, public discussion*

1. Introduction

As one of fundamental problems in cryptography, we will consider the problem of secret key agreement in this paper. Following the studies on confidential message transmissions over noisy channels [8], [14], [15], [24], the problem of the key agreement in the information theory was formulated by Maurer [16] and independently by Ahlswede and Csisźar [1]. They considered the interactive model of secret key agreement from an initially shared partially secret string by communication over a public channel. They also defined the *secret-key rate* as the maximum of key rates, which are ratios between the length of shared secret keys and the length of the initially shared partially secret string. Then they derived a lower bound and upper bound of the *secret-key rate*.

As an example of the key agreement, Maurer [16] considered the following model. Two parties, Alice and Bob, who want to share a secret key, and the wire-tapper, Eve, receive the bits randomly generated by a satellite over independent binary symmetric channels (BSC) respectively. We call this model Maurer's model. Maurer [16] proposed an interactive key agreement protocol in his model, and showed that the key generation rate of the interactive protocol can be higher than that of non-interactive key agreement protocols. More precisely, he showed that the key rate of the interactive protocol can be positive even though the key rates of non-interactive protocols are 0.

In Maurer's original model and protocol, channels are assumed to be BSC, and received signals are assumed to be digital signals. However, signals in practical channels are analogue. Recently, the key agreement over wireless channel is experimentally studied by Aono et al. [2]. However, information theoretic analysis of the key agreement over analogue channels has not sufficiently conducted. In order to close the gap between Maurer's results and the experimental study, we will modify Maurer's model to use Gaussian channels instead of BSC, which we call Gaussian Maurer's model. Then we analyze key rates of protocols in Gaussian Maurer's model. It should be noted that Wolf [22] considered a model in which a satellite sends a random bit to Alice, Bob, and Eve over independent analogue outputs channel (not necessarily Gaussian). By reducing his model to a model in which Alice and Bob are connected to the satellite with BSCs and Eve is connected to the satellite with a binary symmetric and erasure channel, he showed that the *secret-key rate* is positive if the conditional mutual information between Alice and Bob's received analogue signals conditioned by Eve's received analogue signal is positive. However, the key rate is not calculated explicitly.

In Gaussian Maurer's model, Alice and Bob can use the results of soft-decision of analogue received signals. They can determine the reliability information from this results and use it for generating a common secret key. In this paper, we will propose a protocol for secret key agreement using the reliability information. Then, we calculate key rates at which Alice and Bob can agree a secret key in our proposed protocol. It should be noted that Maurer mentioned that Alice and Bob might be able to utilize the reliability information when they are connected to the satellite with Gaussian channels [16, Sect. 5]. However, he did not clarify the key rate of a protocol in which Alice and Bob utilize the reliability information.

Considering the situation that Alice, Bob, and Eve hard-detect the signals that are sent out by the satellite, Maurer's original model can be seen as the special case of Gaussian Maurer's model. Thus, we can compare the protocol in Gaussian Maurer's model and one in BSC Maurer's model.

Manuscript received April 28, 2008.
Manuscript revised September 2, 2008.

†The author is with the Department of Media Science Graduate School of Information Science, Nagoya University, Nagoya-shi, 464-8601 Japan.

††The authors are with the Department of Communications and Integrated Systems, Tokyo Institute of Technology, Tokyo, 152-8552 Japan.

*A part of this paper was presented at 2008 IEEE International Symposium on Information Theory in Toronto, Canada.

a) E-mail: m-naito@sp.m.is.nagoya-u.ac.jp
b) E-mail: shun-wata@it.ss.titech.ac.jp
c) E-mail: ryutaroh@rmatsumoto.org
d) E-mail: uyematsu@ieee.org

DOI: 10.1587/transfun.E92.A.525

In order to show advantage to use reliability information, we will compare the key rate in our proposed protocol and the key rate in Maurer's protocol in which Alice and Bob use only hard-decision. From the result of this comparison, we will show that the higher key rate is obtained by using our proposed protocol than the protocol that only uses hard-decision.

In order to derive the key rate at which Alice and Bob can agree a secret key in our proposed protocol (Theorem 1), we show two lemmas. Lemma 1 is used to show that Alice's key and Bob's key coincide with high probability. Lemma 2 is used to show that Eve's knowledge about the key shared by Alice and Bob is negligible. It should be noted that, during the process of the review, Nascimento et al. published the result [19] similar to Lemma 2. Their result is a generalization of the privacy amplification by Bennett et al. [3] for Eve with continuous random variables. Although Lemma 2 is also a generalization of the privacy amplification for Eve with continuous random variables, the proof method of Lemma 2 is different from [3], [19]. The proof of Lemma 2 is based on Lemma 4. Lemma 4 can be considered as a generalization of [11, Lemma 2.1.1] for conditional distributions conditioned by continuous random variables. Lemma 2.1.1 was used to show general formulae of the random number generation in the information spectrum method [11].

Rest of this paper is organized as follows. In Sect. 2, we will introduce Maurer's model modified to use Gaussian channels instead of BSC. In Sect. 3, we will show our proposed protocol using reliability information. In Sect. 4, we will compare our proposed protocol and Maurer's protocol with hard-decision. In appendices, we will prove the lemmas that are needed for the proof of theorem that derives the key rate at which Alice and Bob can agree a secret key in our proposed protocol.

2. Secret Key Rate in Gaussian Maurer's Model

Consider the following key agreement problem, which we call Gaussian Maurer's model. Assume that a satellite randomly generates signals and sends it to two parties Alice and Bob who want to share secret key and the wire-tapper Eve over three independent memoryless Gaussian channels. Their noises at time i, denoted $N_A^{(i)}$, $N_B^{(i)}$, and $N_E^{(i)}$, are drawn from independently identically distributed (i.i.d.) Gaussian distributions with mean 0 and variances V_A, V_B, and V_E respectively. A sequence of signals that the satellite generates at time 1 to n, denoted $U^n = [U^{(1)}, \ldots, U^{(n)}]$, is drawn from a distribution P_{U^n} on a signal set in $\mathbb{R}^n$ and this sequence of signals satisfies power constraint $\frac{1}{n}\sum_{i=1}^{n}(u^{(i)})^2 \leq 1$ for all sequences u^n. Alice, Bob, and Eve receive $X^n = [X^{(1)}, \ldots, X^{(n)}]$, $Y^n = [Y^{(1)}, \ldots, Y^{(n)}]$,and $Z^n = [Z^{(1)}, \ldots, Z^{(n)}]$, as outputs of these three channels at time 1 to n respectively. They are assumed to know the distribution P_{U^n} and noise variances V_A, V_B, and V_E. Note that capital letters denote random variables and corresponding small letters denote realizations in this paper.

After Alice, Bob, and Eve receive signals, Alice and Bob communicate over a public channel. This channel is assumed to be noiseless and discrete, and its capacity is finite. Every messages communicated between Alice and Bob can be intercepted by Eve, but it is assumed that Eve cannot insert fraudulent messages nor modify messages on this public channel without being detected. Let C be the entire communication held over this public channel. After enough communication over the public channel, Alice computes a secret key S on a key alphabet $\mathcal{S}$ as a function of her received signals X^n and all information C over the public channel. In a similar way, Bob computes a secret key S' on $\mathcal{S}$ as a function of Y^n and C. The secret key rate in this model is defined as follows. Note that we will take all logarithms to be base 2, and hence all the entropies will be measured in bits.

Definition 1 For given noise variances V_A, V_B, and V_E, a rate R is said to be *achievable* if for every $\epsilon > 0$ there exists a protocol for sufficiently large n satisfying

$$\Pr[S \neq S'] \leq \epsilon, \tag{1}$$

$$H(S|CZ^n) \geq \log|\mathcal{S}| - \epsilon \tag{2}$$

and

$$\frac{1}{n}\log|\mathcal{S}| \geq R - \epsilon,$$

where $|\mathcal{S}|$ denotes the number of the elements in $\mathcal{S}$.

Remark 1 In [1], [16], the security of the key is evaluated by

$$\frac{1}{n}H(S|CZ^n) \geq \frac{1}{n}\log|\mathcal{S}| - \epsilon,$$

which is called the weak security criteria, instead of Eq. (2), which is called the strong security criteria. Obviously, the strong security criteria implies the weak security criteria. For discrete random variables, Maurer and Wolf showed that if a key rate is achievable for the weak security criteria then the key rate is also achievable for the strong security criteria [17].

3. Secret Key Agreement by Soft-Decision of Signals

In this section, we will propose a protocol that uses reliability information of signals and show the key rate that is achievable by this protocol.

In our proposed protocol, the satellite selects input signal $U^{(i)}$ i.i.d. according to a distribution $P_U(1) = P_U(-1) = \frac{1}{2}$. Thus, the received signals $X^{(i)}, Y^{(i)}, Z^{(i)}$ are also i.i.d. respectively.

Let $a_1, \ldots, a_K$ be a positive and monotone increasing sequence, and let $E_1, \ldots, E_K$ be sets, where jth level set is defined as $E_j = [-a_j, a_j]$ $(j = 1, \ldots, K)$.

The procedures of our proposed protocol is as follows.

1. From the received signal $X^{(i)}$ at time i, Alice determines reliability information $W_A^{(i)}$ as

$$W_A^{(i)} = \begin{cases} 0 & \text{if } X^{(i)} \in E_1 \\ j & \text{if } X^{(i)} \in E_j^c \backslash E_{j+1}^c \ (j = 1, \ldots, K) \\ K & \text{if } X^{(i)} \in E_K^c \end{cases},$$

where the set E_j^c is the complement of the set E_j in the set of real numbers $\mathbb{R}$, and $E_j^c \backslash E_{j+1}^c = E_j^c \cap E_{j+1}$ is the difference set. Similarly, from the received signal $Y^{(i)}$ at time i, Bob determine reliability information $W_B^{(i)}$ as

$$W_B^{(i)} = \begin{cases} 0 & \text{if } Y^{(i)} \in E_1 \\ j & \text{if } Y^{(i)} \in E_j^c \backslash E_{j+1}^c \ (j = 1, \ldots, K) \\ K & \text{if } Y^{(i)} \in E_K^c \end{cases}.$$

2. Alice and Bob send sequences $W_A^n = [W_A^{(1)}, \ldots, W_A^{(n)}]$ and $W_B^n = [W_B^{(1)}, \ldots, W_B^{(n)}]$ over the public channel. From these messages, they can know the sets containing their received signals.
3. Alice and Bob quantize X^n and Y^n into discrete random variables $\tilde{X}_\Delta^n$ and $\tilde{Y}_\Delta^n$, where $\tilde{X}_\Delta^{(i)}$ is defined as

$$\tilde{X}_\Delta^{(i)} = \begin{cases} 1 & \text{if } X^{(i)} \geq 0, \\ 0 & \text{if } X^{(i)} < 0, \end{cases} \tag{3}$$

and $\tilde{Y}_\Delta^{(i)}$ is similarly defined as

$$\tilde{Y}_\Delta^{(i)} = \begin{cases} 1 & \text{if } Y^{(i)} \geq 0, \\ 0 & \text{if } Y^{(i)} < 0. \end{cases} \tag{4}$$

For given $(W_A^{(i)}, W_B^{(i)}) = (w_A, w_B)$, if Eve's ambiguity $H(\tilde{X}_\Delta|Z, W_A = w_A, W_B = w_B)$ about $\tilde{X}_\Delta^{(i)}$ is smaller than Bob's ambiguity $H(\tilde{X}_\Delta|Y, W_A = w_A, W_B = w_B)$ about $\tilde{X}_\Delta^{(i)}$, then we should discard $\tilde{X}_\Delta^{(i)}$ in our protocol. Indeed, if we keep $\tilde{X}_\Delta^{(i)}$ for such $(W_A^{(i)}, W_B^{(i)}) = (w_A, w_B)$, then a negative term is added to the lower bound on a secret key rate shown in Eq. (11). Furthermore, if the difference between Eve and Bob's ambiguity about $\tilde{X}_\Delta^{(i)}$ is smaller than the difference between Eve's ambiguity $H(\tilde{Y}_\Delta|Z, W_A = w_A, W_B = w_B)$ about $\tilde{Y}_\Delta^{(i)}$ and Alice's ambiguity $H(\tilde{Y}_\Delta|X, W_A = w_A, W_B = w_B)$ about $\tilde{Y}_\Delta^{(i)}$, we should generate a secret key from $\tilde{Y}_\Delta^{(i)}$ instead of $\tilde{X}_\Delta^{(i)}$. For this purpose, we consider the sets $A, B \subset \{1, \ldots, K\} \times \{1, \ldots, K\}$, which are defined as

$$\begin{aligned} A = \{(w_A, w_B)| \\ & H(\tilde{X}_\Delta|Z, W_A = w_A, W_B = w_B) \\ & \quad - H(\tilde{X}_\Delta|Y, W_A = w_A, W_B = w_B) \\ & \qquad \geq \max\{0, H(\tilde{Y}_\Delta|Z, W_A = w_A, W_B = w_B) \\ & \qquad\qquad - H(\tilde{Y}_\Delta|X, W_A = w_A, W_B = w_B)\}\}, \\ B = \{(w_A, w_B)| \\ & H(\tilde{Y}_\Delta|Z, W_A = w_A, W_B = w_B) \\ & \quad - H(\tilde{Y}_\Delta|X, W_A = w_A, W_B = w_B) \\ & \qquad > \max\{0, H(\tilde{X}_\Delta|Z, W_A = w_A, W_B = w_B) \\ & \qquad\qquad - H(\tilde{X}_\Delta|Y, W_A = w_A, W_B = w_B)\}\}. \end{aligned}$$

Note that Alice and Bob can calculate the conditional entropies that appear in the definitions of A and B respectively, because we assumed that they know the noise variances V_A, V_B, and V_E. If given $(W_A^{(i)}, W_B^{(i)})$ is in the set A, we use $\tilde{X}_\Delta^{(i)}$ for generating a secret key, otherwise we discard $\tilde{X}_\Delta^{(i)}$. Similarly, if given $(W_A^{(i)}, W_B^{(i)})$ is in the set B, we use $\tilde{Y}_\Delta^{(i)}$ for generating a secret key, otherwise we discard $\tilde{Y}_\Delta^{(i)}$. Thus, we determine discrete random variables

$$X_\Delta^{(i)} = \begin{cases} \tilde{X}_\Delta^{(i)} & \text{if } (W_A^{(i)}, W_B^{(i)}) \in A, \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

and

$$Y_\Delta^{(i)} = \begin{cases} \tilde{Y}_\Delta^{(i)} & \text{if } (W_A^{(i)}, W_B^{(i)}) \in B, \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

and we use them for generating a secret key instead of $\tilde{X}_\Delta^{(i)}$ and $\tilde{Y}_\Delta^{(i)}$.

4. According to the rule in Eq. (5), Alice determines X_Δ^n from W_A^n, W_B^n, and $\tilde{X}_\Delta^n$. Similarly, Bob determines Y_Δ^n from W_A^n, W_B^n, and $\tilde{Y}_\Delta^n$.
5. Alice compresses her bit sequence X_Δ^n by an encoder φ_n from $\{0, 1\}^n$ to a message set $\mathcal{M}_A$, and sends the message $M_A = \varphi_n(X_\Delta^n)$ to Bob over the public channel. Similarly, Bob compresses his bit sequence Y_Δ^n into the message M_B on $\mathcal{M}_B$, and sends it to Alice over the public channel[†].
6. Alice decodes M_B, X^n, and the reliability information (W_A, W_B) into the estimation $\hat{Y}_\Delta^n$. Similarly, Bob decodes M_A, Y^n, and the reliability information (W_A, W_B), into the estimation $\hat{X}_\Delta^n$.
7. Let $\mathcal{F}$ be a set of universal hash function [5] (see also Appendix B.1) from $\{0, 1\}^n \times \{0, 1\}^n$ to $\mathcal{S}$. Alice randomly choose a hash function $f \in \mathcal{F}$, and publicly tells the choice to Bob. Then, Alice and Bob's final keys are $S = f(X_\Delta^n, \hat{Y}_\Delta^n)$ and $S' = f(\hat{X}_\Delta^n, Y_\Delta^n)$ respectively.

In order to guarantee that Alice and Bob can compute the same key in step 6, we set the rate $\frac{1}{n}\log|\mathcal{M}_A|$ and $\frac{1}{n}\log|\mathcal{M}_B|$ of the public messages (M_A, M_B) according to the following lemma, which is derived by modifying "Slepian-Wolf Coding" [21] for continuous random variables. The lemma is proved in Appendix A.

Lemma 1 Suppose that we set

$$\frac{1}{n}\log|\mathcal{M}_A| > H(X_\Delta|YW_AW_B) \tag{7}$$

and

[†] For example, we can consider the encoder as a parity check matrix and the message as a syndrome.

$$\frac{1}{n}\log|\mathcal{M}_B| > H(Y_\Delta|XW_AW_B), \tag{8}$$

then there exist encoders and decoders such that the decoding error probabilities $\Pr\{\hat{X}^n_\Delta \neq X^n_\Delta\}$ and $\Pr\{\hat{Y}^n_\Delta \neq Y^n_\Delta\}$ tend to 0 as $n \to \infty$.

Thus, Eq. (1) is satisfied for sufficiently large n.

In order to guarantee the security of the protocol, we set the key rate $\frac{1}{n}\log|S|$ according to the following lemma, which is derived by modifying the so-called "left over hash lemma" [3], [4], [12] for continuous random variables. The lemma is proved in Appendix B.

Lemma 2 Suppose that we set

$$\frac{1}{n}\log|S| < H(X_\Delta Y_\Delta|ZW_AW_B) - \frac{1}{n}\log|\mathcal{M}_A||\mathcal{M}_B|, \tag{9}$$

then

$$H(S|Z^nW_A^nW_B^nM_AM_BF) \geq \log|S| - \epsilon \tag{10}$$

is satisfied for sufficiently large n.

Note that F is a random variable on $\mathcal{F}$, and all information C over the public channel correspond to $(W_A^n, W_B^n, M_A, M_B, F)$ in this case.

From Eqs. (7)–(9), we obtain the following theorem that gives the key rate that is achievable by this protocol.

Theorem 1 By using our proposed protocol, the key rate

$$H(X_\Delta Y_\Delta|ZW_AW_B) - H(X_\Delta|YW_AW_B) \\ -H(Y_\Delta|XW_AW_B). \tag{11}$$

is achievable.

Note that from the rule in Eqs. (5), (6), we can rewrite the Eq. (11) as

$$\begin{aligned} &H(X_\Delta Y_\Delta|ZW_AW_B) - H(X_\Delta|YW_AW_B) \\ &\qquad -H(Y_\Delta|XW_AW_B) \\ &= \sum_{w_A,w_B} P_{W_AW_B}(w_A,w_B) \\ &\quad \times \max\{0, H(\tilde{X}_\Delta|Z, W_A = w_A, W_B = w_B) \\ &\qquad - H(\tilde{X}_\Delta|Y, W_A = w_A, W_B = w_B), \\ &\qquad H(\tilde{Y}_\Delta|Z, W_A = w_A, W_B = w_B) \\ &\qquad - H(\tilde{Y}_\Delta|X, W_A = w_A, W_B = w_B)\}. \end{aligned}$$

For fixed $(W_A, W_B) = (w_A, w_B)$, $H(\tilde{X}_\Delta|Z, W_A = w_A, W_B = w_B) - H(\tilde{X}_\Delta|Y, W_A = w_A, W_B = w_B)$ is the key rate that is achievable when we use only $\tilde{X}^n_\Delta$ for generating a secret key, $H(\tilde{Y}_\Delta|Z, W_A = w_A, W_B = w_B) - H(\tilde{Y}_\Delta|X, W_A = w_A, W_B = w_B)$ is the key rate that is achievable when we use only $\tilde{Y}^n_\Delta$ for generating a secret key, and 0 is trivially achievable key rate. By the rule in Eqs. (5), (6), we choose the maximum among these key rates for each (w_A, w_B) in order to make the achievable key rate as high as possible.

Note that encoding in step 5 and decoding in step 6 are implementable by using low-density parity check codes [6], [18].

4. Comparison to a Protocol with Hard-Decision

In this section, we numerically show the key rate that is achievable by our proposed protocol, i.e., Eq. (11). First, we show the relation between signal-to-noise ratio (SNR) and the key rate for several noise-to-noise ratio (NNR). In order to clarify the advantage to use the soft-decision, we also show the comparisons between the key rate that is achievable by our proposed protocol and the key rate that is achievable by Maurer's protocol in which Alice and Bob use only hard-decision for generating a secret key.

The relation between (SNR) and the key rate for several NNR is presented in Fig. 1, where sets E_1, E_2, and E_3 are determined from fixed $a_1 = \frac{1}{3}, a_2 = \frac{2}{3}, a_3 = 1$ in our proposed protocol. Note that SNR is defined as $\frac{1}{V_A}$ and NNR is defined as $\frac{V_E}{V_B}$, and we assume $V_A = V_B$. From this figure, we observe that we do not obtain a high key rate when SNR is too high or too low.

In order to show advantage to use soft-decision, we compare the key rate that is achievable by our proposed protocol and the key rate that is achievable by Maurer's protocol in which Alice and Bob use only hard-decision for generating a secret key. Maurer's protocol consists of an interactive phase, which is the so-called advantage distillation[†], and a non-interactive key agreement phase. The key rate that is achievable by Maurer's protocol depends on the number of the iteration of the interactive phase. For each NNR, we determined the number of the iteration optimally among 0 to 4. In our protocol, we determined the set from fixed $a_1 = \frac{1}{3}, a_2 = \frac{2}{3}, a_3 = 1$. The result of this comparison is presented in Figs. 2(a)–(c). From these figures, we observe that we obtain a larger key rate by our proposed protocol

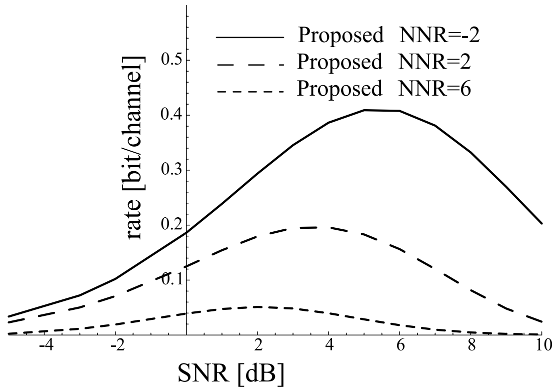
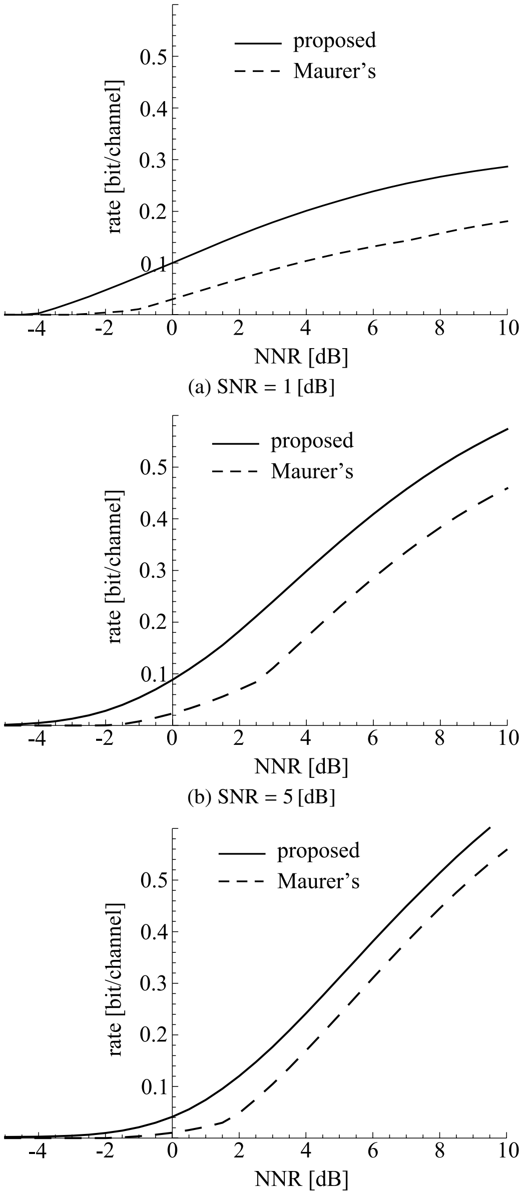


Fig. 1 The relation between SNR and the key rate in our proposed protocol for several NNR.

[†] There is two kinds of advantage distillation, the so-called repeat-code protocol [16] and the parity-check protocol [10] (see also [23]). The key rate of the parity-check protocol is slightly higher than that of the repeat-code protocol. In the comparison, we adopted the parity-check protocol. However, the difference of the key rates for the repeat-code protocol and the parity-check protocol is extremely small.



(a) SNR = 1 [dB]

(b) SNR = 5 [dB]

(c) SNR = 7 [dB]

Fig. 2 The key rates that are achievable by our proposed protocol and Maurer's protocol.

than by Maurer's protocol with all value of NNR. Note that in Gaussian Maurer's model, we should calculate the key rate by Maurer's protocol for Eve who can use continuous random variables Z^n to guess the secret key. However, the numerical calculation of the key rate by Maurer's protocol in Gaussian Maurer's model is difficult when the number of the iteration of the interactive phase is larger than 1. Thus, we calculate the key rate in BSC Maurer's model instead of Gaussian Maurer's model when the number of the iteration of the interactive phase is larger than 1. In the calculation of the key rate in BSC Maurer's model, we consider the situation that Alice, Bob, and Eve hard-detect received signals according to the similar rule as in Eqs. (3) and (4). In this situation, we can convert three Gaussian channels into independent binary symmetric channels with error probabilities $\epsilon_A, \epsilon_B, \epsilon_E$ given by

$$\epsilon_A = \frac{1}{2}erfc\left(\sqrt{\frac{1}{V_A}}\right), \qquad \epsilon_B = \frac{1}{2}erfc\left(\sqrt{\frac{1}{V_B}}\right),$$
$$\epsilon_E = \frac{1}{2}erfc\left(\sqrt{\frac{1}{V_E}}\right), \tag{12}$$

where the complementary error function $erfc(z)$ is defined as

$$erfc(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-t^2}. \tag{13}$$

Note that this way of the comparison gives Maurer's protocol advantage because a wire-tapper in Gaussian Maurer's model is more powerful than in BSC Maurer's model[†]. Hence, the key rate that is guaranteed to be achievable by Maurer's protocol in Gaussian Maurer's model is lower than that presented in Figs. 2(a)–(c).

5. Conclusion

In this paper, we have proposed Gaussian Maurer's model and the protocol with reliability information based on the result of the soft-decision in this model. As a result, we have obtained a higher key rate than Maurer's protocol. This is because that the correlation between X_Δ in Eq. (5) and Y and between Y_Δ in Eq. (6) and X obtained by using the reliability information is stronger than the correlation between $\tilde{X}_\Delta$ in Eq. (3) and $\tilde{Y}_\Delta$ in Eq. (4) obtained by using the hard-decision.

However, we do not know the optimal way to determine sets $E_1, \ldots, E_K$ and its number K. Intuitively, one may think that the more sets we use, the higher rate we obtain. However, this intuition does not seem to be always true. We determined the number of the sets to be 3 according to our rule of thumb. Furthermore, we have to find the optimal signal constellation used by the satellite. These problems are future research agenda.

Acknowledgments

We would like to thank Prof. Motohiko Isaka and Dr. Jun Muramatsu for valuable discussions. We also thank the anonymous reviewers for their constructive comments and suggestions. This research is also partly supported by the Japan Society for the Promotion of Science under Grants-in-Aid No. 00197137.

References

[1] R. Ahlswede and I. Csiszar, "Common randomness in information theory and cryptography — part 1: Secret sharing," IEEE Trans. Inf. Theory, vol.39, no.4, pp.1121–1132, 1993.

[†] The wire-tapper in Gaussian Maurer's model can use continuous random variables Z^n to guess the secret key, but one in BSC Maurer's model can only use quantized versions of them.

[2] T. Aono, K. Higuchi, T. Ohira, B. Komiyama, and H. Sasaoka, “Wireless secret key generation exploiting reactance-domain scalar response of multipat fading channel,” IEEE Trans. Antennas Propag., vol.53, no.11, pp.3776–3784, Nov. 2005.

[3] C.H. Bennett, G. Brassard, C. Crépeau, and U. Maurer, “Generalized privacy amplification,” IEEE Trans. Inf. Theory, vol.41, no.6, pp.1915–1923, Nov. 1995.

[4] C.H. Bennett, G. Brassard, and J.M. Robert, “Privacy amplification by public discussion,” SIAM J. Comput., vol.17, no.2, pp.210–229, April 1988.

[5] J.L. Carter and M.N. Wegman, “Universal classes of hash functions,” J. Comput. Syst. Sci., vol.18, pp.143–154, 1979.

[6] T.P. Coleman, A.H. Lee, M. Médard, and M. Effros, “Low-complexity approaches to Slepian & Wolf near-lossless distributed data compression,” IEEE Trans. Inf. Theory, vol.52, no.8, pp.3546–3561, Aug. 2006.

[7] T.M. Cover and J.A. Thomas, Elements of Information Theory, 2nd ed., John Wiley & Sons, 2006.

[8] I. Csiszár and J. Körner, “Broadcast channels with confidential messages,” IEEE Trans. Inf. Theory, vol.24, no.3, pp.339–348, May 1979.

[9] I. Csiszár and P. Narayan, “Secrecy capacities for multiple terminals,” IEEE Trans. Inf. Theory, vol.50, no.12, pp.3047–3061, Dec. 2004.

[10] M.J. Gander and U. Maurer, “On the secret-key rate of binary random variables,” IEEE International Symposium on Information Theory, p.351, 1994.

[11] T.S. Han, Information-Spectrum Methods in Information Theory, Springer, 2003.

[12] R. Impagliazzo, L.A. Levin, and M. Luby, “Pseudo-random generation from one-way function,” Proc. 21st Annual ACM Symposium on Theory of Computing (STO’89), pp.12–24, ACM Press, 1989.

[13] K. Itou, Introduction to probability theory, Cambridge University Press, 1984.

[14] S.K. Leung-Yan-Cheong, Multi-User and Wire-tap Channels Including Feedback, Ph.D. Thesis, Stanford University, 1976.

[15] S.K. Leung-Yan-Cheong and M.E. Hellman, “The Gaussian wire-tap channel,” IEEE Trans. Inf. Theory, vol.24, no.4, pp.451–456, July 1978.

[16] U. Maurer, “Secret key agreement by public discussion from common information,” IEEE Trans. Inf. Theory, vol.39, no.3, pp.733–742, May 1993.

[17] U. Maurer and S. Wolf, “Infromation-theoretic key agreement: From weak to strong secrecy for free,” Advances in Cryptology — EUROCRYPT 2000, Lecture Notes in Computer Science, vol.1807, pp.351–368, Springer-Verlag, 2000.

[18] J. Muramatsu, T. Uyematsu, and T. Wadayama, “Low-density parity-check matrices for coding of correlated sources,” IEEE Trans. Inf. Theory, vol.51, no.10, pp.3645–3654, Oct. 2005.

[19] A.C.A. Nascimento, J. Barros, S. Skludarek, and H. Imai, “The commitment capacity of the gaussian channel is infinite,” IEEE Trans. Inf. Theory, vol.54, no.6, pp.2785–2789, July 2008.

[20] R. Renner, Security of quantum key distribution, Ph.D. Thesis, Dipl. Phys. ETH, Switzerland, 2005. arXiv:quant-ph/0512258.

[21] D. Slepian and J.K. Wolf, “Noiseless coding of correlated information sources,” IEEE Trans. Inf. Theory, vol.19, no.4, pp.471–480, July 1973.

[22] S. Wolf, Information-Theoretically and Computationally Secure Key Agreement in Cryptography, Ph.D Thesis, Dipl. Math. ETH, Switzerland, 1999.

[23] S. Wolf, “Unconditional security in cryptography,” Lectures on Data Security, Lecture Notes in Computer Science, vol.1561, pp.217–250, Springer, 1999.

[24] A.D. Wyner, “The wire-tap channel,” Bell Syst. Tech. J., vol.54, no.8, pp.1355–1387, 1975.

Appendix A: Proof of Lemma 1

We only prove that if we set the rate $\frac{1}{n}\log|\mathcal{M}_A|$ of public message according to Eq. (7), then there exist encoders and decoders such that the decoding error probabilities $\Pr\{\hat{X}^n_\Delta \neq X^n_\Delta\}$ tends to 0 as $n \to \infty$. The proof for the rate $\frac{1}{n}\log|\mathcal{M}_B|$ of public message follows by symmetry.

We use the so-called “bin coding” proposed by Cover [7] in this proof. The procedures of bin coding is as follows.

Assign every $x^n_\Delta \in \mathcal{X}^n_\Delta$ to one of $|\mathcal{M}_A|$ bins independently according to the uniform distribution on $\mathcal{M}_A$.

Alice sends the index i of the bin to which x^n_Δ belongs. Then let $\bar{\varphi}_n(x^n_\Delta) = i$.

For each $(y^n, \mathbf{w}^n)$, we define the set $S_n(y^n, \mathbf{w}^n) \subset \mathcal{X}^n_\Delta$ as

$$\begin{aligned}S_n(y^n, \mathbf{w}^n) := \Big\{ x^n_\Delta : \frac{1}{n}\log\frac{1}{P_{X^n_\Delta|Y^n,\mathbf{W}^n}(x^n_\Delta|y^n, \mathbf{w}^n)} \\ \leq H(X_\Delta|Y\mathbf{W}) + \gamma \Big\},\end{aligned}$$

where $\gamma > 0$ is an arbitrary fixed small constant, and we denote the pair (W^n_A, W^n_B) as $\mathbf{W}^n$. Then, for given y^n, $\mathbf{w}^n$, and the received index i, declare $\bar{\psi}_n(i, y^n, \mathbf{w}^n) = x^n_\Delta$ if there is one and only one pair $(x^n_\Delta, y^n, \mathbf{w}^n)$ such that $\bar{\varphi}_n(x^n_\Delta) = i$ and $x^n_\Delta \in S_n(y^n, \mathbf{w}^n)$. Otherwise, declare an error.

We will evaluate the decoding error probability averaged over randomly chosen encoders as follows. We have an error if X^n_Δ is not in $S_n(Y^n, \mathbf{W}^n)$ or if there is another symbol $\hat{x}^n_\Delta \in \mathcal{X}^n_\Delta$ in the same bin. Thus, we can define the events of error

$$\begin{aligned}E^{(0)}_n &:= \{X^n_\Delta \notin S_n(Y^n, \mathbf{W}^n)\},\\ E^{(1)}_n &:= \{\exists \hat{x}^n \neq X^n_\Delta : \bar{\varphi}_n(\hat{x}^n_\Delta) = \bar{\varphi}_n(X^n_\Delta) \\ &\qquad \text{and } \hat{x}^n_\Delta \in S_n(Y^n, \mathbf{W}^n)\},\end{aligned}$$

Then the decoding error probability averaged over randomly chosen encoders $\Pr\{X^n_\Delta \neq \bar{\psi}_n(\bar{\varphi}_n(X^n_\Delta), Y^n, \mathbf{W}^n)\}$ is upper bounded as

$$\begin{aligned}&\Pr\{X^n_\Delta \neq \bar{\psi}_n(\bar{\varphi}_n(X^n_\Delta), Y^n, \mathbf{W}^n)\}\\ &\quad = \Pr\{E^{(0)}_n \cup E^{(1)}_n\}\\ &\quad \leq \Pr\{E^{(0)}_n\} + \Pr\{E^{(1)}_n\}. \qquad \text{(A$\cdot$1)}\end{aligned}$$

$\Pr\{E^{(0)}_n\}$ is evaluated as

$$\begin{aligned}\Pr\{E^{(0)}_n\} &= \Pr\{X^n_\Delta \notin S_n(Y^n, \mathbf{W}^n)\}\\ &= \Pr\Big\{\frac{1}{n}\log\frac{1}{P_{X^n_\Delta|Y^n\mathbf{W}^n}(X^n_\Delta|Y^n\mathbf{W}^n)}\\ &\qquad > H(X_\Delta|Y\mathbf{W}) + \gamma\Big\}\\ &= \Pr\Big\{\frac{1}{n}\sum_{i=1}^{n}\log\frac{1}{P_{X_\Delta|Y\mathbf{W}}(X^{(i)}_\Delta|Y^{(i)}\mathbf{W}^{(i)})}\\ &\qquad > H(X_\Delta|Y\mathbf{W}) + \gamma\Big\}, \qquad \text{(A$\cdot$2)}\end{aligned}$$

which tends to 0 as $n \to \infty$ by the weak law of large numbers. To bound $\Pr\{E^{(1)}_n\}$, we rewrite it as

$$\begin{aligned}&\Pr\{E_n^{(1)}\}\\&= \Pr\{\exists \hat{x}_\Delta^n \neq X_\Delta^n : \bar{\varphi}_n(\hat{x}_\Delta^n) = \bar{\varphi}_n(X_\Delta^n)\\&\qquad \text{and } \hat{x}_\Delta^n \in S_n(Y^n, \mathbf{W}^n)\}\\&= \int_{\mathcal{Y}^n} p_{Y^n}(y^n) \sum_{(x_\Delta^n, \mathbf{w}^n) \in \mathcal{X}_\Delta^n \times \mathcal{W}_A^n \times \mathcal{W}_B^n}\\&\quad P_{X_\Delta^n \mathbf{W}^n | y^n}(x_\Delta^n, \mathbf{w}^n) g_n(x_\Delta^n, y^n, \mathbf{w}^n)\, dy^n, \end{aligned} \tag{A· 3}$$

where

$$\begin{aligned}&g_n(x_\Delta^n, y^n, \mathbf{w}^n)\\&= \Pr\{\exists \hat{x}_\Delta^n \neq x_\Delta^n : \bar{\varphi}_n(\hat{x}_\Delta^n) = \bar{\varphi}_n(x_\Delta^n)\\&\qquad \text{and } (\hat{x}_\Delta^n) \in S_n(y^n, \mathbf{w}^n)\}. \end{aligned} \tag{A· 4}$$

Furthermore, we can rewrite (A· 4) as

$$\begin{aligned} g_n(x_\Delta^n, y^n, \mathbf{w}^n) &= \sum_{\substack{\hat{x}_\Delta^n \neq x_\Delta^n \\ \hat{x}_\Delta^n \in S_n(y^n, \mathbf{w}^n)}} \Pr\{\bar{\varphi}_n(\hat{x}_\Delta^n) = \bar{\varphi}_n(x_\Delta^n)\}\\ &= \sum_{\substack{\hat{x}_\Delta^n \neq x_\Delta^n \\ \hat{x}_\Delta^n \in S_n(y^n, \mathbf{w}^n)}} \frac{1}{|\mathcal{M}_A|}\\ &\leq \sum_{\hat{x}_\Delta^n \in S_n(y^n, \mathbf{w}^n)} \frac{1}{|\mathcal{M}_A|}\\ &= \frac{|S_n(y^n, \mathbf{w}^n)|}{|\mathcal{M}_A|} \end{aligned} \tag{A· 5}$$

If $\hat{x}_\Delta^n \in S_n(y^n, \mathbf{w}^n)$, then from the definition of $S_n(y^n, \mathbf{w}^n)$, we have

$$P_{X_\Delta^n | y^n, \mathbf{w}^n}(\hat{x}_\Delta^n) \geq 2^{-n(H(X_\Delta | Y \mathbf{W}) + \gamma)}.$$

Thus, we have

$$\begin{aligned} 1 &\geq \sum_{\hat{x}_\Delta^n \in S_n(y^n, \mathbf{w}^n)} P_{X_\Delta^n | Y^n \mathbf{W}^n}(x_\Delta^n | y^n, \mathbf{w}^n)\\ &\geq |S_n(y^n, \mathbf{w}^n)| 2^{-n(H(X_\Delta | Y \mathbf{W}) + \gamma)}. \end{aligned}$$

Hence, we have

$$|S_n(y^n, \mathbf{w}^n)| \leq 2^{n(H(X_\Delta | Y \mathbf{W}) + \gamma)}. \tag{A· 6}$$

From Eqs. (A· 3)–(A· 6), we upper bound $\Pr\{E_n^{(1)}\}$ as

$$\begin{aligned} \Pr\{E_n^{(1)}\} &\leq \int_{\mathcal{Y}^n} p_{Y^n}(y^n) \sum_{(x_\Delta^n, \mathbf{w}^n) \in \mathcal{X}_\Delta^n \times \mathcal{W}_A^n \times \mathcal{W}_B^n}\\ &\qquad P_{X_\Delta^n \mathbf{W}^n | y^n}(x_\Delta^n, \mathbf{w}^n) \frac{2^{n(H(X_\Delta | Y \mathbf{W}) + \gamma)}}{|\mathcal{M}_A|}\, dy^n\\ &\leq \frac{2^{n(H(X_\Delta | Y \mathbf{W}) + \gamma)}}{|\mathcal{M}_A|}\\ &= 2^{-\log |\mathcal{M}_A|} 2^{n(H(X_\Delta | Y \mathbf{W}) + \gamma)}, \end{aligned} \tag{A· 7}$$

which exponentially tends to 0 as $n \to \infty$ if $\frac{1}{n} \log |\mathcal{M}_A| > H(X_\Delta | Y \mathbf{W}) + \gamma$.

Since the decoding error probability $\Pr\{X_\Delta^n \neq \bar{\psi}_n(\bar{\varphi}_n(X_\Delta^n), Y^n, \mathbf{W}^n)\}$ of randomly chosen code tends to 0 as $n \to \infty$, there exist at least one pair of an encoder and a decoder such that the decoding error probability $\Pr\{\hat{X}_\Delta^n \neq X_\Delta^n\}$ tends to 0 as $n \to \infty$.

Appendix B: Proof of Lemma 2

In this Appendix, we will show the proof of lemma 2. In Sect. B.1, we introduce a universal hash family, which is used for computation of a secret key. In Sect. B.2, we define the security of the protocol in the sense of the variational distance, and we show the relation between the security of the protocol in the sense of the variational distance and the condition Eq. (10). This relation implies that if the security of the protocol in the sense of the variational distance is satisfied, then the condition Eq. (10) is satisfied. In Sect. B.3, we relate the size $|\mathcal{S}|$ of a secret key S and the size $|\mathcal{M}_A \times \mathcal{M}_B|$ of public messages $\mathbf{M} = (M_A, M_B)$ to the security of the protocol, and we show that if we set $\frac{1}{n} \ln |\mathcal{S}| < H(X_\Delta Y_\Delta | Z W_A W_B) - \frac{1}{n} \ln |\mathcal{M}_A \times \mathcal{M}_B|$, then there exists at least one hash function f that satisfy Eq. (10) for sufficiently large n.

For the simplicity of notation, we denotes the integral over $\mathbb{R}^n$ as $\int$ unless otherwise specified, and we abbreviates $P_{\mathbf{R}^n \mathbf{M}^n | Z^n \mathbf{W}^n}(\cdot, \cdot | z^n, \mathbf{w}^n)$ as $P_{\mathbf{R}^n \mathbf{M}^n | z^n, \mathbf{w}^n}(\cdot, \cdot)$. The variational distance $\|P_1 - P_2\|$ between the probability distribution P_1 and P_2 on $\mathcal{V}$ is defined as

$$\|P_1 - P_2\| := \sum_{v \in \mathcal{V}} |P_1(v) - P_2(v)|. \tag{A· 8}$$

B.1 Universal Hash Family

In order to extract an almost secret string (secret key S) from a partially secret strings (a pair $\mathbf{R}^n$ of random variables X_Δ^n and Y_Δ^n), we use a universal hash family $\mathcal{F}$. A set $\mathcal{F}$ of functions $f : \mathcal{X}_\Delta^n \times \mathcal{Y}_\Delta^n \to \mathcal{S}$ is said to be a *universal hash family* if we have

$$P_F\left(\{f \in \mathcal{F} \mid f(\mathbf{r}^n) = f(\mathbf{r}'^n)\}\right) \leq \frac{1}{|\mathcal{S}|} \tag{A· 9}$$

for any $\mathbf{r}^n \neq \mathbf{r}'^n \in \mathcal{X}_\Delta^n \times \mathcal{Y}_\Delta^n$, where F denotes a random variable on $\mathcal{F}$ and P_F denotes the uniform distribution on $\mathcal{F}$. For given Eve's received signals $z^n \in \mathbb{R}^n$ and reliability information $\mathbf{w}^n \in \mathcal{W}_A \times \mathcal{W}_B$, the jointly conditional distribution $P_{S \mathbf{M} | z^n, \mathbf{w}^n}(s, \mathbf{m})$ of a secret key $S = f(\mathbf{R}^n)$ and public message $\mathbf{M}$ is given by

$$\begin{aligned} P_{S \mathbf{M} | z^n, \mathbf{w}^n}(s, \mathbf{m}) &:= \sum_{\mathbf{r}^n \in f^{-1}(s)} P_{\mathbf{R}^n \mathbf{M} | z^n, \mathbf{w}^n}(\mathbf{r}^n, \mathbf{m})\\ &= P_{\mathbf{R}^n \mathbf{M} | z^n, \mathbf{w}^n}(f^{-1}(s), \mathbf{m}), \end{aligned}$$

where $f^{-1}(s) := \{\mathbf{r}^n \in \mathcal{X}_\Delta^n \times \mathcal{Y}_\Delta^n \mid f(\mathbf{r}^n) = s\}$ is the subset of a set $\mathcal{X}_\Delta^n \times \mathcal{Y}_\Delta^n$ such that $f(\mathbf{r}^n) = s$. Note that since S depends on a hash function f, it should be referred as S_f. But, we use the above notation for convenience in this paper.

B.2 The Security of the Protocol in the Sense of the Variational Distance

In order to prove lemma 2, we define the security of the protocol in the sense of the variational distance in this section.

If a secret key S is independent of Eve's information and its distribution P_S is close to the uniform distribution $P_{\bar{S}}$ on $\mathcal{S}$, we decide that the secret key S is secure in the sense of the variational distance. In the other words, we define the security of the protocol as

$$\Delta_f := \int p_{Z^n}(z^n) \sum_{\mathbf{w}^n \in \mathcal{W}_A^n \times \mathcal{W}_B^n} P_{\mathbf{W}^n|z^n}(\mathbf{w}^n) \\ \|P_{S\mathbf{M}|z^n,\mathbf{w}^n} - P_{\bar{S}} \times P_{\mathbf{M}|z^n,\mathbf{w}^n}\| dz^n, \tag{A$\cdot$ 10}$$

where $P_{\mathbf{M}|z^n,\mathbf{w}^n}$ is the marginal distribution of $P_{S\mathbf{M}|z^n,\mathbf{w}^n}$, and $P_{\bar{S}} \times P_{\mathbf{M}|z^n,\mathbf{w}^n}$ is the product distribution of $P_{\bar{S}}$ and $P_{\mathbf{M}|z^n,\mathbf{w}^n}$

As an extension of [9, Lemma 1] to continuous random variable, the following lemma relates the security of the protocol in the sense of the variational distance to the security of the protocol in the sense of the entropy shown in Eq. (2).

Lemma 3 The conditional entropy $H(S|Z^n\mathbf{W}^n\mathbf{M}F)$ is lower bounded by

$$H(S|Z^n\mathbf{W}^n\mathbf{M}F) \geq (1 - \mathbb{E}_f[\Delta_f]) \ln |\mathcal{S}| \\ -\mathbb{E}_f[\Delta_f] \log \frac{1}{\mathbb{E}_f[\Delta_f]}. \tag{A$\cdot$ 11}$$

Note that since $\mathbf{W}^n = (W_A^n, W_B^n)$ and $\mathbf{M} = (M_A, M_B)$, the conditional entropy $H(S|Z^n\mathrm{W}^n\mathbf{M}F)$ is equivalent to $H(S|Z^nW_A^nW_B^nM_A, M_BF)$ in Eq. (10). From this lemma, if $\mathbb{E}_f[\Delta_f]$ is sufficiently small, a secret key S is secure in the sense of the entropy.

Proof. Let

$$\Delta_{f,\mathbf{m},z^n,\mathbf{w}^n} := \|P_{S|\mathbf{m},z^n,\mathbf{w}^n} - P_{\bar{S}}\|. \tag{A$\cdot$ 12}$$

Then, we can rewrite Δ_f as

$$\Delta_f = \int p_{Z^n}(z^n) \sum_{\mathbf{m},\mathbf{w}^n} \\ P_{\mathbf{M}\mathbf{W}^n|z^n}(\mathbf{m}, \mathbf{w}^n)\Delta_{f,\mathbf{m},z^n,\mathbf{w}^n}\, dz^n \tag{A$\cdot$ 13}$$

For given $z^n \in \mathbb{R}^n$, $\mathbf{w}^n \in \mathcal{W}_A^n \times \mathcal{W}_B^n$, and $\mathbf{m} \in \mathcal{M}_A \times \mathcal{M}_B$, we obtain

$$H(S|\mathbf{M} = \mathbf{m}, Z^n = z^n, \mathbf{W}^n = \mathbf{w}^n, F = f) \\ \geq \log |\mathcal{S}| - \Delta_{f,\mathbf{m},z^n,\mathbf{w}^n} \log \frac{|\mathcal{S}|}{\Delta_{f,\mathbf{m},z^n,\mathbf{w}^n}}, \tag{A$\cdot$ 14}$$

which follows from the continuity of entropy [7] in the similar way as [9, Lemma 1].

The second term of Eq. (A$\cdot$ 14) is upper bonded as follow. Since $t \log \frac{1}{t}$ is a concave function, we obtain

$$\sum_{\mathbf{m},\mathbf{w}^n} P_{\mathbf{M}\mathbf{W}^n|z^n}(\mathbf{m}, \mathbf{w}^n)\Delta_{f,\mathbf{m},z^n,\mathbf{w}^n} \log \frac{|\mathcal{S}|}{\Delta_{f,\mathbf{m},z^n,\mathbf{w}^n}} \\ \leq \Delta_{f,z^n} \log \frac{|\mathcal{S}|}{\Delta_{f,z}} \tag{A$\cdot$ 15}$$

from Jensen's inequality for $\mathbf{w}^n, \mathbf{m}$, where we let $\Delta_{f,z^n} := \sum_{\mathbf{m},\mathbf{w}^n} P_{\mathbf{M}\mathbf{W}^n|z^n}(\mathbf{m}, \mathbf{w}^n)\Delta_{f,\mathbf{m},z^n,\mathbf{w}^n}$. Averaging Eq. (A$\cdot$ 15) over z^n, we obtain

$$\int p_{Z^n}(z^n)\Delta_{f,z^n} \log \frac{|\mathcal{S}|}{\Delta_{f,z^n}} dz^n \leq \Delta_f \log \frac{|\mathcal{S}|}{\Delta_f} \tag{A$\cdot$ 16}$$

from Jensen's inequality for z^n. Moreover, averaging Eq. (A$\cdot$ 16) over f, we obtain

$$\mathbb{E}_f \left[\Delta_f \log \frac{|\mathcal{S}|}{\Delta_f}\right] \leq \mathbb{E}_f[\Delta_f] \log \frac{|\mathcal{S}|}{\mathbb{E}_f[\Delta_f]} \tag{A$\cdot$ 17}$$

from Jensen's inequality for f. □

Note that when we use Jensen's inequality for a continuous random variable, the condition of absolutely integrable

$$\int p_{Z^n}(z^n)|\Delta_{f,z^n}| dz^n < \infty \tag{A$\cdot$ 18}$$

must be satisfied [13]. In this case, from the fact that $0 \leq \Delta_{f,z^n} \leq 2$, this condition is satisfied.

B.3 The Relation between the Size of a Secret Key and the Security of the Protocol

The following lemma relates the size $|\mathcal{S}|$ of a secret key S and the size $|\mathcal{M}_A \times \mathcal{M}_B|$ of public messages $\mathbf{M}$ to the security of the protocol.

Lemma 4 For the size $|\mathcal{S}|$ of a secret key S, the size $|\mathcal{M}_A \times \mathcal{M}_B|$ of public messages $\mathbf{M}$, and the security of the protocol Δ_f, we have

$$\mathbb{E}_f[\Delta_f] \\ \leq \sqrt{\frac{|\mathcal{S}||\mathcal{M}_A \times \mathcal{M}_B|}{2^{\alpha n}}} \\ +2 \int p_Z^n(z^n) \sum_{\mathbf{w}^n} P_{\mathbf{W}^n|z^n}(\mathbf{w}^n) \\ \times P_{\mathbf{R}^n|z^n\mathbf{w}^n}\left(\left\{\mathbf{r}^n \in \mathcal{X}_\Delta^n \times \mathcal{Y}_\Delta^n \mid \\ -\frac{1}{n} \log P_{\mathbf{R}^n|z^n\mathbf{w}^n}(\mathbf{r}^n) < \alpha\right\}\right) dz^n, \tag{A$\cdot$ 19}$$

where $\mathbb{E}_f$ denotes expectation for a uniform distribution on $\mathcal{F}$.

Proof. This proof is based on the techniques in [20, Chapter 5]. In the following, we will prove

$$\mathbb{E}_f[\Delta_{f,z^n,\mathbf{w}^n}] \\ \leq \sqrt{\frac{|\mathcal{S}||\mathcal{M}_A \times \mathcal{M}_B|}{2^{\alpha n}}} \\ +2P_{\mathbf{R}^n|z^n\mathbf{w}^n}\left(\left\{\mathbf{r}^n \in \mathcal{X}_\Delta^n \times \mathcal{Y}_\Delta^n \mid \\ -\frac{1}{n} \log P_{\mathbf{R}^n|z^n\mathbf{w}^n}(\mathbf{r}^n) < \alpha\right\}\right), \tag{A$\cdot$20}$$

where

$$\Delta_{f,z^n,\mathbf{w}^n} = \|P_{S\mathbf{M}|z^n\mathbf{w}^n} - P_{\bar{S}} \times P_{\mathbf{M}|z^n\mathbf{w}^n}\|, \tag{A$\cdot$ 21}$$

Averaging Eq. (A· 20) over z^n and $\mathbf{w}^n$, we obtain Eq. (A· 19).

For given $z^n \in \mathbb{R}^n$ and $\mathbf{w}^n \in \mathcal{W}_A^n \times \mathcal{W}_B^n$, we define the set $A_n \subset \mathcal{X}_\Delta^n \times \mathcal{Y}_\Delta^n$ as

$$A_n := \left\{ \mathbf{r}^n \in \mathcal{X}_\Delta^n \times \mathcal{Y}_\Delta^n \mid -\frac{1}{n}\log P_{\mathbf{R}^n|z^n\mathbf{w}^n}(\mathbf{r}^n) \geq \alpha \right\},$$

and we define the set A_n^c as the complement of A_n on $\mathcal{X}_\Delta^n \times \mathcal{Y}_\Delta^n$. Then, $\Delta_{f,z^n,\mathbf{w}^n}$ for given $f \in \mathcal{F}$ is upper bounded by

$$\begin{aligned}
&\|P_{S\mathbf{M}|z^n\mathbf{w}^n} - P_{\bar{S}} \times P_{\mathbf{M}|z^n\mathbf{w}^n}\| \\
&= \sum_{s,\mathbf{m}} |P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(f^{-1}(s),\mathbf{m}) \\
&\quad -P_{\bar{S}}(s)P_{\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{m})| && \text{(A· 22)} \\
&= \sum_{s,\mathbf{m}} |P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(f^{-1}(s)\cap A_n,\mathbf{m}) \\
&\quad -P_{\bar{S}}(s)P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(A_n,\mathbf{m}) \\
&\quad +P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(f^{-1}(s)\cap A_n^c,\mathbf{m}) \\
&\quad -P_{\bar{S}}(s)P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(A_n^c,\mathbf{m})| && \text{(A· 23)} \\
&\leq \sum_{s,\mathbf{m}} h_n(s,\mathbf{m}) + \sum_{s,\mathbf{m}} P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(f^{-1}(s)\cap A_n^c,\mathbf{m}) \\
&\quad +\sum_{s,\mathbf{m}} P_{\bar{S}}(s)P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(A_n^c,\mathbf{m}) && \text{(A· 24)} \\
&= \sum_{s,\mathbf{m}} h_n(s,\mathbf{m}) + 2P_{\mathbf{R}^n|z^n\mathbf{w}^n}(A_n^c). && \text{(A· 25)}
\end{aligned}$$

where

$$\begin{aligned}
h_n(s,\mathbf{m}) &= |P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(f^{-1}(s)\cap A_n,\mathbf{m}) \\
&\quad -P_{\bar{S}}(s)P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(A_n,\mathbf{m})|. && \text{(A· 26)}
\end{aligned}$$

Equation (A· 22) follows from the definition of the variational distance and $f^{-1}(s)$. Eq. (A· 23) follows from the fact that $(f^{-1}(s)\cap A_n)\cap(f^{-1}(s)\cap A_n^c)=\emptyset$, $f^{-1}(s)=(f^{-1}(s)\cap A_n)\cup(f^{-1}(s)\cap A_n^c)$, and $P_{\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{m}) = P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(A_n,\mathbf{m}) + P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(A_n^c,\mathbf{m})$. Eq. (A· 24) follows from the triangle inequality. Eq. (A· 25) follows from the fact that $\cup_{s\in\mathcal{S}} f^{-1}(s) = \mathcal{X}_\Delta^n \times \mathcal{Y}_\Delta^n$. By regarding the first term in Eq. (A· 25) as an inner product, and by using the Cauchy-Schwarz inequality, we can upper bound the first term in Eq. (A· 25) by

$$\sum_{s,\mathbf{m}} h_n(s,\mathbf{m}) \leq \sqrt{|\mathcal{S}||\mathcal{M}_A \times \mathcal{M}_B| \sum_{s,\mathbf{m}} h_n(s,\mathbf{m})^2} \qquad \text{(A· 27)}$$

Furthermore, we can rewrite the inside of the root of Eq. (A· 27) as

$$\begin{aligned}
&\sum_{s,\mathbf{m}} h_n(s,\mathbf{m})^2 \\
&= \sum_{s,\mathbf{m}} \Big\{ P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(f^{-1}(s)\cap A_n,\mathbf{m})^2 \\
&\qquad -2P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(f^{-1}(s)\cap A_n,\mathbf{m}) \\
&\qquad P_{\bar{S}}(s)P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(A_n,\mathbf{m}) \\
&\qquad +P_{\bar{S}}(s)^2 P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(A_n,\mathbf{m})^2 \Big\} \\
&= \sum_{s,\mathbf{m}} P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(f^{-1}(s)\cap A_n,\mathbf{m})^2 \\
&\quad -\sum_{\mathbf{m}} \frac{1}{|\mathcal{S}|} P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(A_n,\mathbf{m})^2, && \text{(A· 28)}
\end{aligned}$$

where Eq. (A· 28) follows from the fact that $P_{\bar{S}}(s) = \frac{1}{|\mathcal{S}|}$ and $\sum_s P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(f^{-1}(s)\cap A_n,\mathbf{m}) = P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(A_n,\mathbf{m})$. Then, we can rewrite the first term of Eq. (A· 28) as

$$\begin{aligned}
&\sum_{s,\mathbf{m}} P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(f^{-1}(s)\cap A_n,\mathbf{m})^2 \\
&= \sum_{s,\mathbf{m}} \sum_{\mathbf{r}^n,\mathbf{r}'^n \in f^{-1}(s)\cap A_n} P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{r}^n,\mathbf{m}) \\
&\qquad P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{r}'^n,\mathbf{m}) \\
&= \sum_{\mathbf{m}} \sum_{\mathbf{r}^n,\mathbf{r}'^n \in A_n} \delta_{f(\mathbf{r}^n),f(\mathbf{r}'^n)} P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{r}^n,\mathbf{m}) \\
&\qquad P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{r}'^n,\mathbf{m}), && \text{(A· 29)}
\end{aligned}$$

where $\delta_{f(\mathbf{r}^n),f(\mathbf{r}'^n)}$ is Kronecker's delta. On the other hand, we can rewrite the second term of Eq. (A· 28) as

$$\begin{aligned}
&\sum_{\mathbf{m}} \frac{1}{|\mathcal{S}|} P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(A_n,\mathbf{m})^2 \\
&= \sum_{\mathbf{m}} \sum_{\mathbf{r}^n,\mathbf{r}'^n \in A_n} \frac{1}{|\mathcal{S}|} P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{r}^n,\mathbf{m}) \\
&\qquad P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{r}'^n,\mathbf{m}). && \text{(A· 30)}
\end{aligned}$$

Thus, averaging Eq. (A· 28) over f, we obtain

$$\begin{aligned}
&\sum_{\mathbf{m}} \sum_{\mathbf{r}^n,\mathbf{r}'^n \in A_n} \mathbb{E}_f \left[\delta_{f(\mathbf{r}^n),f(\mathbf{r}'^n)} - \frac{1}{|\mathcal{S}|} \right] \\
&P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{r}^n,\mathbf{m}) P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{r}'^n,\mathbf{m}). && \text{(A· 31)}
\end{aligned}$$

Since f is chosen from a universal-hash-family, we obtain

$$\mathbb{E}_f \left[\delta_{f(\mathbf{r}^n),f(\mathbf{r}'^n)} - \frac{1}{|\mathcal{S}|} \right] \leq \begin{cases} 1 & \text{for } \mathbf{r}^n = \mathbf{r}'^n \\ 0 & \text{for } \mathbf{r}^n \neq \mathbf{r}'^n \end{cases}$$

from its definition (shown in Eq. (A· 9)). Thus, Eq. (A· 31) is upper bounded by

$$\begin{aligned}
&\sum_{\mathbf{m}} \sum_{\mathbf{r}^n \in A_n} P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{r}^n,\mathbf{m}) P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{r}^n,\mathbf{m}) \\
&\leq \sum_{\mathbf{m}} \sum_{\mathbf{r}^n \in A_n} P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{r}^n,\mathbf{m}) \frac{1}{2^{\alpha n}} && \text{(A· 32)} \\
&\leq \sum_{\mathbf{r}^n,\mathbf{m}} P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{r}^n,\mathbf{m}) \frac{1}{2^{\alpha n}} && \text{(A· 33)} \\
&= \frac{1}{2^{\alpha n}}, && \text{(A· 34)}
\end{aligned}$$

where Eq. (A· 32) follows from the fact that $P_{\mathbf{R}^n\mathbf{M}|z^n\mathbf{w}^n}(\mathbf{r}^n,\mathbf{m}) \leq P_{\mathbf{R}^n|z^n\mathbf{w}^n}(\mathbf{r}^n) \leq \frac{1}{2^{\alpha n}}$ for any $\mathbf{r}^n \in A_n$. Since the root function $\sqrt{\cdot}$ is concave function, by combining Eqs. (A· 22)–(A· 32) and averaging over f, we obtain

$$\begin{aligned}
&\mathbb{E}_f[\Delta_{f,z^n,\mathbf{w}^n}] \\
&\le \sqrt{|\mathcal{S}||\mathcal{M}_A \times \mathcal{M}_B| \sum_{s,\mathbf{m}} h_n(s,\mathbf{m})^2} \\
&\quad +2P_{\mathbf{R}^n|z^n\mathbf{w}^n}\left(\left\{\mathbf{r}^n \in \mathcal{X}^n_\Delta \times \mathcal{Y}^n_\Delta \mid \right.\right. \\
&\quad \left.\left. -\frac{1}{n}\log P_{\mathbf{R}^n|z^n\mathbf{w}^n}(\mathbf{r}^n) < \alpha\right\}\right) \\
&\le \sqrt{\frac{|\mathcal{S}||\mathcal{M}_A \times \mathcal{M}_B|}{2^{\alpha n}}} \\
&\quad +2P_{\mathbf{R}^n|z^n\mathbf{w}^n}\left(\left\{\mathbf{r}^n \in \mathcal{X}^n_\Delta \times \mathcal{Y}^n_\Delta \mid \right.\right. \\
&\quad \left.\left. -\frac{1}{n}\log P_{\mathbf{R}^n|z^n\mathbf{w}^n}(\mathbf{r}^n) < \alpha\right\}\right). \qquad (\mathrm{A}\cdot 35)
\end{aligned}$$

□

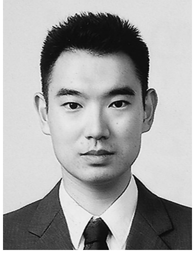
Corollary 1 Suppose that we set $\frac{1}{n}\log|\mathcal{S}| = H(\mathbf{R}|Z\mathbf{W}) - \frac{1}{n}\log|\mathcal{M}_A \times \mathcal{M}_B| - 2\delta$, $\mathbb{E}_f[\Delta_f]$ is exponentially small for sufficiently large n.

Proof. Suppose that we set $\alpha = H(\mathbf{R}|Z\mathbf{W}) - \delta$ for $\delta > 0$, the second term of Eq. (A· 19) exponentially tends to 0 as $n \to \infty$ by using the Chernoff bound [7]. On the other hand, suppose that we set $\frac{1}{n}\log|\mathcal{S}| = H(\mathbf{R}|Z\mathbf{W}) - \frac{1}{n}\log|\mathcal{M}_A \times \mathcal{M}_B| - 2\delta$, the first term of Eq. (A· 19) is $e^{-\delta n}$ and tends to 0 as $n \to \infty$. Thus, suppose that we set $\frac{1}{n}\log|\mathcal{S}| = H(\mathbf{R}|Z\mathbf{W}) - \frac{1}{n}\log|\mathcal{M}_A \times \mathcal{M}_B| - 2\delta$, $\mathbb{E}_f[\Delta_f]$ exponentially tends to 0 as $n \to \infty$. □

If $\mathbb{E}_f[\Delta_f]$ is exponentially small, then the security of the protocol in the sense of entropy is guaranteed by lemma 3. From this fact and corollary 1, suppose that we set $\frac{1}{n}\log|\mathcal{S}| < H(X_\Delta Y_\Delta|ZW_AW_B) - \frac{1}{n}\log|\mathcal{M}_A \times \mathcal{M}_B|$, Eq. (10) is satisfied for sufficiently large n.

Masashi Naito was born in Aichi, Japan, on July 5, 1985. He received the B.E. degree from Tokyo Institute of Technology in 2008. He is currently a master's student in the Department of Media Science of the Graduate School of Information Science of Nagoya University. His current research interests are in the areas of pattern recognition.

Shun Watanabe was born in Tokyo, Japan, on February 2, 1983. He received the B.E. and M.E. degrees from Tokyo Institute of Technology in 2005 and 2007 respectively. He is currently a doctoral student in the Department of Communications and Integrated Systems of Tokyo Institute of Technology. His current research interests are in the areas of information theory, quantum information theory, and quantum cryptography.



Ryutaroh Matsumoto was born in Nagoya, Japan, on November 29, 1973. He received the B.E. degree in computer science, the M.E. degree in information processing, and the Ph.D. degree in electrical and electronic engineering, all from Tokyo Institute of Technology, Japan, in 1996, 1998, 2001, respectively. He was an Assistant Professor from 2001 to 2004, and has been an Associate Professor since 2004 in the Department of Communications and Integrated Systems of Tokyo Institute of Technology. His research interest includes error-correcting codes, quantum information theory, and communication theory. Dr. Matsumoto received the Young Engineer Award from IEICE and the Ericsson Young Scientist Award from Ericsson Japan in 2001. He received the Best Paper Awards from IEICE in 2001 and 2008.

Tomohiko Uyematsu received the B.E., M.E. and Dr.Eng. degrees from Tokyo Institute of Technology in 1982, 1984 and 1988, respectively. From 1984 to 1992, He was with the Department of Electrical and Electronic Engineering of Tokyo Institute of Technology, first as research associate, next as lecturer, and lastly as associate professor. From 1992 to 1997, he was with School of Information Science of Japan Advanced Institute of Science and Technology as associate professor. Since 1997, he returned to Tokyo Institute of Technology as associate professor, and currently he is with the Department of Communications and Integrated Systems as professor. In 1992 and 1996, he was a visiting researcher at the Centre National de la Recherche Scientifique, France and Delft University of Technology, Netherlands, respectively. He received Shinohara Memorial Young Engineer Award in 1989, and the Best Paper Award in 1993, 1996, 2002, and 2007 all from IEICE. His current research interests are in the areas of information theory, especially Shannon theory and multi-terminal information theory. Dr. Uyematsu is a senior member of IEEE.