Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/080226

Generalized Upper Bound of Agreement Probability for Extracting Common Random Bits From Correlated Sources

Young-Sik Kim[1] *and Dae-Woon Lim* [2,*]

[1] Department of Information and Communication Engineering, Chosun University, Gwangju 501-759, Korea
[2] Department of Information and Communication Engineering, Dongguk University, Seoul 100-715, Korea

Received: 30 Mar. 2013, Revised: 31 Jul. 2013, Accepted: 2 Aug. 2013
Published online: 1 Mar. 2014

Abstract: Suppose that both Alice and Bob receive independent random bits without any bias, which are influenced by an independent noise. From the received random bits, Alice and Bob are willing to extract common randomness, without any communication. The extracted common randomness can be used for authentication or secrets. Recently, Bogdanov and Mossel derived an upper bound of the agreement probability, based on the min-entropy of outputs. In this paper, we derive a generalized upper bound of the probability of extracting common random bits from correlated sources, using the Rènyi entropy of order $1/(1-\varepsilon)$, where ε is the error probability of the binary symmetric noise. It is shown that the generalized upper bound is always less than or equal to the previous bound.

Keywords: Rènyi entropy, common randomness, agreement probability, secret extraction, information reconciliation.

1 Introduction

Recent research has studied the exploitation of processing deviations for the unique identification of electronic devices, by extracting unique information, which is influenced by the internal electronic characteristics of the device [1]. However, since the extraction process can be influenced by random noise, which is affected by temperature, electro-magnetic waves, or interference between other devices, a small deviation in the extraction process can result in a disagreement between the final results on both sides. Therefore, a good information distillation algorithm is required, to increase the agreement probability for a pair of extraction processes.

In this paper, we derive an upper bound of agreement probability of common random bits that are independently extracted by Alice and Bob, without any communication from correlated sources, when noise independently and uniformly occurs between each bit. There are several studies for the theoretical performance of single bit extraction by Alice and Bob [2],[3]. A trivial protocol is known, in which both sides use the first k-bit of the generated n-bit random data. Recently, Bogdanov and Mossel derived an upper bound, based on the min-entropy k of outputs [3]. In their result, it is shown that there is no protocol that can achieve higher agreement probability than $2^{-k\varepsilon/(1-\varepsilon)}$. In this paper, we derive a new upper bound, based on the Rènyi entropy of order $1/(1-\varepsilon)$, where ε is the error probability of the binary symmetric noise. It is shown that the new upper bound is always less than or equal to the previous upper bound, based on the min-entropy of the generated outputs.

This paper is organized as follows. In Section 2, we briefly summarize the related work, and explain basic notions of entropies for self-containedness. In Section 3, the new upper bound of agreement probability that the same random bits are extracted from correlated sources is derived. In addition, because the Rènyi entropy can be approximated by the Shannon entropy for small error probability ε, we will represent the new bound in terms of the Shannon entropy. In Section 4, numerical simulations on the upper bounds of agreement probability with the various error probability ε and output bias will be presented Finally, we conclude this paper in Section 5.

* Corresponding author e-mail: daewoonlim@gmail.com

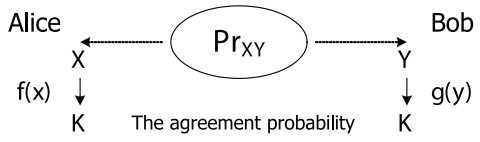

Fig. 1: Security models of random agreement.

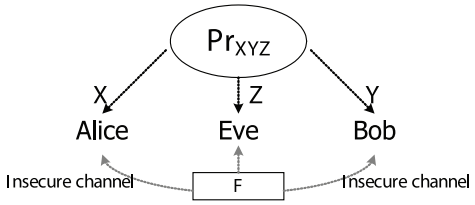

Fig. 2: Security models of information reconciliation problem.

2 Preliminaries

2.1 Related Works

In this subsection, two cases are considered as related research fields of this paper. Firstly, physically unclonable functions (PUFs) are widely studied for various applications [1]. The PUF is usually embedded in a physical electronic device, and generates output that is easily evaluated and hard to predict, based on the user input. It is known that every device has slightly distinct characteristics, due to fabricational variations that cannot be completely removed, despite many tries of getting rid of it, for decades. Instead of trying to prevent it, the PUF exploits the phenomenon to generate a unique signature of the device, which can be used for authentication or the generation of a secret.

Consider that Alice extracts a signature sequence from a PUF at time t_1, and later Bob also extracts the signature sequence from the same PUF at time t_2. For successful authentication, the extracted outputs should be the same. However, since electronic characteristics are easily affected by their environments, such as temperature and electro-magnetic fields, the generated output can be slightly changed. This situation can be modeled by extracting output from a random source with noise, which is the case considered in this paper, and is depicted in Fig. 1. In this case, it is important to ensure that the generated output is always the same with the same input, for the use of authentication.

In the seminal work by Shannon [4], the well-known pessimistic conclusion on perfect secrecy states that prefect secrecy is possible only when both sides share a secret key that is as long as the user message to be protected, if an adversary can perfectly access the cipher-text. After that, Wyner [5] and Csiszár and Körner [6] showed that it is possible to securely communicate between Alice and Bob, under the assumption of the presence of noise in communication channels from Alice to Bob, and to an adversary Eve. By taking various assumptions on the adversary's ability, many applications have been derived in the field of physical layer security.

Among them, information reconciliation processes are proposed in order to solve many problems in cryptography and security issues [7,8,9,11,12,13,14,15,16]. In this problem, it is assumed that Alice and Bob want to share a secret key by generating random numbers from a correlated source with noise, instead of securely transferring from Alice to Bob. Here, an adversary Eve can access the correlated random source with independent noise and observe the noiseless discussion message F. Through an interactive protocol, Alice, Bob, and even Eve receive correlated information, which is characterized by independent repetitions of a random experimented $\Pr_{XYZ}$, as depicted in Fig. 2. Maurer [7] and Ahlswede and Csiszár [8] showed that it is possible to generate common randomness K between Alice and Bob, without disclosing any information on K to Eve. In their results, Alice and Bob have an insecure communication channel in both directions. Information reconciliation problems are similar to randomness extraction, or correlation distillation problems. However, the major issue for information reconciliation problems is to make sure that the adversary Eve cannot obtain any information on K, as well as to extract common information K on both sides.

2.2 Definitions of Entropies and Related Notions

In this subsection, we briefly introduce the definitions of various entropies. Suppose that the generated k output bits are considered as a symbol. Then, the randomness of the output can be measured by various entropies, based on the k-bit blocks. For a random variable of the k-bit random bits, the Shannon entropy is defined as

$$H_k(X) = -\sum_{z \in F_{2^k}} \Pr(z) \log_2 \Pr(z).$$

Note that the maximum value of Shannon entropy is k-bit, which can be obtained when the probability of each symbol occurring is equal to $1/2^k$. There are other entropy definitions as measures of randomness or uncertainty. The Rènyi entropy of order α for the random variable X is defined as [10]

$$R_{\alpha,k}(X) = \frac{1}{1-\alpha} \log_2 \sum_{z \in F_{2^k}} \Pr(z)^\alpha \qquad (1)$$

where X is a random variable representing k-bit random values. Note that the Rènyi entropy produces different values, according to the value of order α. Another widely used entropy notion is the min-entropy defined as

$$M_k(X) = -\log_2 \sup_{z \in F_{2^k}} \Pr(z).$$

The statistical distance between two distributions D and D' over a sample space Ω is defined as

$$\sum_{\omega\in\Omega} |\Pr_{D}(\omega) - \Pr_{D'}(\omega)|.$$

A probability distribution D is said to be δ-close to min-entropy t, if there is a probability distribution D' of min-entropy t, where the statistical distance between D and D' is less than δ. Then, Bogdanov and Mossel proved the following theorem.

Theorem 1. *[3] For any two functions $f,g : \{0,1\}^n \to \{0,1\}^k$ that are δ-close to min-entropy t and every $\varepsilon \leq 1/2$, we have*

$$\Pr_{(x,y)_\varepsilon}[f(x) = g(y)] < 2^{-t\varepsilon/(1-\varepsilon)} + 2\delta.$$

If the outputs $f(x)$ and $g(y)$ at each side are exactly uniform, the size k of the output is equal to the min-entropy t and $\delta = 0$. Then, for generated $1/\varepsilon$ common random bits (i.e., $k = 1/\varepsilon$), the agreement probability of the outputs of Alice and Bob is less than $1/2$ [3].

Suppose that x and y are binary sequences with length n. Each sequence is generated by the following random process: x_i, the i-th bit of x, is independently selected from the set $\{0,1\}$. In this case, y_i, the i-th bit of y, is the same as x_i with the probability $1-\varepsilon$, and $1-x_i$ with the probability ε. That is, the probability that each side extracts a different bit is ε. Without loss of generality, we can set the error probability as $\varepsilon \leq 1/2$.

Suppose that there is a protocol to generate the same k-bit on both sides of Alice and Bob, which can be represented by a pair of functions f and g, such as $f,g : \{0,1\}^n \to \{0,1\}^k$, where $f(x)$ and $g(y)$ mean the k-bit outputs of Alice and Bob from n-bit inputs x and y, respectively. It is assumed that the inputs x and y are uniform.

3 Upper Bound of Common Random Bit Extraction

The same auxiliary lemmas are required to derive the new upper bound, based on the Rènyi entropy, which is given in [3]. The first lemma is a consequence of the fact that $E_{(x,y)_\varepsilon}[f(x),g(y)]$ is an inner product of f and g.

Lemma 1. *[3] For every pair of functions $f,g : \{0,1\}^n \to \mathbb{R}$, we have*

$$E_{(x,y)_\varepsilon}[f(x)g(y)] \leq \sqrt{E_{(x,y)_\varepsilon}[f(x)(f(y)]} \times \sqrt{E_{(x,y)_\varepsilon}[g(x)(g(y)]}.$$

The next lemma is a result of the hypercontractive inequality.

Lemma 2. *[3] For a function $h : \{0,1\}^n \to \{0,1\}$, we have*

$$E[h(x)h(y)] \leq E[h(x)]^{1/(1-\varepsilon)}.$$

A probability distribution D is said to be δ-close to Rènyi entropy r_α of order α if there is a probability distribution D' of Rènyi entropy r_α of order α, where the statistical distance between D and D' is less than δ. Using the above lemmas, we can similarly prove the new upper bound, as in the following theorem.

Theorem 2. *Suppose that the k-bit outputs functions $f,g : \{0,1\}^n \to \{0,1\}^k$ are δ-close to Rènyi entropy r_α of order $\alpha = 1/(1-\varepsilon)$. Then, for $\varepsilon \leq 1/2$, we have*

$$\Pr[f(x) = g(y)] \leq 2^{-\varepsilon r_\alpha/(1-\varepsilon)} + 2\delta.$$

Proof: The first steps of this proof can be carried out in the same way as given in [3]. However, we repeat them for the self-containedness. Suppose that the k-bit outputs functions f and g have the Rényi entropy of order α. For $z \in \{0,1\}^k$, let f_z and g_z be the functions $\{0,1\}^n \to \{0,1\}$, defined as

$$f_z(x) = \begin{cases} 1, & \text{if } f(x) = z \\ 0, & \text{otherwise} \end{cases}$$

and similarly,

$$g_z(x) = \begin{cases} 1, & \text{if } g(x) = z \\ 0, & \text{otherwise.} \end{cases}$$

By using the same method in [3], the agreement probability can be bounded as

$$\Pr[f(x) = g(y)] = \sum_{z\in\{0,1\}^k} \Pr[(f(x) = z) \wedge (g(y) = z)] \quad (2)$$

$$= \sum_{z\in\{0,1\}^k} E[f_z(x), g_z(y)] \quad (3)$$

$$\leq \sum_{z\in\{0,1\}^k} \sqrt{E[f_z(x)f_z(y)]E[g_z(x)g_z(y)]} \quad (4)$$

$$\leq \sum_{z\in\{0,1\}^k} \sqrt{E[f_z(x)]^{1/(1-\varepsilon)}E[g_z(x)]^{1/(1-\varepsilon)}} \quad (5)$$

$$\leq \sqrt{\sum_{x\in\{0,1\}^k} E[f_z(x)]^{1/(1-\varepsilon)}} \times \sqrt{\sum_{x\in\{0,1\}^k} E[g_z(x)]^{1/(1-\varepsilon)}}. \quad (6)$$

The first equality in (2) comes from $f(x) = g(y) = z$, and the second equality in (3) is satisfied from the fact that $f(x) = z$ and $g(y) = z$ if and only if $f_z(x)g_z(y) = 1$. The inequalities in (4) and (5) are obtained from Lemmas 1 and 2, respectively. Finally, the last inequality

in (6) can be derived from the Cauchy-Schwarz inequality.

From the assumption of this theorem, f and g generate k-bit outputs that are δ-close to Rènyi entropy r_α of order $\alpha = 1/(1-\varepsilon)$. Let p_z be the expectations of $E[f_z(x)] = E[g_z(y)]$. Then, the right side of the last inequality can be rewritten as

$$\sum_{z\in\{0,1\}^k} p_z^{1/(1-\varepsilon)} = 2^{-r_\alpha(1-\alpha)} = 2^{-\varepsilon r_\alpha/(\varepsilon-1)}.$$

This is because of the definition of the Rènyi entropy of order α given in (1). Note that this is a more straightforward representation than the previous approach [3], which uses the inequality $p_z \le 2^t$, by the definition of the min-entropy.

Next, by the same argument in [3], we can prove the case that the functions f and g are close Rényi entropy of order α distributions Let $\delta' > \delta$. By taking a larger value of n, there exist f' and g' of Renyi entropy of order α, such that $\Pr[f \neq f'] \le \delta'$ and $\Pr[g \neq g'] \le \delta'$. Clearly, we have $\Pr[f = g] \times \Pr[f = f'] \times \Pr[g = g'] = \Pr[f' = g']$. Then for $\delta' \ll 1$, we have the following inequality

$$\begin{aligned}
\Pr[f(x) = g(x)] &\le \Pr[f'(x) = g'(x)](1-\delta')^{-2} \\
&= \Pr[f'(x) = g'(x)](1 + \sum_{k=2}^{\infty} k\delta'^{k-1}) \\
&= \Pr[f'(x) = g'(x)] + \Pr[f'(x) = g'(x)] \sum_{k=2}^{\infty} k\delta'^{k-1} \\
&\le \Pr[f'(x) = g'(x)] + 2\delta' \\
&\le 2^{-\varepsilon r_\alpha/(\varepsilon-1)} + 2\delta'
\end{aligned}$$

when $\Pr[f'(x) = g'(x)] \le 2(1-\delta')^2/(2-\delta')$. Because δ is an arbitrary value, the theorem is proved.

□

Because the order of the Rènyi entropy is equal to $\alpha = 1/(1-\varepsilon)$, for $\varepsilon \to 0$, we have $\alpha \to 1$. Therefore, we can obtain an upper bound with respect to the Shannon entropy of the extracted output, with very small error probability ε.

Corollary 1. *For $\varepsilon \to 0$, we have*

$$\Pr[f(x) = g(y)] \le 2^{-\varepsilon H_k}$$

where H_k is the Shannon entropy for the k-bit random blocks.

Proof:

It is well-known that the Renyi's entropy converges to the Shannon entropy, which can be checked for $\alpha \to 1$,

$$\begin{aligned}
R_k(X) &= \frac{1}{1-\alpha} \log_2 \sum_{z\in\{0,1\}^k} [\Pr(z)]^\alpha \\
&= -\frac{\log_2 \sum_{z\in\{0,1\}^k} \Pr(z) - \log_2 \sum_{z\in\{0,1\}^k} [\Pr(z)]^\alpha}{1-\alpha} \\
&= -\frac{d}{d\alpha} \log_2 \sum_{z\in\{0,1\}^k} [\Pr(z)]^\alpha \,|_{\alpha=1} \\
&= -\sum_{z\in\{0,1\}^k} \Pr(z) \log_2 \Pr(z) \\
&= H_k(X).
\end{aligned}$$

Therefore, for $\alpha \to 1$, the Rènyi entropy of order α converges to the Shannon entropy. From Theorem 3, for $\varepsilon \to 0$, we have $\alpha = 1/(1-\varepsilon) \to 1$, then we have

$$\Pr[f(x) = g(y)] \le 2^{-\varepsilon H_k}.$$

□

Now, it is possible to show that the upper bound in Theorem 3 is always less than or equal to that in [3].

Theorem 3. *For $\alpha > 1$ and two arbitrary functions $f, g : \{0,1\}^n \to \{0,1\}^k$, suppose that the Rènyi entropy of order α and min-entropy of the outputs from f and g are equal to r_α and t, respectively. Then we have*

$$2^{-\varepsilon r_\alpha/(1-\varepsilon)} \le 2^{-\varepsilon t/(1-\varepsilon)}.$$

Proof:
We have

$$\begin{aligned}
\log_2 \sum_{z\in F_{2^k}} p_z^\alpha &\le \log_2 \sum_{z\in F_{2^k}} p_z^{\alpha-1} \sum_{z\in F_{2^k}} p_z \\
&= (\alpha-1) \log_2 \sup_{z\in F_{2^k}} p_z.
\end{aligned}$$

Therefore if the factor $(\alpha-1)$ on the right side is moved to the left side and multiplied by -1 on both sides, we have

$$r_\alpha = \frac{1}{1-\alpha} \log_2 \sum_{z\in F_{2^k}} p_z^\alpha \ge t.$$

□

4 Numerical Results

If the functions f and g produce common random bits with statistical bias η, the bias can reduce the amount of output entropies. Fig. 3 shows a comparison between the upper bounds of the agreement probability, based on the min-entropy and Rènyi entropy of order α, with respect to the bias η of the output bits. In this case, it is assumed that the output bias η is equal to the error probability ε.

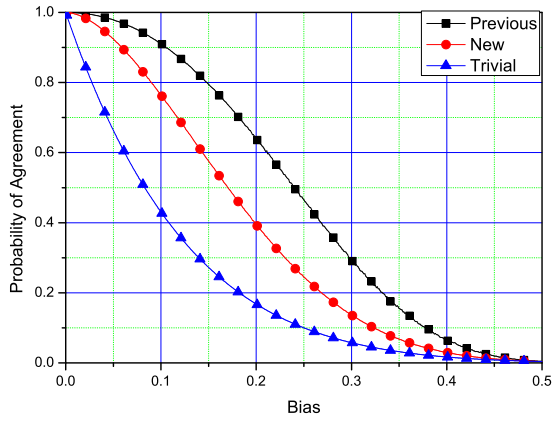

Fig. 3: The new upper bound of agreement probability in terms of bias of output bits.

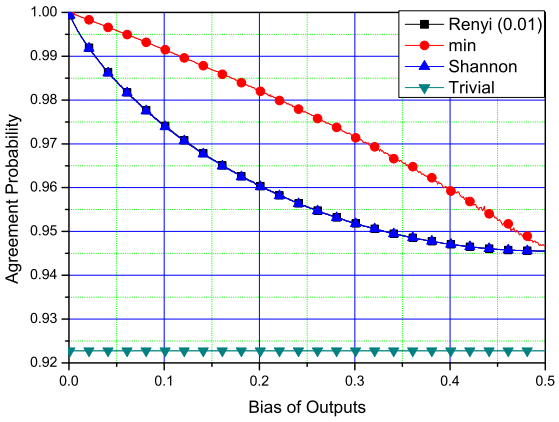

Fig. 5: Comparison of the upper bounds, based on Rènyi and min-entropies with the error probability $\varepsilon = 0.01$.

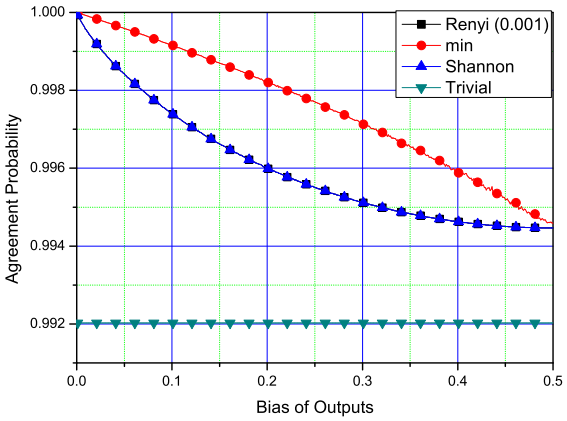

Fig. 4: Comparison of the upper bounds, based on Rènyi and min-entropies with the error probability $\varepsilon = 0.001$.

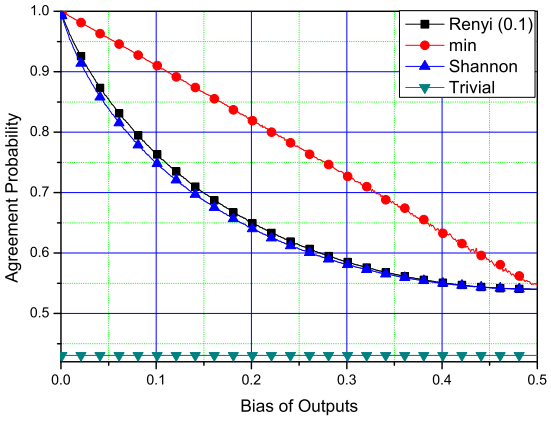

Fig. 6: Comparison of the upper bounds, based on Rènyi and min-entropies with the error probability $\varepsilon = 0.1$.

This corresponds to the case that the functions f and g produce the output values with statistical bias that is proportional to the error probability. The top line with the block squares in Fig. 3 corresponds to the previous upper bound of the agreement probability presented in [3]. The center line with the red circles in Fig. 3 shows the new upper bound, based on the Rènyi entropy of order $\alpha = 1/(1-\varepsilon)$ presented in this paper. Finally, for comparison, the bottom line with the blue triangles depicts the agreement probability of the trivial extraction that is choosing the first k bits from the inputs. Consider that by taking a k-bit part from the n-bit random input, the uniformity of the input sequence can be broken. The effect of the bias at the output is illustrated in Fig. 3 for the trivial case. Note that the proposed upper bound is always less than or equal to the previous one presented in Theorem 3.

The next four figures from Fig. 4 to Fig. 7 show the comparison results between the Rènyi entropy and min-entropy, with respect to the four fixed error probabilities. Moreover, these figures show the main statement in Corollary 1, which states that the new upper

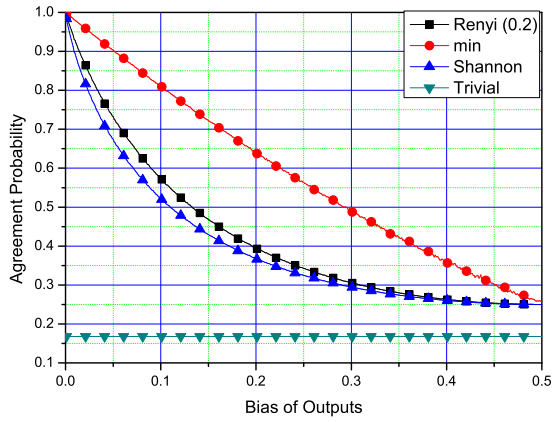

Fig. 7: Comparison of the upper bounds, based on Rènyi and min-entropies with the error probability $\varepsilon = 0.2$.

bound can be approximated by using the Shannon entropy, when the error probability ε is sufficiently low. Therefore, in Figs. 4 and 5, the upper bound, based on the Shannon entropy, is almost equal to the proposed upper bound, based on the Rènyi entropy of order $1/(1-\varepsilon)$. However, for the cases of the error probabilities $\varepsilon = 0.1$ or 0.2, there are some deviations between the two lines, based on the Rènyi and Shannon entropies, respectively. Note that the lines obtained by substituting Rènyi entropy values with the Shannon entropy values in Figs. 6 and 7 cannot be legitimate upper bounds. However, because there are still large differences between the upper bounds and the trivial case (i.e., just taking the first k-bits from the correlated n-bit inputs), we can think that it may be possible to find better strategies for the agreement, although the new bound is closer to the trivial case than the previous one. Also, note that as pointed out in [3], when the error probability is large enough, it is more possible to find better protocols for the agreement, because the deviation between the upper bound and the trivial case is greater than that with the lower error probability.

5 Conclusion

For the problem of extracting common randomness by Alice and Bob from correlated sources that are influenced by independent random noise without any bias, we derive the new upper bound of the agreement probability on both sides, based on the Rènyi entropy. We then prove that the new upper bound is always tighter than the previous bound, based on the min-entropy, because of the properties of Rènyi entropy. Moreover, for $\varepsilon \to 0$, the Shannon entropy based tighter upper bound can be used instead, because the Rènyi entropy converges to the Shannon entropy, for the same size of random blocks.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (NRF-2011-0016664). This work was supported by the Power Generation and Electricity Delivery of the KETEP grant funded by the Korea government Ministry of Trade, Industry and Energy (20131020400760).
The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

[1] G. E. Suh and S. Devadas, “Physical Unclonable Functions for Device Authentication and Key Generation,” in *Proc. the 44th Design Automation Conf.*, June, (2007).
[2] K. Yang, “On the impossibility of non-interactive correlation distribution,” *Theoret. Comput. Sci.*, **382**, 157–166 (2007).
[3] A. Bogdanov and E. Mossel, “On Extracting Common Random Bits From Correlated Sources,” *IEEE. Trans. Inf. Theory*, **57**, 6351–6355 (2011).
[4] C. E. Shannon, “Communication theory of secrecy systems,” *Bell Syst. Tech. J.*, **28**, 656–715 (1949).
[5] A. D. Wyner, “The wire-tap channel,” *Bell Syst. Tech. J.*, **54**, 1355–1387 (1975).
[6] I. Csiszár and J. Körner, “Broadcast channels with confidential messages,” *IEEE Trans. Inf. Theory*, **IT-24**, 339–348 (1978).
[7] U. Maurer, “Secret key agreement by public discussion from common information,” *IEEE Trans. Inf. Theory*, **39**, 733–742 (1993).
[8] R. Ahlswede and I. Csiszár, “Common randomness in information theory and cryptography. I. Secret sharing,” *IEEE Trans. Inf. Theory*, **39**, 1121–1132 (1993).
[9] U. Maurer and S. Wolf, “Secret-key agreement over unauthenticated public channels–part I: definitions and a completeness result,” *IEEE Trans. Inf. Theory*, **49**, (2003).
[10] A. Rènyi, “On measures of entropy and information,” in *Proc. 4th Berkeley Symp. Mathematical Statistics and Probability*, **1**, 547–561 (1961).
[11] U. Maurer, R. Renner, and S. Wolf, “Unbreakable keys from random noise,” in *Security with Noisy Data*, P. Tuyls, B. Skoric, and T. Kevennar, Eds. Springer-Verlag, 21–44 (2007).
[12] A. A. Gohari and V. Anantharam, “Information-theoretic key agreement of multiple terminals—part I,” *IEEE Trans. Inf. Theory*, **56**, 3973–3996 (2010).
[13] A. Gohari and V. Anantharam, “Information-theoretic key agreement of multiple terminals—part II: Channel models,” *IEEE Trans. Inf. Theory*, **56**, 3997–4010 (2010).

[14] I. Csiszár and P. Narayan, "Common randomness and secrey key generation with a helper," *IEEE Trans. Inf. Theory*, **46**, 344–366 (2000).
[15] A. Khisti, S. N. Diggavi, and G. Wornell, "Secret-key generation with correlated sources and noisy channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, Canada, 1005–1009 (2008).
[16] S. Watanabe and Y. Oohama, "Secrey key agreement from vector Gaussian sources by rate limited public communications," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, 2597–2601 (2010).

Young-Sik Kim received B.S., M.S., and Ph.D. degrees in electrical engineering and computer science from Seoul National University in 2001, 2003, and 2007, respectively. He joined Semiconductor Division, Samsung Electronics and carried out research and development for secure hardware IPs for various embedded systems, especially for smart-cards until the end of August in 2010. He is an assistant professor at Chosun University, Gwangju, Korea. He is an Editor of the Journal of Communications and Networks (JCN) from 2013. His research interests include cryptographic engineering and information theory including hardware security, embedded security, physical layer security, data hiding, channel coding, and signal design.

Dae-Woon Lim received the B.S. and M.S. degrees in department of electrical engineering from KAIST, Daejeon, Korea, in 1994 and 1997, respectively. In 2006, he received the Ph.D. degree in electrical engineering and computer science from Seoul National University. From 1997 to 2002 he was with LG Industrial Systems as a senior research engineer, where he developed recognition algorithm, real-time tracking algorithm, and electric toll collection system. He is currently an associate professor in department of information and communication engineering at Dongguk University, Seoul, Korea. His research interests are in the area of signal processing, wireless communications, cryptography, and security.