Finite-Block-Length Analysis in Classical and Quantum Information Theory

Masahito Hayashi

Abstract

Transmitting an information is a fundamental technology. However, there are several demands for this transmission. The research area to study such problems is called information theory. In the information transmission, the information is transmitted via a physical media. Hence, the analysis of this problem might depend on the property of the physical media. Indeed, while it is ideal that the analysis does not depend on this property, it depends on the following classification of the physical media. Currently, we have two kinds of physical objects, the first one is a macroscopic object, i.e., an object subject to classical physics, and the second one is a microscopic object, i.e., an object subject to quantum physics. Since these two objects have completely different behaviors, we need to build up information theory dependently of these two information media. That is, we have classical information theory and quantum information theory. In both information theory, there are very elegant theoretical results with an ideal assumption. That is, we often assume that infinitely large size of the system is available while the real situation does not satisfy this assumption. Hence, to discuss the real case, we need to care about the finite size effect of the system size. This paper reviews this finite size effect in classical and quantum information theory with respect to various topics.

Index Terms

channel coding, finite block-length, quantum information theory, classical information theory

I. Introduction

It is a fundamental technology to transmit the information correctly via a noisy channel. To resolve this issue, we have a technology of channel coding, which is composed of two parts, an encoder and a decoder. The key point of this technology is adding the redundancy to the original information to protect the original information from the noise. The most simple channel coding is transmitting the same information at three times repetitively as Fig. 1. That is, when we need to send one bit 0 or 1, we transmit the three bits information 0, 0, 0 or 1, 1, 1. When an error occurs only in one bit among three bits, we can easily recover the original bit. The conversion from 0 or 1 to 0, 0, 0 or 1, 1, 1 is called an encoder. The conversion from the noisy three bits to the original one bit is called a decoder. The pair of an encoder and a decoder is called a code. In this code, when two bits are flipped during the transmission, we cannot recover the original information. This code has a large redundancy and the range of correctable errors is limited. For practical use, we need improvement of this code. That is, it is required to decrease the amount of the redundancy and to enlarge the range of the correctable errors. The reason why the above code is so bad is that the block-length of the code is only 3. Here, the block-length is the length of bits of one block in the code. For example, the above code has block-length 3.

In 1948, Shannon [1] discovered that increase of the block-length $n$ can improve the redundancy and the range of the correctable errors. That is, he clarified the minimum rate of the redundancy to correct the error with probability almost 1 with infinitely large block-length $n$. To discuss this problem, for a probability distribution $P$, he introduced the quantity $H(P)$, which is called entropy or Shannon entropy and expresses the uncertainty of the probability distribution $P$. He showed that we can recover the original information by a suitable code when the noise is generated subject to the probability distribution $P$, the rate of redundancy is the entropy $H(P)$, and the block-length $n$ is infinitely large. This fact is called

M. Hayashi is with Graduate School of Mathematics, Nagoya University, Furocho, Chikusaku, Nagoya, 464-8602, Japan, and Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117542. (e-mail: masahito@math.nagoya-u.ac.jp)
channel coding theorem. Under this limit, the limit of the minimum error probability depends only on whether the rate of the redundancy is larger than the entropy $H(P)$ or not.

We can consider a similar problem when the channel is given as the additive white Gaussian noise. In this case, we cannot use the term of the redundancy because its meaning is not clear. In the following, instead of this term, we employ the transmission rate to characterize the speed of the transmission, which expresses the number of transmitted bits per one use of the channel. In the case of the additive white Gaussian channel, the optimal transmission rate is $\frac{1}{2} \log(1 + \frac{S}{N})$, where $\frac{S}{N}$ is the signal-noise ratio [2, Theorem 7.4.4], as channel coding theorem. However, we cannot directly apply the channel coding theorem to the actual information transmission because this theorem guarantees only the existence of a code with the above ideal performance. To construct a practical code, we need another type of theory, which is often called coding theory. Until now, dependently of the strength of the noise of the channel, so many practical codes have been proposed and have been used in the real communication system. These codes realize sufficiently small error probability, but no code could attain the optimal transmission rate. In 1990s, turbo codes and low density parity check (LDPC) codes were actively studied as useful codes [3], [4]. It was theoretically shown that they can attain the optimal transmission rate when the block-length $n$ goes to infinity. However, still no really constructed code could attain the optimal transmission rate. Hence, many researchers doubted what is the real optimal transmission rate. Here, we should emphasize that any really constructed code has finite block-length and it is not necessarily attain the conventional asymptotic transmission rate.

On the other hand, in 1962, Strassen [5] has already addressed this problem by discussing the coefficient with the order $\frac{1}{\sqrt{n}}$ of the transmission rate, which is called the second order asymptotic theory. The calculation of the second order coefficient approximately gives the solution of the above problem, i.e., the real optimal transmission rate with finite block-length $n$. Although he derived the second order coefficient for the discrete channel, he could not derive it for the additive white Gaussian channel. Also, in spite of the importance of his result, many researchers overlooked his result because his paper was written in German. Therefore, the successive researchers had to recover his result without use of his derivation. This paper explains how this problem has been resolved even for additive white Gaussian channel by tracing the long history for classical and quantum information theory. Currently, the finite block-length theory is one of hottest topics in information theory and is more precisely discussed in various situations.

In addition to correct information transmission, information theory studies the data compression (source coding) and (secure) uniform random number generation. In these problems, we address a code with block-length $n$. When the information source is subject to the distribution $P$ and the block-length $n$ is infinitely large, the optimal conversion rate is $H(P)$ in both problems. This paper also reviews these problems.

II. QUANTUM INFORMATION TRANSMISSION

When we discuss the information transmission problem ultimately, we need to address the property of the physical media carrying the information. When we address the ultimate limit of the information transmission rate as a theoretical problem, we need to consider the case when individual particles expresses each information. That is, we focus on the information transmission rate under such an ultimate
situation. To realize the ultimate transmission rate, we need to use every photon to describe the respective information. Since the physical media to transmit the information has quantum behavior under such a situation, the description of the information system needs to reflect the quantum property.

Several researchers, Takahasi [6] etc started to consider this problem to study the limit of the optical communication in 1960s. In 1967, Helstrom [8], [7] started to systematically formulate this problem as a new type of information processing system. That is, he discussed the information transmission system based on quantum theory instead of the conventional information transmission system based on classical mechanics, i.e., the conventional probability theory. Studies of information transmission based on such a quantum media is called quantum information theory. In particular, researches for channel coding based on such a quantum media is called quantum channel coding. When we need to distinguish information theory based on the conventional probability theory from quantum information theory, it is called classical information theory because it is based on classical mechanics. Quantum information theory in the earlier stage has been studied more deeply by Holevo and systematically summarized in his book [9] in 1980.

Fig. 2. Classical channel coding for optical communication

Here, we point out that the current optical communication system is treated in the framework of classical information theory. In fact, the optical communication can be treated in both of classical and quantum information theory as follows (Figs. 2 and 3). Since the framework of classical information theory cannot treat the quantum system, to treat the optical communication within the classical information theory, we need to fix the modulator converting the input signal to the input quantum state and the detector converting the output quantum state to the outcome as Fig. 2. Once we fix them, we have the conditional distribution connecting the input and output signals, which describes the channel in the framework of classical information theory. That is, we can apply classical information to the channel. The encoder is the process converting the message (to be sent) to the input signal, and the decoder is the process recovering the message from the outcome.

On the other hand, when we discuss optical communication within the framework of quantum information theory as Fig. 3, we combine a classical encoder and a modulator into a quantum encoder, in which, the message is directly converted to the input quantum state. Similarly, we combine a classical...
encoder and a detector into a quantum decoder, in which, the message is directly recovered from the output quantum state. Once the optical communication is treated in the framework of quantum information theory, information transmission can be treated in the combination of a quantum encoder and a quantum decoder. Since this framework allows us to employ physical process across plural pulses as a quantum encoder and/or a quantum decoder, quantum information theory clarifies how much such a process enhances information transmission speed. It is also possible to fix only a modular and discuss the combination of a classical encoder and a quantum decoder, which is called classical-quantum channel coding as Fig. 4.

Fig. 4. Classical-quantum channel coding for optical communication

Here, we remark that the framework of quantum information theory mathematically contains the framework of classical information theory as the commutative special case. This character is contrastive with the fact that quantum Turing machine does not contain the conventional Turing machine as the commutative special case. Hence, when we obtain a new result in quantum information theory and it is still new even in the commutative special case, it is automatically new even in classical information theory. This point is a big advantageous point and became a driving force for latter unexpected theoretical developments.

As an remarkable achievement of the early stage, in 1979, Holevo obtained a partial result for classical-quantum channel coding [10], [11]. However, this research direction has entered the period of stagnation. In 1990s, quantum information theory has entered a new phase and was studied from a new viewpoint. For example, Schumacher introduced the concept of typical sequence in quantum system. This idea brought us new development and enabled us to extend data compression to the quantum setting [13]. Based on this idea, independently, Holevo [12] and Schumacher and Westmoreland [14] proved the classical-quantum channel coding theorem, which had been unsolved at that time.

Unfortunately, a quantum operation in the framework of quantum information theory is not necessarily available in the current technology in general. Hence, these achievements remain in the range of more theoretical results than classical channel coding theorem. However, such theoretical results brought us more practical results latter in a sense.

III. INFORMATION SPECTRUM

The early stage of the actual development of finite block-length study has been started from a completely different motivation as the method of information spectrum by Han and Verdú[15], [17]. The conventional study in information theory usually imposes the independent and identical condition or the memoryless condition to the information source or the channel. However, the information source nor the channel is rarely independent in the actual case and they often have correlation. Hence, information theory had been needed to be adapted for such a situation.

To resolve such a problem, Verdú and Han have succeeded in discussing the optimal performance of several topics in classical information theory including channel coding by using the behavior of the logarithmic likelihood as Fig. 5[16]. However, at that time, they have discussed only the case when the block-length $n$ approaches infinity, and have not studied the case with finite block-length. It is notable that
this study clarified that the analysis of the independent and identical distributed case is reduced to the law of large number. In this way, the method of information spectrum has clarified the mathematical structure of many topics in information theory, which worked as the silent trigger of the future development.

Fig. 5. Structure of information spectrum

Another important contribution of the method of information spectrum is connecting the statistical simple hypothesis testing to many topics in classical information theory [17]. Here, the statistical simple hypothesis testing is the problem to decide which candidate is the true distribution with an asymmetric treatment of two kinds of errors when two candidates of the true distribution are given. In particular, the method of information spectrum has revealed that the performances of data compression and uniform random number generation are given as the behavior of a random variable called the logarithmic likelihood.

In the next stage, Nagaoka and the author have extended the method of information spectrum to the quantum case [18], [19]. In this extension, their contributions not only are the non-commutative extension but also work as the redevelopment of information theory. Especially, they have more deeply clarified the explicit relation between the statistical simple hypothesis testing and the channel coding, which is called dependence test bound in the latter study [19, Remark 15]. In this context, Nagaoka [20] has developed another explicit relation between the statistical simple hypothesis testing and the channel coding, which is called meta converse in the latter study. These two relations clarified the relation between the statistical simple hypothesis testing and the channel coding, which completed the preparation for the next step.

IV. FOLKLORE IN SOURCE CODING

Han [22] tackled the problem of folklore in source coding by using the method of information spectrum, which states that the compressed data up to the entropy rate is the uniform random number. For a certainty, we cannot deny such a possibility because both optimal conversion rates are the entropy $H(P)$. However, such a reason does not works the proof of the statement. Hence, this statement had been regarded as a folklore. Han focused on the normalized divergence as the criterion to measure the difference of the generated random number from the uniform random number, and showed that folklore in source coding is valid by using the method of information spectrum [22]. However, the normalized divergence is too loosed criterion to guarantee the quality of the uniform random number because it is possible to distinguish the generated random number from the true uniform random number even though the random number is considered with this criterion. In particular, when the random number is used for cryptography, we need to employ more rigorous criterion to judge the quality of the uniform random number.

1Unfortunately, due to the page limitation, this paper does not describe the detail derivation. However, the detail discussion is available in Section 4.6 of the book [21].
The most popular criterion is the criterion called the variational distance, which gives the statistical
distinguishability between the true uniform random number and the given random number [23]. That is,
when this criterion takes the value 0, the random number must be the true uniform distribution. Hence,
when we use a random number, we need to accept only a random number passing this criterion. Also,
Han [22] has proved that folklore in source coding is not valid when we adopt the variational distance as
our criterion.

Fig. 6. Asymptotic trade-off relation between errors of data compression and uniform random number generation.

On the other hand, to clarify the incompatibility between the data compression and the uniform random
number generation, the author [24] have developed the theory for finite block-length codes for both
topics. In this analysis, applying the method of information spectrum to the second order as Fig. 5
he has asymptotically expanded the conversion rates of both topics with respect to the block-length \( n \)
as an approximation of the finite block-length analysis for large \( n \). It had been known by the existing
studies that the coefficients of the first order are the entropy in both topics. Although the optimal rates
of both topics are characterized by the logarithmic likelihood, their coefficients of the second order \( \frac{1}{\sqrt{n}} \)
are different because this order corresponds to the central limit theorem and the coefficients depend on
the error probability. Indeed, since the task of the data compression has the opposite direction to that of
uniform random number generation, the second order analysis explicitly clarifies the trade-off relation for
their errors rather than compatibility. That is, when we fix the conversion rate up to the second order \( \frac{1}{\sqrt{n}} \)
whose minimum error of the data compression is \( \epsilon \), the error of uniform random number generation is
greater than \( 1 - \epsilon \). In this way, this analysis clarifies that data compression and uniform random number
generation are incompatible task to each other. The evaluation of the optimal performance up to the
second order coefficient shows the existence of their trade-off relation as well as the approximation of
the finite-length analysis. Since the evaluation for the uniformity of the random number is closely related
to the security evaluation, this type evaluation has been applied to security analysis [25]. This trade-off
relation plays an important role when we use the compressed data as the scramble random variable for
another information [26].

V. QUANTUM CRYPTOGRAPHY

Section II has explained that the ultimate performance of optical communication can be treated as
quantum channel coding. When the communication media has quantum property, it opens up a possibility
to a new communication style that cannot be realized in the existing technology. Quantum cryptography
was proposed by Bennett and Brassard [27] in 1984 as a technology to distribute secure keys by using a
quantum media. When the key is eavesdropped during the distribution, this method enables us to detect the
existence of the eavesdropper with high probability. Hence, this method realizes a secure key distribution,
and is called quantum key distribution (QKD).

However, this method supposes noiseless quantum communication by a single photon. So, the obtained
keys are not necessarily secure when the channel has noise or two photons are transmitted. When the
communication device has such a imperfectness, it is natural to consider that there exists a partial
information leakage according to the amount of the imperfectness. Privacy amplification is a method
to disable such a leaked information. Shor and Preskill \cite{28} and Mayers \cite{29} showed that privacy amplification generates a secure final key even when the channel has noise when the light source correctly generates a single photon. Gottesman et al \cite{30} showed that this final key can be secure even when the light source partially generates a multiple photon when the ratio of a multiple photon is sufficiently small. The light source used in the actual quantum optical communication is the weak coherent light, which partially generates a multiple photon inevitably as Fig. \ref{fig:7}. Hence, this kind of extension had been essentially required for practical use. Hwang \cite{31} and several researchers \cite{32}, \cite{33}, \cite{34}, \cite{35}, \cite{36}, \cite{37} proposed efficient methods to estimate the ratio of multiple photon. These results assume the combination of error correction and privacy amplification of infinitely large block-length $n$. They did not give the quantitative evaluation of the security with finite block-length $n$. They also did not sufficiently discuss the concrete construction of privacy amplification so that their discussions were not sufficient for realization of quantum key distribution system.

![Fig. 7. Multiple photon in weak coherent light](image)

To resolve this issue, the author \cite{38} approximately evaluated the security with finite block-length $n$ when the channel has noise and the light source correctly generates a single photon. This idea has two key points. His first contribution is the evaluation of information leakage via the phase error probability when virtual error correction to phase basis, which has the dual relation to the bit basis. This evaluation is based on the duality relation in quantum theory, which typically appears in the relation between position and momentum. The other contribution is the approximate evaluation of the phase error probability via the application of central limit theorem. This analysis is essentially equivalent to the derivation of the coefficient the transmission rate up to the second order $1/\sqrt{n}$. Later, he exactly derived this evaluation without any approximation with Tsurumaru \cite{39}. After the analysis of the single photon case, he extended his analysis to the case when the light source partially generates a multiple photon \cite{40}, \cite{41}. The transmission rate in the typical case is calculated as Fig. \ref{fig:8}. Further, jointly with Tsurumaru, he developed more efficient privacy amplification with less random seeds for practical use as Fig. \ref{fig:9} \cite{42}, \cite{43}. These series achievements satisfy the minimum requirement for theoretical study to implement quantum key distribution with the current technology.

On the other hand, in 2008, several universities and several companies organized a project for QKD network system (SECOQC project) in Europe. The author belongs to a project for the development of QKD system by the National Institute of Information and Communication Technology (NICT). The NICT project organized a similar project in Tokyo (Tokyo QKD Network) in 2010 \cite{45}. Also, similar projects were organized in Hefei in 2012 and in Jinan in 2013 \cite{44}. In 2013, a US company, Battelle, implemented QKD system for commercial use in Ohio state by using the device for QKD by id Quantique \cite{46}. It has a plan to establish a QKD system between Ohio and Washington DC with distance 700km \cite{46}. Also, Chinese government has a plan to establish a QKD system connecting Shanghai, Hefei, Jinan, and Beijing
with distance 2,000km until December 2016. Indeed, these implemented QKD networks are composed of a collection of QKD communications with relatively short distance. To enhance the security of the total system of these QKD network systems, we need further study of information theoretical design for quantum network.

![Key generation rate](image)

**Fig. 8.** Key generation rate with imperfect photon source: We employ two intensities, signal intensity and decoy intensity. Using the difference between detection rates of the pulses with two different intensities, we can estimate the ratio of the multiple photon in the detected pulses. Here, we set the signal intensity to be 1. This graph shows the key generation rate dependently of the decoy intensity. This graph is based on the calculation formula given in [41].

Here, we should remark the security analysis based on leftover hash lemma [47], [48], [49], [50] as another research stream of quantum key distribution. This method came from cryptography theory and was started by group of Swiss Federal Institute of Technology in Zurich (ETH), whose leader is Renner [51]. The advantage of this method is the direct evaluation of information leakage without discussing the virtual phase error probability. Hence, they do not need to care about the virtual phase error probability. This method also enables the security analysis with finite block-length [52]. However, their finite block-length analysis is loser than the analysis by the author and Tsurumaru [39] because their bound [52] cannot yield the second order rate based on central limit theorem in the limit while it can be recovered with the limit from the bound by the author and Tsurumaru [39]. Further, although their method is rather precise, it has so many parameters to be estimated in the finite block-length analysis. Many parameters to be estimated increase the error of channel estimation. Hence, they need to decrease the number of parameters to be estimated. In their finite block-length analysis, they simplified their analysis so that only the virtual phase error probability to be estimated. Thus, the approach based on leftover hash lemma is better because it gives security evaluation based on the virtual phase error probability more directly. However, the approach based on leftover hash lemma influenced the security analysis in the classical setting latter [87], [88], [89], [70], [90]. Further, this approach did not discuss the security of the case when the light source partially generates a multiple photon.

### VI. SECOND ORDER CHANNEL CODING

Now, we back to the classical channel coding. In the channel coding, it is important to clarify the difference between the asymptotic transmission rate and the actual transmission rate dependently of the block-length as Fig. [10] The author extended the theory of second order analysis to the classical channel.
Fig. 9. Classes of (Dual) Universal\textsubscript{2} hash functions and the Security: Hash function is a function to realize the privacy amplification. This picture shows the relation classes of hash functions and the security. In cryptography theory, it is considered that strong security is requirement for hash function [23]. The papers [49], [50] proposed the class of universal\textsubscript{2} hash functions. Using leftover hash lemma [47], [48], Renner [51] proposed to use this class for quantum cryptography. Tomamichel etal [53] proposed to use the class of \(\epsilon\)-almost universal\textsubscript{2} hash functions when \(\epsilon\) is close to 1. Tsurumaru etal [43] proposed to use the class of \(\epsilon\)-almost dual universal\textsubscript{2} hash functions when \(\epsilon\) is constant or increases polynomially. As an example of \(\epsilon\)-almost dual universal\textsubscript{2} hash function, the author with his collaborators [42] constructed a secure Hash Function with less random seed and less calculation. Although the security analysis in [52] is based on universal\textsubscript{2} hash function, that in [40], [41], [39] is based on \(\epsilon\)-almost dual universal\textsubscript{2} hash function.

In fact, a group in Princeton University, mainly Verdú and Polyanskiy, tackled this problem, indepen-
In their papers [56], [57], they considered the relation between channel coding and the statistical simple hypothesis testing and independently derived two relations so called dependence test bound and meta converse, which are the same as the classical special case of the author and Nagaoka [19] and Nagaoka [20]. Since their results [56] are limited to the classical case, the applicable region of their results are more narrow than the preceding results in [19], [20]. Then, they rederived Strassen’s result without direct use of the method of information spectrum by direct evaluation these two bounds. They also independently derive the second order coefficient of the optimal transmission rate for the additive white Gaussian noise channel the paper [56] in 2010. Since the result by Princeton group called a large impact in the community of information theory at that time, their paper has received the best paper award of IEEE Information theory society in 2011 jointly with the preceding paper by the author [54].

As explained the above, Japanese group obtained some of the same results prior several years over Princeton group. However, the outgoing force for their own result is much weaker than Princeton group. Princeton group caught up the demand in the community of information theory and present their result in very effectively. In contrast to Princeton group, Japanese group studied the same topic far from the above demand. Since the research activity of Princeton group is limited to the community of information theory, their target of audiences is suitably concentrated. Hence, they could succeed in creating a scientific boom of this direction. Since Japanese group was studied in quantum information theory, they had to play an active role in the community of quantum information theory so that they presented their result to the audience that are less interested in their results. Hence, such a research direction has been discouraged. Since the Japanese group is related to so wide research area, as a result, they could not appeal their own result effectively. As their research fund comes from quantum information, they cannot pay their effort to appeal their own result to the community of classical information theory. In particular, the organization the author belongs to has no researcher to understand this research direction. Hence, almost no Japanese researcher knew this paper when his paper received the best paper award of IEEE Information theory society in 2011.

When the award is awarded to these two papers, they have less citation. After this award, this research direction becomes much more popular and applied to so many topics in information theory [58], [59], [61], [66], [60], [63], [64], [70]. In particular, the third order analysis has been done for the channel coding [67]. These activities were reviewed in the book [68].

However, these studies consider the finite block-length analysis for the optimal rate with respect to all codes that including too higher calculation complexity to implement. Hence, the obtained rate is not necessarily realized with implementable codes. To resolve this issue, we need to discuss the optimal rate among codes whose calculation complexity is not so high. No existing study discusses such a type of finite block-length analysis. So, such a study is strongly appreciated as a future study.

After this breakthrough, Princeton group extended their idea to many topics in channel coding and data compression [58], [59], [60], [62]. On the other hand, jointly with Watanabe, the author extended his result to the channels with additive Markovian noise [63]. He also, jointly with Tomamichel, extended his result to the quantum system [65]. In the paper [65], they provided a unified framework of the second order theory in the quantum system for data compression with side information, secure uniform random number generation, and simple hypothesis testing. In the same time, Li [69] directly derived the second order analysis only for statistical simple hypothesis testing in the quantum case. However, the second order theory for statistical simple hypothesis testing has less meaning for itself. It is more meaningful for the relation with other topics in information theory.

VII. EXTENSION TO PHYSICAL LAYER SECURITY

The above explained quantum cryptography offers secure key distribution based on physical law. Hence, we cannot assume other reasonable assumption so that the total process becomes very complicated. On the other hand, the classical counterpart of the quantum cryptography is physical layer security, which offers information theoretical security based on several physical assumption with classical mechanics. This
area was formulated as wire-tap channel model by Wyner [71], and has been investigated more deeply by Csiszár and Kőnner[72]. This model assumes two channels, a channel from the authorized sender (Alice) to the authorized receiver (Bob) and a channel from the authorized sender to the eavesdropper (Eve). When the original signal of Alice has stronger correlation with the received signal than that with Eve, the authorized users can communicate without any information leakage. In this case, we need to assume that the channel to Eve is known. We often assume that the channel is stationary. Although this model is a typical model for information theoretic security, it is somehow unnatural. So, it is quite hard to find a realistic situation that the model can be applied to. Fortunately, this model has a more realistic derivative, which is called secure network coding [75], [76], [77], [78], [79], [80]. Secure network coding guarantees the security when the information is sent via plural paths and only a part of them are eavesdropped. It cannot guarantee the security when all of paths are eavesdropped by Eve. However, since it is quite hard to tap all of the paths, this kind of security is sufficiently useful for practical use of ordinary people as cost-benefit performance.

As another type of information theoretical security, Ahlswede and Csiszár[74] and Maurer[73] proposed secure key distillation. When the mutual information between two authorized users is larger than that between one authorized user and Eve, they can extract secure final key. Quantum cryptography has the same operation after quantum communication. Since we cannot assume any assumption except for basic law of quantum theory in quantum cryptography, quantum cryptography requires so many complicated treatments. Secure key distillation with physical layer security does not need to consider such a delicate problem so that it offers much more tighter evaluation because attacks considered for physical layer security are more limited than those for quantum cryptography due to the difference between their purposes. As the most biggest difference, secure key distillation can often assume the independent and identical distributed condition, but quantum cryptography cannot assume this condition usually.

![Fig. 11. Secure wireless communication with Eve’s attack](image)

Recently, depending on the promotion of the use of wireless communication, secure wireless communication has been very actively studied [82], [83], [84], [85]. In the case of wireless communication, the simple application of wire-tap channel cannot guarantee secure communication when Eve put the antenna between the Alice and Bob. However, when the noise in Bob’s output signal is independent of the noise in Eve’s output signal, the mutual information between Alice and Bob is larger than that between Eve and Bob. So, they can generate secure keys. Even though they are correlated, if it is restricted in a sense as Fig. 11 we have the same conclusion [81]. For example, we can naturally assume that the noises generated inside the detector are independent and only the noises generated outside the detector are correlated. While such a constraint does not necessarily hold in general, it is practical under cost-benefit performance.

Here, we summarize the advantage over the modern cryptography based on computation complexity [91]. When the modern cryptography based on computation complexity is broken by a computer, any information

---

2 In the study of cryptography, We call the authorized sender, the authorized receiver, and the eavesdropper Alice, Bob, and Eve, respectively.
transmission with this cryptography will be eavesdropped by using the computer. To eavesdrop the physical layer security with the above type, Eve has to prepare an antenna for each communication. Further, the antenna must be very expensive because it has to break the above assumption. Maybe, it is not impossible to break a limited number of specific communications for very limited persons. However, due to the cost constraint, it is impossible to eavesdrop all of communications. In this way, the physical layer security offers a different type of security from the computational security.

Recently, security analysis for physical layer security has been developed by importing the methods developed or motivated in quantum cryptography [86], [87], [88], [89], [70], [90], [25], [92]. In particular, the finite block-length analysis has been much developed for physical layer security including the Markovian case [87], [88], [89], [70], [90]. The hash function originally constructed for quantum cryptography can be used for privacy amplification process even in physical layer security [49], [50], [42]. Since we are allowed to assume several natural assumptions for physical layer security, physical layer security has much advanced security analysis. Its finite block-length analysis is different from that for channel coding in the following point. The obtained finite block-length analysis for channel coding discusses only the optimal performance among all codes including impractical code whose calculation complexity is too high. However, in the finite block-length analysis for physical layer security, the obtained bound can be attained by practical protocol whose calculation complexity is linear for the block-length. In this sense, the finite block-length analysis for physical layer security is almost completed. The remaining task is how to apply the obtained security evaluation to real communication systems.

VIII. CONCLUSIONS AND FUTURE PROSPECTS

In this review article, we have discussed the developments of finite block-length theory in classical and quantum information theory: classical and quantum channel coding, data compression, (secure) random number generation, quantum cryptography, and physical layer security. These subareas have been developed with strong interactions with each other in unexpected ways.

The required future studies for channel coding and data compression are completely different from those for security topics. In the former topics, existing finite block-length theory discusses only the minimum error among all codes without discussion for the calculation complexity of its construction. Hence, for practical use, we need a finite block-length theory for realizable codes whose construction has less calculation complexity. Such a type of finite block-length theory is strongly required.

In contrast, in the latter topics, finite block-length theory has been already established with taking account into the calculation complexity of its construction. Hence, the obtained finite block-length theory is more practical. However, these types of security protocols are not realized. In the case of quantum cryptography, we need more development for device side. Also, to realize secure communication over 200km, we need another type of information-scientific combinatorics. In the case of physical layer security, we need more studies to fill the gap between information theoretical security analysis and the device developments. The recent study [81] is a candidate of such a study.

Furthermore, the idea of finite block-length theory is fundamental and can be extended to other areas beyond information theory. For example, it has been applied to a statistical mechanical rederivation of thermodynamics [93], [94], conversion of entangled states [95], [96], [97], [98], the analysis of the gap between two classes of local operations [99]. Therefore, we can expect more applications of finite block-length theory to other areas.

ACKNOWLEDGMENTS

The works reported here were supported in part by a MEXT Grant-in-Aid for Scientific Research (A) No. 23246071, a MEXT Grant-in-Aid for Young Scientists (A) No. 20686026, a MEXT Grant-in-Aid for Scientific Research on Priority Area “Deepening and Expansion of Statistical Mechanical Informatics (DEX-SMI)” No. 18079014, the National Institute of Information and Communication Technology (NICT), Japan, the Okawa Research Grant, and the Kayamori Foundation of Informational Science Advancement.
REFERENCES

[1] C. E. Shannon, “A mathematical theory of communication,” Bell Syst. Tech. J., 27, 623–656 (1948).
[2] R. G. Gallager, Information Theory and Reliable Communication. New York: Wiley, 1968.
[3] C. Berrou, A. Glavieux, and P. Thitimajshima, “Near Shannon Limit Error-correcting Coding and Decoding: Turbo-codes 1,” Communications, 1993. ICC ’93 Geneva. Technical Program, Conference Record, IEEE International Conference on (Volume:2) 1064 - 1070
[4] D. J.C. MacKay and R. M. Neal, “Near Shannon Limit Performance of Low Density Parity Check Codes,” Electronics Letters, July 1996
[5] V. Strassen, “Asymptotische Abschätzungen in Shannon’s Informationstheorie,” In Transactions of the Third Prague Conference on Information Theory etc, Czechoslovak Academy of Sciences, Prague, pp. 689-723, 1962. (In German); English translation in 2009 is available at [http://www.math.cornell.edu/~pmlib/strassen.pdf](http://www.math.cornell.edu/~pmlib/strassen.pdf)
[6] H. Takahasi Information theory of quantum-mechanical channels Advan. Commun. Systems, 1 (1965), pp. 227-310
[7] C. W. Helstrom, “Detection theory and quantum mechanics,” Inf. Contr., 10, 254–291 (1967).
[8] C. W. Helstrom, “Minimum-mean-square error quantum statistics,” Phys. Lett., 25A, 101–102 (1967).
[9] A. S. Holevo, Probabilistic and Statistical Aspects of Quantum Theory, (North-Holland, Amsterdam, 1982); originally published in Russian (1980).
[10] A. S. Holevo, “Bounds for the quantity of information transmitted by a quantum communication channel,” Problemy Peredachi Informatsii, 9, 3–11 (1973) (in Russian). (English translation: Probl. Inf. Transm., 9, 177–183 (1975)).
[11] A. S. Holevo, “On the capacity of quantum communication channel,” Problemy Peredachi Informatsii, 15, 4, 3–11 (1979) (in Russian). (English translation: Probl. Inf. Transm., 15, 247–253 (1979)).
[12] A. S. Holevo, “The capacity of the quantum channel with general signal states,” IEEE Trans. Inf. Theory, 44, 269 (1998).
[13] B. Schumacher, “Quantum coding,” Phys. Rev. A, 51, 2738–2747 (1995).
[14] B. Schumacher, M. D. Westmoreland, “Sending classical information via noisy quantum channels,” Phys. Rev. A, 56, 131 (1997).
[15] T. S. Han and S. Verdú, “A general formula for channel capacity,” IEEE Trans. Inf. Theory, 40, 1147–1157 1994.
[16] T. S. Han, Information-Spectrum Methods in Information Theory, (Springer, Berlin, 2003). (Original Japanese version was published from Baifukan in 1998)
[17] H. Nagaoka and M. Hayashi, “An information-spectrum approach to classical and quantum hypothesis testing,” IEEE Trans. Inf. Theory, 53, 534–549 (2007).
[18] M. Hayashi and H. Nagaoka: “General formulas for capacity of classical-quantum channels,” IEEE Trans. Inf. Theory, 49, 1753–1768 (2003).
[19] H. Nagaoka, “Strong converse theorems in quantum information theory,” Proc. ERATO Conference on Quantum Information Science (EQIS) 2001, 33 (2001). (also appeared as Chap. 3 of Asymptotic Theory of Quantum Statistical Inference, M. Hayashi eds.).
[20] M. Hayashi, Quantum Information: An Introduction, Springer 2006. (Original Japanese version was published from Saizensu-sha in 2004)
[21] T.S. Han, “Folklore in source coding: information-spectrum approach,” IEEE Transactions on Information Theory (Volume:51, Issue: 2) 747 - 753 (2005)
[22] R. Renner, and R. Koenig, “Universally composable privacy amplification against quantum adversaries.” Theory of Cryptography: Second Theory of Cryptography Conference, TCC 2005, J.Kilian (ed.) Springer Verlag 2005, vol. 3378 of Lecture Notes in Computer Science, pp. 407–425.
[23] M. Hayashi, “Second-Order Asymptotics in Fixed-Length Source Coding and Intrinsic Randomness,” IEEE Trans. Inform. Theory, vol. 54, no. 10, 4619 – 4637, 2008.
[24] S. Watanabe, R. Matsumoto, and T. Uyematsu “Strongly Secure Privacy Amplification Cannot Be Obtained by Encoder of Slepian-Wolf Code,” IEICE Trans. Fundamentals, vol. EA-93, no. 9, 1650-1659, 2010.
[25] M. Hayashi, R. Matsumoto, “Secure Multiplex Coding with Dependent and Non-Uniform Multiple Messages,” Accepted for publication in IEEE Transactions on Information Theory.
[26] C. H. Bennett, G. Brassard, “Quantum cryptography: public key distribution and coin tossing,” Proc. IEEE International Conference on Computers, Systems and Signal Processing, (Bangalore, India), pp. 175–179 (1984).
[27] P. W. Shor, J. Preskill, “Simple proof of security of the BB84 quantum key distribution protocol,” Phys. Rev. Lett., 85, 441–444 (2000).
[28] M. Hayashi, “Unconditional security in Quantum Cryptography,” Journal of the ACM 48 351 (2001)
[29] D. Gottesman, H.-K. Lo, N. Litkeinen, and J. Preskill, “Security of quantum key distribution with imperfect devices,” Quant. Inf. Comput. 5, pp.325-360 (2004).
[30] W.-Y. Hwang, “Quantum Key Distribution with High Loss: Toward Global Secure Communication” Phys. Rev. Lett. 91 057901 (2003)
[31] H.-K. Lo, X. Ma, and K. Chen “Decoy State Quantum Key Distribution,” Phys. Rev. Lett. 94 230504 (2005)
[32] X.-B. Wang, “Beating the Photon-Number-Splitting Attack in Practical Quantum Cryptography,” Phys. Rev. Lett. 94 230503 (2005)
[33] X.-F. Ma, B. Qi, Y. Zhao, and H.-K. Lo, “Practical decoy state for quantum key distribution,” Phys. Rev. A 72 012326 (2005)
[34] X.-B. Wang, “A decoy-state protocol for quantum cryptography with four intensities of coherent states,” Phys. Rev. A 72 012322 (2005)
[35] M. Hayashi, “General theory for decoy-state quantum key distribution with an arbitrary number of intensities,” New J. Phys. 9 284 (2007).
[36] T. Tsurumaru, A. Soujaeff, and S. Takeuchi “Exact minimum and maximum of yield with a finite number of decoy light intensities,” Phys. Rev. A 77 022319 (2008)
[37] M. Hayashi, “Practical Evaluation of Security for Quantum Key Distribution,” Phys. Rev. A, vol. 74, 022307, 2006.
[38] M. Hayashi and T. Tsurumaru “Concise and Tight Security Analysis of the Bennett-Brassard 1984 Protocol with Finite Key Lengths,” New Journal of Physics, Vol. 14, 093014, (2012),
[77] N. Cai and R. W. Yeung, “Secure network coding,” in Proc. 2002 IEEE ISIT, Lausanne, Switzerland, Jul. 2002, p. 323. [Online]. Available: http://iest2.ie.cuhk.edu.hk/~whyeung/publications/secure.pdf

[78] ——, “A security condition for multi-source linear network coding,” in Proc. 2007 IEEE ISIT, Nice, France, Jun. 2007, pp. 561–565.

[79] ——, “Secure network coding on a wiretap network,” IEEE Trans. Inform. Theory, vol. 57, no. 1, pp. 424–435, Jan. 2011.

[80] R. Matsumoto and M. Hayashi, “Secure Multiplex Network Coding” Proceedings of 2011 International Symposium on Network Coding (NetCod), Beijing, China, 25–27 July 2011. (DOI: 10.1109/ISNEXTCOD.2011.5979076.)

[81] M. Hayashi, “Security guaranteed wireless communication over common noise attack under spatial assumption,” http://arxiv.org/abs/1604.00635 (2016).

[82] C. Ye, S. Mathur, A. Reznik, Y. Shah, W. Trappe, and N.B. Mandayam, “Information-theoretically secret key generation for fading wireless channels,” IEEE Transactions on Information Forensics and Security, 5 (2), 240-254 (2010)

[83] J.W. Wallace and R. K. Sharma, Automatic secret keys from reciprocal MIMO wireless channels: Measurement and analysis IEEE Transactions on Information Forensics and Security, 5 (3) 381 - 392 (2010)

[84] S. N. Premnath, S. Jana, J. Croft, and P.L. Gowda “Secret key extraction from wireless signal strength in real environments,” IEEE Transactions on Mobile Computing (Volume:12, Issue: 5) 917 - 930 (2013).

[85] C. Chen and M.A. Jensen, “Secret key establishment using temporally and spatially correlated wireless channel coefficients,” IEEE Transactions on Mobile Computing (Volume:10, Issue: 2) 205 - 215 (2011)

[86] M. Hayashi, “General non-asymptotic and asymptotic formulas in channel resolvability and identification capacity and its application to wire-tap channel,” IEEE Trans. Inform. Theory, vol. 52, no. 4, 1562–1575, 2006.

[87] M. Hayashi, “Exponential decreasing rate of leaked information in universal random privacy amplification,” IEEE Trans. Inform. Theory, vol. 57, no. 6, 3989–4001, 2011.

[88] M. Hayashi, “Tight exponential analysis of universally composable privacy amplification and its applications,” IEEE Trans. Inform. Theory, vol. 59, No. 11, 7728 – 7746, 2013.

[89] W. Kumagai and M. Hayashi, “Tight asymptotic bounds on local hypothesis testing between a pure bipartite state and the white noise state,” Proceedings of IEEE International Symposium on Information Theory (ISIT2015), Hong Kong, June 14 - June 19, 2015. pp. 691-695