Coding Theorems on the Threshold Scheme for a General Source

Hiroki Koga, *Member, IEEE*

***Abstract*—In this paper, coding theorems on the (t, m)-threshold scheme for a general source are discussed, where m means the number of the shares and t means a threshold. The (t, m)-threshold scheme treated in this paper encrypts n source outputs X^n to m shares at once and is required to satisfy the two conditions that 1) X^n is reproduced from arbitrary t shares, and 2) almost no information of X^n is revealed from any $t-1$ shares. It is shown that the (t, m)-threshold scheme must satisfy certain inequalities including the limit inferiors in probability. One of the inequalities is closely related to the minimum length of the fair random bits needed to a dealer for realizing the (t, m)-threshold scheme. In addition, it is shown that a certain variation of Shamir's threshold scheme meets the two conditions. The same approach can be taken to the problems of Shannon's cipher system with the perfect secrecy and fixed-length source coding with vanishing decoding error probability. It is shown that the same kind of inequalities, which indicate the converse coding theorems, hold in both two cases.**

***Index Terms*—Fixed-length source coding, information-spectrum methods, secret sharing scheme, Shannon's cipher system, threshold scheme.**

I. Introduction

A secret sharing scheme, which was independently proposed by Shamir [11] and Blakley [1] in 1979, provides a method for keeping a secret information securely by encryption. In particular, the threshold scheme [11] is one of important secret sharing schemes. In the threshold scheme with m (≥ 2) participants, a dealer encrypts a secret information X to m shares $W^{(1)}, \ldots, W^{(m)}$ and distributes $W^{(i)}$ to the ith participant for $i = 1, 2, \ldots, m$. The threshold scheme has a remarkable property that, letting t be a threshold satisfying $2 \leq t \leq m$, X is reproduced from an arbitrary collection of t shares while no information on X is obtained from any less than t shares. We call the threshold scheme with m participants and a threshold t the *(t, m)-threshold scheme*.

In studies of secret sharing schemes the sizes of shares are often analyzed. In particular, [8] gives a basic result on the (t, m)-threshold scheme that $H(W^{(i)}) \geq H(X)$ for all $i = 1, 2, \ldots, m$, where $H(\cdot)$ denotes the entropy. The sizes of shares are also evaluated in secret sharing schemes with general access structures (e.g., [3]). Randomness needed to a dealer for realizing the (t, m)-threshold scheme is discussed by Blundo *et al.* [2]. It is shown in [2] that the randomness, which is defined as the conditional entropy $H(W^{(1)}W^{(2)} \cdots W^{(m)}|X)$, is lower bounded by $(t-1) \log |\mathcal{X}|$, where $\mathcal{X}$ is the alphabet of X and $|\mathcal{X}|$ denotes its cardinality. On the other hand, from a Shannon-theoretic viewpoint, [15] treats a communication system which can be regarded as the threshold scheme with two or three shares in a certain case and gives an achievable region for the rates of the shares. A generalization of the results in [15] is discussed in [10]. The rate for the uniform random number that is available to a dealer is also analyzed in [10] for the cases of $m = 2$ and $m = 3$.

In this paper, we unveil certain fundamental properties of the (t, m)-threshold scheme from a viewpoint of information-spectrum methods. Information-spectrum methods, which originate from [7] and are described in detail in [6], provide methods to treat coding of vast classes of sources and channels. In fact, in information-spectrum methods, instead of the ordinary entropy and mutual information, quantities defined by using the limit inferior or superior *in probability* play key roles such as the channel capacity [14]. In this paper we are interested in the minimum rate of the uniform random number needed to a dealer for realizing the (t, m)-threshold scheme. We consider the situation where an n-tuple of secrets X^n is generated from a general source and is encrypted to m shares $W_n^{(1)}, W_n^{(2)}, \ldots, W_n^{(m)}$. Here, the class of general sources includes all the classes of sources such as stationary memoryless sources, stationary ergodic sources, stationary sources and even nonstationary/nonergodic sources [6]. In addition, the size of the alphabet $\mathcal{X}$ can be countably infinite.

The (t, m)-threshold scheme considered in this paper can cause decoding error. That is, we consider the situation where an arbitrary collection of t shares $W_n^{(i_1)}, W_n^{(i_2)}, \ldots, W_n^{(i_t)}$ does not always reproduce X^n. However, we require that the probability of such decoding error vanishes as $n \to \infty$ for all $\{i_1, i_2, \ldots, i_t\} \subset \{1, 2, \ldots, m\}$. We introduce the vanishing decoding error probability so that we can obtain meaningful asymptotic results as $n \to \infty$. Recall that in problems of two-terminal and multi-terminal source coding we often introduce the vanishing decoding error probability and obtain fundamental bounds on the rates of codes that are asymptotically attainable and are described by using in terms of information-theoretic quantities. As a byproduct, this setting enables us to discuss construction of the (t, m)-threshold scheme for sources with a countably infinite alphabet.

Manuscript received February 16, 2007; revised November 29, 2007. This work is supported in part by Grants-in-Aid for Scientific Research for Young Scientists No. 17700005 from Ministry of Education, Culture, Sports, Science and Technology in Japan. The material in this paper was presented in part at 2005 IEEE Information Theory Workshop on Theory and Practice in Information Theoretic Security, Awaji Island, Japan, October 2005.

The author is with Graduate School of Systems and Information Engineering, University of Tsukuba, Ibaraki 305-8573, Japan (e-mail: koga@esys.tsukuba.ac.jp).

Communicated by H. Imai, Guest Editor for Special Issue on Information Theoretic Security.

Digital Object Identifier 10.1109/TIT.2008.921860

We impose a nonconventional security criterion on the (t, m)-threshold scheme. We require that (t, m)-threshold scheme must satisfy a security criterion such that almost no information of X^n is revealed from any collection of less than t shares. Usually, such a criterion is described in terms of the mutual information, say $I(X^n; W_n^{(i_1)} \cdots W_n^{(i_{t-1})}) = 0$. However, the criterion considered in this paper is written as an inequality including the limit superior in probability. Roughly speaking, the criterion requires that with probability close to one X^n is almost independent of $W_n^{(i_1)}, W_n^{(i_2)}, \ldots, W_n^{(i_{t-1})}$ provided that n is sufficiently large.

We first prove that any sequence of encoders and decoders realizing the (t, m)-threshold scheme must satisfy certain inequalities including the limit inferior in probability. The inequalities are closely related to the rates of j shares $(1 \leq j \leq t)$ and the uniform random number needed to a dealer. In particular, it is shown that the rate of the uniform random number is lower-bounded by $(t-1)H(X)$ in the asymptotic sense as $n \to \infty$ for the case of a stationary memoryless source with the entropy $H(X) < \infty$.

Next, we give a construction of the (t, m)-threshold scheme. It is shown that a certain variation of Shamir's threshold scheme [11] meets the requirements as the (t, m)-threshold scheme under a certain assumption. We can also prove that the construction satisfies $I(X^n; W_n^{(i_1)} \cdots W_n^{(i_t)}) \to 0$ as $n \to \infty$, where $I(\cdot;\cdot)$ denotes the mutual information. Note that this result does not conflict with [4] showing the nonexistence of the (t, m)-threshold scheme for a source with a countably infinite alphabet. In fact, while [4] does not permit decoding error, we permit negligible decoding error probability. Permitting negligible decoding error probability enables us to realize the (t, m)-threshold scheme for sources with countably infinite alphabets.

We can treat Shannon's cipher system with the perfect secrecy from the same viewpoint as the above (t, m)-threshold scheme [12], [16], [17]. We consider the following setting. Given an n-tuple of outputs X^n from a general source, an encoder encrypts X^n to a cryptogram W_n under a key E_n. A decoder, which shares a key E_n with the encoder, decrypts the cryptogram W_n to $\hat{X}^n$ under E_n. The encoder and decoder must satisfy the two condition that 1) the decoding error probability vanishes as $n \to \infty$, and 2) W_n is almost independent of X^n. We give two fundamental inequalities including the limit inferior in probability that Shannon's cipher system with the perfect secrecy must satisfy. It should be noted that the two inequalities hold for the class of stochastic encoders. Since the encoder using homophonic coding (e.g., [9]) is regarded as a stochastic encoder, this kind of extension is meaningful. The two inequalities suggest the converse theorem on the rates of the cryptogram and the key as a byproduct.

We can take the same approach to the fixed-length coding as well. We consider the following setting. Given an n-tuple of outputs X^n from a general source, an encoder encodes X^n to a codeword W_n. The encoder can be stochastic. On the other hand, a decoder decodes the codeword W_n to $\hat{X}^n$ by using a deterministic mapping. We require that the decoding error probability vanishes as $n \to \infty$. Actually, the problem of the fixed-length source coding is a special case of Shannon's cipher system where E_n is a constant and we do not care the above condition 2). In this case as well, we can obtain an inequality including the limit inferior in probability that characterizes a fundamental relationship between X^n and W_n. We also give a general formula of the infimum-achievable coding rate for stochastic encoders. The formula coincides with the formula in [7] that treats only deterministic encoders.

The organization of this paper is as follows. In Section II we formulate the problem of the (t, m)-threshold scheme with introducing notations. Section III is devoted to description of results for the case of a stationary memoryless source with a finite alphabet. Results in Section III is expressed in terms of the entropy and the mutual information. Main results of this paper are stated in Section IV. After formulating the (t, m)-threshold scheme by using terminologies of information-spectrum methods, we give the inequalities that the (t, m)-threshold scheme must satisfy. In Section V we give a construction of the (t, m)-threshold scheme that meets the conditions given in Section IV. Results on Shannon's cipher system and fixed-length source coding, which are obtained as byproducts of the methods developed in Sections IV and V, are given in Sections VI and VII, respectively.

II. Preliminaries

Let $\mathcal{X}$ be a finite or a countably infinite source alphabet. For each $n \geq 1$ denote by $\mathcal{X}^n$ the nth Cartesian product of $\mathcal{X}$. Let X^n be a random variable that takes values in $\mathcal{X}^n$, where X^n means n source outputs. The probability distribution of X^n is denoted by P_{X^n}. In addition, the probability that x^n is generated from the source is denoted by $P_{X^n}(x^n)$. We express X^n, $n \geq 1$, in the form of $\boldsymbol{X} = \{X^n\}_{n=1}^{\infty}$ and call $\boldsymbol{X}$ a *general source*. The class of general sources includes all classes of sources such as memoryless sources, stationary ergodic sources, stationary sources, and even nonstationary/nonergodic sources. In particular, if $\boldsymbol{X}$ is stationary and memoryless, it holds that $P_{X^n}(x^n) = \prod_{i=1}^{n} P_X(x_i)$ for all $n \geq 1$ and $x^n \in \mathcal{X}^n$, where $x^n = (x_1, x_2, \ldots, x_n)$ and P_X is a probability distribution on $\mathcal{X}$,

For each $n \geq 1$ let E_n denote an output from a random number generator taking values in a finite set $\mathcal{E}_n$. The set $\mathcal{E}_n$ may not be a Cartesian product, but is dependent on n. If E_n is subject to the uniform distribution on $\mathcal{E}_n$, the probability distribution P_{E_n} of E_n satisfies $P_{E_n}(e_n) = 1/|\mathcal{E}_n|$ for all $e_n \in \mathcal{E}_n$, where $|\cdot|$ denotes the cardinality of the set. Assume that E_n is independent of X^n for each $n \geq 1$.

Throughout the paper let $\mathcal{P} = \{1, 2, \ldots, m\}$ be a set of m participants and t a threshold satisfying $2 \leq t \leq m$. We assume that m and t are fixed integers that do not depend on n. For each $i = 1, \ldots, m$ let $\mathcal{W}_n^{(i)}$ be a finite set in which a share distributed to participant i takes values. We define an encoder as a deterministic mapping $f_n : \mathcal{X}^n \times \mathcal{E}_n \to \mathcal{W}_n^{(1)} \times \cdots \times \mathcal{W}_n^{(m)}$. A dealer generates m shares $(W_n^{(1)}, \ldots, W_n^{(m)})$ by

$$(W_n^{(1)}, \ldots, W_n^{(m)}) = f_n(X^n, E_n)$$

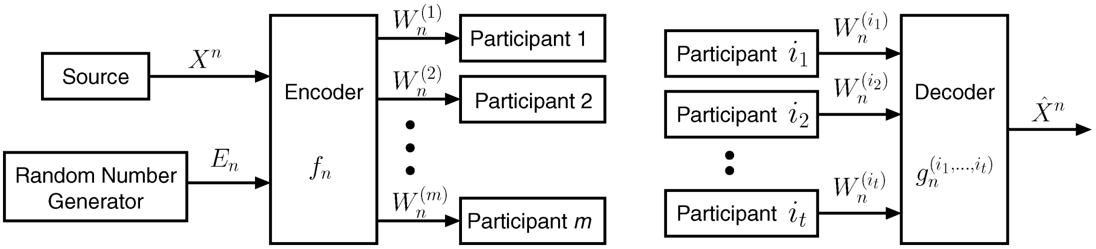

Fig. 1. Encoding and decoding in the (t,m)-threshold scheme.

where $W_n^{(i)} \in \mathcal{W}_n^{(i)}$ means a share which is securely distributed to participant i (see Fig. 1). For the sake of notational convenience, for any $1 \le j \le m$ and $\{i_1,\ldots,i_j\} \subset \mathcal{P}$, we use $W_n^{(i_1,\ldots,i_j)}$ and $\mathcal{W}_n^{(i_1,\ldots,i_j)}$ instead of $(W_n^{(i_1)},\ldots,W_n^{(i_j)})$ and $\mathcal{W}_n^{(i_1)}\times\cdots\times\mathcal{W}_n^{(i_j)}$, respectively. Here, $\{i_1,\ldots,i_j\} \subset \mathcal{P}$ means an arbitrary subset of $\mathcal{P}$ with j elements satisfying $i_1 < \cdots < i_j$. In addition, restriction of $f_n(X^n,E_n)$ to the i_1th, …, i_jth components is denoted by $f_n^{(i_1,\ldots,i_j)}(X^n,E_n)$. That is

$$W_n^{(i_1,\ldots,i_j)} = f_n^{(i_1,\ldots,i_j)}(X^n,E_n).$$

In the (t,m)-threshold scheme, the source output X^n is reproduced from arbitrary t shares $W_n^{(i_1,\ldots,i_t)}$. This means that we need to consider $\binom{m}{t}$ decoders depending on $\{i_1,\ldots,i_t\} \subset \mathcal{P}$ in general. We define a decoder for $W_n^{(i_1,\ldots,i_t)}$ as a deterministic mapping $g_n^{(i_1,\ldots,i_t)} : \mathcal{W}_n^{(i_1,\ldots,i_t)} \to \mathcal{X}^n$ (see Fig. 1). If there is no confusion, we use the term a decoder g_n in the sense of the collection of all $g_n^{(i_1,\ldots,i_t)}$, $\{i_1,\ldots,i_t\} \subset \mathcal{P}$.

We permit decoding error in the (t,m)-threshold scheme. The decoding error probability caused by the encoder f_n and the decoder $g_n^{(i_1,\ldots,i_t)}$ is written as

$$\varepsilon_n^{(i_1,\ldots,i_t)} = \Pr\{g_n^{(i_1,\ldots,i_t)}(f_n^{(i_1,\ldots,i_t)}(X^n,E_n)) \neq X^n\}.$$

Here, throughout this paper $\Pr\{\cdot\}$ denotes the (joint) probability with respect to the random variable(s) contained between the parentheses. In the ordinary setting (e.g., [8], [11]) of the (t,m)-threshold scheme, the decoding error probability is assumed to be equal to zero. In this paper, however, we mainly consider the case where $\varepsilon_n^{(i_1,\ldots,i_t)} \to 0$ as $n \to \infty$ for all $\{i_1,\ldots,i_t\} \subset \mathcal{P}$. This requirement on the decoding error probability makes the problem of the (t,m)-threshold scheme more Shannon-theoretic. We will investigate properties of the (t,m)-threshold scheme that have been unknown so far, say the minimum sizes of shares and construction of the (t,m)-threshold scheme for sources with countably infinite alphabet under this weekended requirement on the decoding error probability.

We impose a condition on $f_n, n \ge 1$, such that any collection of less than t shares reveals almost no information on X^n. In fact, we impose a condition that X^n is almost independent of $W_n^{(i_1,\ldots,i_{t-1})}$ for all $\{i_1,\ldots,i_{t-1}\} \subset \mathcal{P}$. We will give two of such conditions in Sections III and IV.

We define the entropy, the conditional entropy and the mutual information in the ordinary sense [5]. All the logarithms are to the base 2. We define the limit inferior and the limit superior *in probability* according to [6]. That is, for a sequence of real-valued random variables $\boldsymbol{U} = \{U_n\}_{n=1}^{\infty}$ we define

$$\operatorname*{p-lim\,inf}_{n\to\infty} U_n = \sup\left\{\alpha : \lim_{n\to\infty} \Pr\{U_n \ge \alpha\} = 1\right\} \quad (1)$$

$$\operatorname*{p-lim\,sup}_{n\to\infty} U_n = \inf\left\{\beta : \lim_{n\to\infty} \Pr\{U_n \le \beta\} = 1\right\}. \quad (2)$$

In particular, given a general source $\boldsymbol{X} = \{X^n\}_{n=1}^{\infty}$, we define

$$\underline{H}(\boldsymbol{X}) = \operatorname*{p-lim\,inf}_{n\to\infty} \frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} \quad (3)$$

$$\overline{H}(\boldsymbol{X}) = \operatorname*{p-lim\,sup}_{n\to\infty} \frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} \quad (4)$$

which are called the spectral inf-entropy rate and the spectral sup-entropy rate respectively. It is known that $\overline{H}(\boldsymbol{X})$ has the operational meaning as the infimum achievable fixed-length coding rate for the case that the decoding error probability vanishes as $n \to \infty$ [7]. On the other hand, $\underline{H}(\boldsymbol{X})$ means the supremum achievable rate of the uniform random number in the intrinsic randomness problem [13]. If $\underline{H}(\boldsymbol{X}) = \overline{H}(\boldsymbol{X}) \overset{\text{def}}{=} H(\boldsymbol{X})$, then $\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)}$ converges in probability to $H(\boldsymbol{X})$. In particular, it follows from the law of large numbers that $H(\boldsymbol{X}) = H(X)$ for a stationary memoryless source satisfying $H(X) < \infty$, where $H(X)$ denotes the entropy of the source. See [6] for more details.

It is known that for any sequences $\boldsymbol{U} = \{U_n\}_{n=1}^{\infty}$ and $\boldsymbol{V} = \{V_n\}_{n=1}^{\infty}$ of real-valued random variables it holds that

$$\operatorname*{p-lim\,sup}_{n\to\infty} (-U_n) = -\operatorname*{p-lim\,inf}_{n\to\infty} U_n, \quad (5)$$

$$\operatorname*{p-lim\,inf}_{n\to\infty} (U_n + V_n) \ge \operatorname*{p-lim\,inf}_{n\to\infty} U_n + \operatorname*{p-lim\,inf}_{n\to\infty} V_n, \quad (6)$$

$$\operatorname*{p-lim\,inf}_{n\to\infty} (U_n + V_n) \le \operatorname*{p-lim\,inf}_{n\to\infty} U_n + \operatorname*{p-lim\,sup}_{n\to\infty} V_n \quad (7)$$

[14]. As are clear from (5)–(7), the limit superior and inferior in probability have properties similar to the ordinary limit superior and inferior, respectively. Throughout this paper we use the convention that the limit inferiors and superiors in probability are defined with respect to the (joint) probability of the included random variable(s). For example, the limit inferior on the left side of (6) is defined with respect to the joint probability of U_n and V_n.

We also use the following two facts in the following sections. Proofs of these facts are given in Appendix A for readers who are not familiar to the limit superior and inferior in probability.

Fact 1: Let A and B be constants. Then

$$\operatorname*{p-lim\,inf}_{n\to\infty} U_n \ge A \Leftrightarrow \lim_{n\to\infty} \Pr\{U_n \ge A - \gamma\} = 1 \quad \text{for any constant } \gamma > 0 \quad (8)$$

$$\operatorname*{p\text{-}lim\,sup}_{n\to\infty} U_n \le B \Leftrightarrow \lim_{n\to\infty} \Pr\{U_n \le B+\gamma\} = 1 \quad \text{for any constant } \gamma > 0. \tag{9}$$

Fact 2: Let $\alpha > 0$ be an arbitrary constant. If $\operatorname*{p\text{-}lim\,inf}_{n\to\infty} U_n \ge 0$, then $\operatorname*{p\text{-}lim\,inf}_{n\to\infty} \frac{1}{n^\alpha} U_n \ge 0$. Similarly, if $\operatorname*{p\text{-}lim\,sup}_{n\to\infty} U_n \le 0$, then $\operatorname*{p\text{-}lim\,sup}_{n\to\infty} \frac{1}{n^\alpha} U_n \le 0$.

III. Converse Theorem for Memoryless Sources

When we consider the (t,m)-threshold scheme, the following (A), (B), and (C) are implicitly assumed: (A) $n=1$, (B) $\mathcal{X}$ is a finite alphabet, and (C) no decoding error occurs. In particular, if $\mathcal{X}$ has the algebraic structure as a finite field, Shamir's threshold scheme realizes the (t,m)-threshold scheme in the following sense. That is, letting $E^{(1)}, E^{(2)}, \dots, E^{(t-1)} \in \mathcal{X}$ be the $t-1$ random variables subject to the uniform distribution and $\alpha_1, \alpha_2, \dots, \alpha_m$ be m distinct nonzero elements of $\mathcal{X}$, $W^{(1)}, W^{(2)}, \dots, W^{(m)}$ generated according to

$$W^{(j)} = h(\alpha_j), \quad j = 1,2,\dots,m,$$

become m shares of the (t,m)-threshold scheme satisfying $\varepsilon^{(i_1,\dots,i_t)} = 0$ for all $\{i_1,\dots,i_t\} \subset \mathcal{P}$ and $I(X; W^{(i_1,\dots,i_{t-1})}) = 0$ for all $\{i_1,\dots,i_{t-1}\} \subset \mathcal{P}$, where

$$h(u) = X + E^{(1)}u + E^{(2)}u^2 + \cdots + E^{(t-1)}u^{t-1}$$

is a random polynomial of degree at most $t-1$, $I(\cdot;\cdot)$ denotes the mutual information and all the subscripts 1 are omitted. Note that $I(X; W^{(i_1,\dots,i_{t-1})}) = 0$ means that $W^{(i_1,\dots,i_{t-1})}$ is independent of X. Shamir's threshold scheme guarantees that the fair random bits of length $(t-1)\log|\mathcal{X}|$ are enough for the dealer to realize the (t,m)-threshold scheme. This result coincides with the attainable lower bound of the dealer's randomness in [2, Th. 2.13], where the randomness is defined as $H(W^{(1,\dots,m)}|X)$.

However, if we consider block coding of sufficiently large blocklength n and permit negligible decoding error probability, then the problem of finding the minimum length of the fair random bits needed to the dealer for realizing the (t,m)-threshold scheme becomes more difficult. Suppose that $\mathcal{X}$ is a finite alphabet and $X^n \in \mathcal{X}^n$ is generated from a stationary memoryless source. If $\mathcal{X}^n$ is embedded to a finite field, say $\boldsymbol{Z}_{p_n} = \{0,1,\dots,p_n-1\}$, and elements of $\mathcal{X}^n$ are identified with elements in $\boldsymbol{Z}_{p_n}$, then Shamir's threshold scheme realizes the (t,m)-threshold scheme, where p_n is the smallest prime number satisfying $|\mathcal{X}|^n \le p_n$. It is clear that this scheme satisfies $I(X^n; W_n^{(i_1,\dots,i_{t-1})}) = 0$ for all $\{i_1,\dots,i_{t-1}\} \subset \mathcal{P}$. In this scheme, the length of the fair random bits required to the dealer equals to $(t-1)\log p_n \approx n(t-1)\log|\mathcal{X}|$.

On the other hand, this length of the fair random bits can be decreased if we consider a scheme based on the typical set. Let $\gamma > 0$ be an arbitrarily small constant and define the typical set by

$$\mathcal{A}_n = \left\{ x^n \in \mathcal{X}^n : \left| \frac{1}{n} \log \frac{1}{P_{X^n}(x^n)} - H(X) \right| \le \gamma \right\} \tag{10}$$

where $H(X)$ denotes the entropy of the source. It is known that $\Pr\{X^n \in \mathcal{A}_n\} \to 1$ as $n \to \infty$ and $|\mathcal{A}_n| \le 2^{n(H(X)+\gamma)}$ for all $n \ge 1$ [5]. If we embed $\mathcal{A}_n$ to a finite field $\boldsymbol{Z}_{p'_n}$ and apply Shamir's threshold scheme, then we can obtain a scheme that is close to the (t,m)-threshold scheme with the vanishing decoding error probability. Here, p'_n is the smallest prime number satisfying $p'_n \ge 2^{n(H(X)+\gamma)}$. In this scheme, the length of the fair random bits needed to the dealer is roughly equal to $n(t-1)H(X)$, which is smaller than $n(t-1)\log|\mathcal{X}|$. However, this scheme does not guarantee $I(X^n; W_n^{(i_1,\dots,i_{t-1})}) = 0$. That is, we cannot know in what sense this scheme is "close" to the (t,m)-threshold scheme. In addition, we cannot say nothing whether $(t-1)H(X)$ is the minimum length of the fair random bits per source symbol or not.

The objective of this section is investigation of the length of the fair random bits required to the dealer subject to the condition that the decoding error probability vanishes as $n \to \infty$. We assume that $\boldsymbol{X}$ is a stationary memoryless source with a finite alphabet $\mathcal{X}$.

We begin with the definition of the (t,m)-threshold scheme in an extended sense.

Definition 1: Let a stationary memoryless source $\boldsymbol{X} = \{X^n\}_{n=1}^\infty$ be given. If a sequence $\{(f_n, g_n)\}_{n=1}^\infty$ of encoders and decoders satisfies the following two conditions, we say that $\{(f_n, g_n)\}_{n=1}^\infty$ realizes the (t,m)-threshold scheme for $\boldsymbol{X}$:

M1) For any $\{i_1,\dots,i_t\} \subset \mathcal{P}$

$$\lim_{n\to\infty} \varepsilon_n^{(i_1,\dots,i_t)} = 0$$

M2) For any $\{i_1,\dots,i_{t-1}\} \subset \mathcal{P}$

$$\lim_{n\to\infty} \frac{1}{n} I(X^n; W_n^{(i_1,\dots,i_{t-1})}) = 0.$$

We have the following theorem that gives lower bounds on the entropies $W_n^{(i_1,\dots,i_j)}$, $j=1,2,\dots,t$, and E_n in the asymptotic sense as $n \to \infty$.

Theorem 1: Let $\{(f_n, g_n)\}_{n=1}^\infty$ be an arbitrary sequence of encoders and decoders of the (t,m)-threshold scheme for a stationary memoryless source $\boldsymbol{X} = \{X^n\}_{n=1}^\infty$ satisfying conditions M1) and M2). Then, for any $j = 1,2,\dots,t$ and $\{i_1,\dots,i_j\} \subset \mathcal{P}$ it holds that

$$\liminf_{n\to\infty} \frac{1}{n} H(W_n^{(i_1,\dots,i_j)}) \ge jH(X) \tag{11}$$

where $H(X)$ denotes the entropy of the source. In addition, it holds that

$$\liminf_{n\to\infty} \frac{1}{n} H(E_n) \ge (t-1)H(X). \tag{12}$$

We prove Theorem 1 by using the following three lemmas.

Lemma 1: Let $\{(f_n, g_n)\}_{n=1}^\infty$ be an arbitrary sequence of encoders and decoders of the (t,m)-threshold scheme for a stationary memoryless source $\boldsymbol{X}$ satisfying conditions M1) and

M2). Then, for any $j = 1, 2, \ldots, t$ and $\{i_1, \ldots, i_j\} \subset \mathcal{P}$ it holds that

$$\frac{1}{n} H(W_n^{(i_j)} | W_n^{(i_1, \ldots, i_{j-1})}) \geq H(X) + o(1)$$

where $o(1)$ denotes the terms which goes to zero as $n \to \infty$ and the left side is interpreted as $\frac{1}{n} H(W_n^{(i_1)})$ for the case of $j = 1$.

Proof: Fix $\{i_1, \ldots, i_t\} \subset \mathcal{P}$ arbitrarily. Then, for any $j = 1, 2, \ldots, t$ it follows that

$$\begin{aligned} &H(W_n^{(i_j)} | W_n^{(i_1, \ldots, i_{j-1})}) \\ &\quad \geq H(W_n^{(i_j)} | W_n^{(I_j)}) - H(W_n^{(i_j)} | X^n W_n^{(I_j)}) \\ &\quad = I(X^n; W_n^{(i_j)} | W_n^{(I_j)}) \\ &\quad = H(X^n | W_n^{(I_j)}) - H(X^n | W_n^{(i_1, \ldots, i_t)}) \\ &\quad = H(X^n) - I(X^n; W_n^{(I_j)}) - H(X^n | W_n^{(i_1, \ldots, i_t)}) \end{aligned} \quad (13)$$

where $W_n^{(I_j)} = W_n^{(i_1, \ldots, i_{j-1}, i_{j+1}, \ldots, i_t)}$ and the inequality follows because $H(W_n^{(i_j)} | X^n W_n^{(I_j)}) \geq 0$ and conditioning does not increase the entropy. By dividing both sides of (13) by n and using $H(X^n) = nH(X)$, we have

$$\begin{aligned} &\frac{1}{n} H(W_n^{(i_j)} | W_n^{(i_1, \ldots, i_{j-1})}) \\ &\quad = H(X) - \frac{1}{n} I(X^n; W_n^{(I_j)}) - \frac{1}{n} H(X^n | W_n^{(i_1, \ldots, i_t)}). \end{aligned} \quad (14)$$

Note that the second term on the right side of (14) goes to zero as $n \to \infty$ from condition M2). In order to evaluate the third term on the right side of (14), we use Fano's inequality (e.g., [5]). Fano's inequality tells us that

$$\begin{aligned} &\frac{1}{n} H(X^n | g_n^{(i_1, \ldots, i_t)}(W_n^{(i_1, \ldots, i_t)})) \\ &\qquad \leq \varepsilon_n^{(i_1, \ldots, i_t)} \log |\mathcal{X}| + \frac{1}{n} h_2(\varepsilon_n^{(i_1, \ldots, i_t)}) \end{aligned} \quad (15)$$

where $h_2(\cdot)$ denotes the binary entropy. Notice here that we have

$$\begin{aligned} &H(X^n | W_n^{(i_1, \ldots, i_t)}) \\ &\quad = H(X^n | W_n^{(i_1, \ldots, i_t)}, g_n^{(i_1, \ldots, i_t)}(W_n^{(i_1, \ldots, i_t)})) \\ &\quad \leq H(X^n | g_n^{(i_1, \ldots, i_t)}(W_n^{(i_1, \ldots, i_t)})) \end{aligned} \quad (16)$$

where the equality follows because $g_n^{(i_1, \ldots, i_t)}$ is deterministic. Hence, the combination of (15) and (16) leads to

$$\frac{1}{n} H(X^n | W_n^{(i_1, \ldots, i_t)}) \leq \varepsilon_n^{(i_1, \ldots, i_t)} \log |\mathcal{X}| + \frac{1}{n} h_2(\varepsilon_n^{(i_1, \ldots, i_t)})$$

which goes to zero as $n \to \infty$ owing to condition M1). Thus, we obtain the claim of this lemma. □

Lemma 2: Let $\{(f_n, g_n)\}_{n=1}^{\infty}$ be an arbitrary sequence of encoders and decoders of the (t, m)-threshold scheme for a stationary memoryless source $\boldsymbol{X}$ satisfying conditions M1) and M2). Then, for any $\{i_1, \ldots, i_t\} \subset \mathcal{P}$ it holds that

$$\frac{1}{n} H(W_n^{(i_1, \ldots, i_t)} | X^n) \geq \frac{1}{n} H(W_n^{(i_1, \ldots, i_{t-1})}) + o(1).$$

Proof: Fix $\{i_1, \ldots, i_t\} \subset \mathcal{P}$ arbitrarily. From the chain rule of the entropy, we have

$$\begin{aligned} &H(W_n^{(i_1, \ldots, i_t)} | X^n) \\ &\quad = H(W_n^{(i_1, \ldots, i_{t-1})} | X^n) + H(W_n^{(i_t)} | X^n W_n^{(i_1, \ldots, i_{t-1})}) \\ &\quad \geq H(W_n^{(i_1, \ldots, i_{t-1})} | X^n) \end{aligned} \quad (17)$$

where the inequality follows because $H(W_n^{(i_t)} | X^n W_n^{(i_1, \ldots, i_{t-1})}) \geq 0$. In addition, by the definition of the mutual information, we have

$$H(W_n^{(i_1, \ldots, i_{t-1})} | X^n) = H(W_n^{(i_1, \ldots, i_{t-1})}) - I(X^n; W_n^{(i_1, \ldots, i_{t-1})}). \quad (18)$$

The claim of the lemma follows from the combination of (17), (18) and condition M2). □

Lemma 3: Let $\{(f_n, g_n)\}_{n=1}^{\infty}$ be an arbitrary sequence of encoders and decoders of the (t, m)-threshold scheme for a stationary memoryless source $\boldsymbol{X}$ satisfying conditions (M1) and (M2). Then, for any $\{i_1, \ldots, i_t\} \subset \mathcal{P}$ it holds that

$$H(E_n) \geq H(W_n^{(i_1, \ldots, i_t)} | X^n).$$

Proof: Fix $\{i_1, \ldots, i_t\} \subset \mathcal{P}$ arbitrarily. Since $f_n^{(i_1, \ldots, i_t)}$ is deterministic, it holds that

$$\begin{aligned} H(X^n E_n W_n^{(i_1, \ldots, i_t)}) &= H(X^n E_n) \\ &= H(X^n) + H(E_n) \end{aligned} \quad (19)$$

where the last equality in (19) follows because X^n is independent of E_n. On the other hand, by the chain rule of the entropy, we have

$$\begin{aligned} &H(X^n E_n W_n^{(i_1, \ldots, i_t)}) \\ &\quad = H(X^n W_n^{(i_1, \ldots, i_t)}) + H(E_n | X^n W_n^{(i_1, \ldots, i_t)}) \\ &\quad \geq H(X^n W_n^{(i_1, \ldots, i_t)}) \\ &\quad = H(X^n) + H(W_n^{(i_1, \ldots, i_t)} | X^n), \end{aligned} \quad (20)$$

where the inequality in (20) follows from $H(E_n | X^n W_n^{(i_1, \ldots, i_t)}) \geq 0$. Then, the claim of the lemma follows from the combination of (19) and (20). □

Proof of Theorem 1: Fix $\{i_1, \ldots, i_t\} \subset \mathcal{P}$ arbitrarily. By the chain rule of the entropy and Lemma 1, for each $j = 1, 2, \ldots, t$ it holds that

$$\begin{aligned} \frac{1}{n} H(W_n^{(i_1, \ldots, i_j)}) &= \frac{1}{n} \sum_{k=1}^{j} H(W_n^{(i_k)} | W_n^{(i_1, \ldots, i_{k-1})}) \\ &\geq jH(X) + o(1) \end{aligned} \quad (21)$$

which implies the first claim of Theorem 1.

Next, we prove the second claim of Theorem 1. It follows from Lemmas 2 and 3 that

$$\begin{aligned} \frac{1}{n} H(E_n) &\geq \frac{1}{n} H(W_n^{(i_1, \ldots, i_t)} | X^n) \\ &\geq \frac{1}{n} H(W_n^{(i_1, \ldots, i_{t-1})}) + o(1). \\ &\geq (t-1)H(X) + o(1) \end{aligned} \quad (22)$$

where the last inequality follows from (21) with $j = t-1$. Clearly, (22) implies (12). □

IV. Converse Theorem for General Sources

In the preceding section we have developed asymptotic lower bounds of $\frac{1}{n}H(W_n^{(i_1,\ldots,i_j)})$ and $\frac{1}{n}H(E_n)$ in the (t,m)-threshold scheme for a stationary memoryless source $\boldsymbol{X}$ with a finite alphabet. Recall that Lemma 1 is proved by using $\frac{1}{n}H(X^n) = H(X)$ and Fano's inequality, which cannot be used without the assumptions on the source.

In this section, we consider the (t,m)-threshold scheme for a general source $\boldsymbol{X}$. Here, a class of general sources includes various classes of sources as mentioned in Section II. In addition, the source alphabet $\mathcal{X}$ of $\boldsymbol{X}$ can be countably infinite. Chor and Kushilevitz [4] show that there is no (t,m)-threshold scheme if $|\mathcal{X}|$ is countably infinite. However, the (t,m)-threshold scheme with vanishing decoding error probability is out of the scope in [4]. We will see that we can construct the (t,m)-threshold scheme in a certain sense even if $\mathcal{X}$ is a countably infinite alphabet.

We begin with a new definition of the (t,m)-threshold scheme. We do not use the mutual information as a measure of security of X^n against less than t shares. We do not use the spectral sup-mutual information rate, which is a certain generalization of the mutual information and is defined by using the limit superior in probability [6], either. Instead, we impose a criterion described as an inequality including the limit superior in probability. Note that this kind of criterion has never been discussed in the ordinary framework of information-spectrum methods [6] so far. Under the new definition of the (t,m)-threshold scheme, we can obtain fundamental inequalities that are closely related to the length of the fair random bits required to the dealer.

Throughout this section we assume that E_n is the uniformly distributed random variable on $\mathcal{E}_n$. We define the (t,m)-threshold scheme for a general source $\boldsymbol{X}$ as follows:

Definition 2: If a sequence $\{(f_n,g_n)\}_{n=1}^\infty$ of encoders and decoders satisfies the following two conditions G1) and G2), we say that $\{(f_n,g_n)\}_{n=1}^\infty$ realizes the (t,m)-threshold scheme for a general source $\boldsymbol{X}$:

G1) For any $\{i_1,\ldots,i_t\} \subset \mathcal{P}$

$$\lim_{n\to\infty} \varepsilon_n^{(i_1,\ldots,i_t)} = 0.$$

G2) For any $\{i_1,\ldots,i_{t-1}\} \subset \mathcal{P}$

$$\text{p-}\limsup_{n\to\infty} \log \frac{P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(X^n|W_n^{(i_1,\ldots,i_{t-1})})}{P_{X^n}(X^n)} \le 0 \quad (23)$$

where $P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}$ denotes the conditional probability distribution of X^n given $W_n^{(i_1,\ldots,i_{t-1})}$.

Readers may feel strange to condition G2). However, the meaning of condition G2) becomes clear by considering the following condition:

G2') For any $\alpha > 0$ and $\{i_1,\ldots,i_{t-1}\} \subset \mathcal{P}$

$$\text{p-}\limsup_{n\to\infty} \frac{1}{n^\alpha} \log \frac{P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(X^n|W_n^{(i_1,\ldots,i_{t-1})})}{P_{X^n}(X^n)} \le 0. \quad (24)$$

Note that, if condition G2) is satisfied, then G2') is also satisfied owing to Fact 2 in Section II. Furthermore, since we can easily prove

$$\text{p-}\liminf_{n\to\infty} \frac{1}{n^\alpha} \log \frac{P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(X^n|W_n^{(i_1,\ldots,i_{t-1})})}{P_{X^n}(X^n)} \ge 0 \quad (25)$$

for any $\alpha > 0$ and $\{i_1,\ldots,i_{t-1}\} \subset \mathcal{P}$ similarly to [6, Lemma 3.2.1], G2') actually means that

$$\lim_{n\to\infty} \Pr\{(X^n, W_n^{(i_1,\ldots,i_{t-1})}) \in \mathcal{T}_n\} = 1 \quad (26)$$

where

$$\mathcal{T}_n = \left\{ (x^n, w_n^{(i_1,\ldots,i_{t-1})}) \in \mathcal{X}^n \times \mathcal{W}_n^{(i_1,\ldots,i_{t-1})} : \left| \frac{1}{n^\alpha} \log \frac{P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(x^n|w_n^{(i_1,\ldots,i_{t-1})})}{P_{X^n}(x^n)} \right| \le \gamma \right\}$$

and $\gamma > 0$ is an arbitrarily small constant. Clearly, (26) implies that, if n is sufficiently large, $P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(x^n|w_n^{(i_1,\ldots,i_{t-1})})$ is arbitrarily close to $P_{X^n}(x^n)$, i.e., $W_n^{(i_1,\ldots,i_{t-1})}$ is almost independent of X^n, on the set $\mathcal{T}_n$ with probability arbitrarily close to one. We can regard G2) as the stronger version of G2') with $\alpha \downarrow 0$.

While Theorem 2 below holds under G1) and G2'), we can construct the (t,m)-threshold scheme satisfying G1) and G2) in Section V. That is why we use G2) as a criterion on secrecy.

Remark 1: Note that G2) implies neither $I(X^n; W_n^{(i_1,\ldots,i_{t-1})}) \to 0$ nor $\frac{1}{n}I(X^n; W_n^{(i_1,\ldots,i_{t-1})}) \to 0$ as $n \to \infty$, in general. In fact, $I(X^n; W_n^{(i_1,\ldots,i_{t-1})})$ is defined as the expectation of $i(X^n; W_n^{(i_1,\ldots,i_{t-1})}) \overset{\text{def}}{=} \log \frac{P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(X^n|W_n^{(i_1,\ldots,i_{t-1})})}{P_{X^n}(X^n)}$ with respect to $P_{X^n W_n^{(i_1,\ldots,i_{t-1})}}$, while condition G2) only imposes a certain property on the distribution of $i(X^n; W_n^{(i_1,\ldots,i_{t-1})})$. This is the why characterizations of the (t,m)-threshold scheme under G2) are so successful. Readers can easily check that $I(X^n; W_n^{(i_1,\ldots,i_{t-1})}) \to 0$ (or even $\frac{1}{n}I(X^n; W_n^{(i_1,\ldots,i_{t-1})}) \to 0$) as $n \to \infty$ does not hold only from $\Pr\{(X^n, W_n^{(i_1,\ldots,i_{t-1})}) \in \mathcal{T}_n\} \to 1$ as $n \to \infty$. We need some condition on the boundedness of $i(x^n; w_n^{(i_1,\ldots,i_{t-1})})$ for $(x^n; w_n^{(i_1,\ldots,i_{t-1})}) \notin \mathcal{T}_n$ (recall that $\mathcal{X}$ can be a countably infinite alphabet). □

We have the following theorem under the new definition of the (t,m)-threshold scheme.

Theorem 2: Given a general source $\boldsymbol{X} = \{X^n\}_{n=1}^\infty$, let $\{(f_n,g_n)\}_{n=1}^\infty$ be a sequence of encoders and decoders of the

(t,m)-threshold scheme for $\boldsymbol{X}$ satisfying conditions G1) and G2). Then, for any $j=1,2,\ldots,t$, $\{i_1,\ldots,i_j\}\subset\mathcal{P}$ and constant $\alpha>0$ it holds that

$$\text{p-}\liminf_{n\to\infty}\frac{1}{n^\alpha}\log\frac{(P_{X^n}(X^n))^j}{P_{W_n^{(i_1,\ldots,i_j)}}(W_n^{(i_1,\ldots,i_j)})}\geq 0. \tag{27}$$

In addition, for any constant $\alpha>0$ it holds that

$$\text{p-}\liminf_{n\to\infty}\frac{1}{n^\alpha}\log\left(|\mathcal{E}_n|(P_{X^n}(X^n))^{t-1}\right)\geq 0. \tag{28}$$

Remark 2: In (27) and (28) n^α can be replaced with any sequence a_n, $n\geq 1$, satisfying $a_n\geq 1$ and $a_n\to\infty$ as $n\to\infty$. We use n^α for simplifying notations and facilitating comparison to the ordinary case of $\alpha=1$. □

We can prove Theorem 2 by using the following three lemmas. The first lemma, which is proved in Appendix B, characterizes a property of $\{(f_n,g_n)\}_{n=1}^\infty$ satisfying G1).

Lemma 4: Let $\{(f_n,g_n)\}_{n=1}^\infty$ be an arbitrary sequence of encoders and decoders satisfying condition G1). Then, for any $\{i_1,\ldots,i_t\}\subset\mathcal{P}$ it holds that

$$\text{p-}\limsup_{n\to\infty}\log\frac{1}{P_{X^n|W_n^{(i_1,\ldots,i_t)}}(X^n|W_n^{(i_1,\ldots,i_t)})}=0.$$

The following lemma characterizes an important property on conditional distribution related to t shares $W_n^{(i_1,\ldots,i_t)}$.

Lemma 5: Let $\{(f_n,g_n)\}_{n=1}^\infty$ be an arbitrary sequence of encoders and decoders. Then, for any $j=1,2,\ldots,t$, $\{i_1,\ldots,i_t\}\subset\mathcal{P}$ and constant $\alpha>0$ it holds that

$$\text{p-}\liminf_{n\to\infty}\frac{1}{n^\alpha}\log\frac{P_{W_n^{(i_j)}|W_n^{(I_J)}}(W_n^{(i_j)}|W_n^{(I_j)})}{P_{W_n^{(i_j)}|W_n^{(i_1,\ldots,i_{j-1})}}(W_n^{(i_j)}|W_n^{(i_1,\ldots,i_{j-1})})}\geq 0$$

for any $\alpha>0$, where $W_n^{(I_j)}=W_n^{(i_1,\ldots,i_{j-1},i_{j+1},\ldots,i_t)}$.

The proof of Lemma 5 is essentially the same as the proof of [6, Lemma 3.2.1]. However, we give the proof of Lemma 5 in Appendix C for readers' convenience.

We also use the following lemma in the proof of Theorem 2. This lemma plays a role that is similar to the role of Lemma 1 in the proof of Theorem 1.

Lemma 6: Let $\{(f_n,g_n)\}_{n=1}^\infty$ be an arbitrary sequence of encoders and decoders of the (t,m)-threshold scheme satisfying conditions G1) and G2). Then, for any $j=1,2,\ldots,t$, $\{i_1,\ldots,i_j\}\subset\mathcal{P}$ and constant $\alpha>0$ it holds that

$$\text{p-}\liminf_{n\to\infty}\frac{1}{n^\alpha}\log\frac{P_{X^n}(X^n)}{P_{W_n^{(i_j)}|W_n^{(i_1,\ldots,i_{j-1})}}(W_n^{(i_j)}|W_n^{(i_i,\ldots,i_{j-1})})}\geq 0$$

where the denominator on the left side is interpreted as $P_{W_n^{(i_1)}}(W_n^{(i_1)})$ for the case of $j=1$.

Proof: Fix $\{i_1,\ldots,i_t\}\subset\mathcal{P}$ arbitrarily. Define $\mathcal{D}_n$ by

$$\mathcal{D}_n=\left\{(x^n,w_n^{(i_1,\ldots,i_t)})\in\mathcal{X}^n\times\mathcal{W}_n^{(i_1,\ldots,i_t)}:\right.\\ \left.\frac{1}{n^\alpha}\log\frac{1}{P_{X^n|W_n^{(i_1,\ldots,i_t)}}(x^n|w_n^{(i_1,\ldots,i_t)})}\leq\gamma\right\}.$$

For each $j=1,2,\ldots,t-1$ we define $\mathcal{U}_n^{(i_j)}$ and $\mathcal{V}_n^{(i_j)}$ as follows:

$$\mathcal{U}_n^{(i_j)}=\left\{(x^n,w_n^{(i_1,\ldots,i_t)})\in\mathcal{X}^n\times\mathcal{W}_n^{(i_1,\ldots,i_t)}:\right.\\ \left.\frac{1}{n^\alpha}\log\frac{P_{X^n|W_n^{(I_j)}}(x^n|w_n^{(I_j)})}{P_{X^n}(x^n)}\leq\gamma\right\}$$

$$\mathcal{V}_n^{(i_j)}=\left\{(x^n,w_n^{(i_1,\ldots,i_t)})\in\mathcal{X}^n\times\mathcal{W}_n^{(i_1,\ldots,i_t)}:\right.\\ \left.\frac{1}{n^\alpha}\log\frac{P_{W_n^{(i_j)}|W_n^{(I_j)}}(w_n^{(i_j)}|w_n^{(I_j)})}{P_{W_n^{(i_j)}|W_n^{(i_1,\ldots,i_{j-1})}}(w_n^{(i_j)}|w_n^{(i_1,\ldots,i_{j-1})})}\geq-\gamma\right\}$$

where $w_n^{(I_j)}=w_n^{(i_1,\ldots,i_{j-1},i_{j+1},\ldots,i_t)}$. Then, it follows from Lemma 4 and Fact 2 that

$$\lim_{n\to\infty}\Pr\{(X^n,W_n^{(i_1,\ldots,i_t)})\in\mathcal{D}_n\}=1. \tag{29}$$

In addition, condition G2) guarantees that

$$\lim_{n\to\infty}\Pr\{(X^n,W_n^{(i_1,\ldots,i_t)})\in\mathcal{U}_n^{(i_j)}\}=1\\ \text{for all } j=1,2,\ldots,t. \tag{30}$$

Furthermore, in view of Lemma 5 we obtain

$$\lim_{n\to\infty}\Pr\{(X^n,W_n^{(i_1,\ldots,i_t)})\in\mathcal{V}_n^{(i_j)}\}=1\\ \text{for all } j=1,2,\ldots,t. \tag{31}$$

Therefore, the combination of (29), (30) and (31) yields

$$\lim_{n\to\infty}\Pr\left\{(X^n,W_n^{(i_1,\ldots,i_t)})\in\mathcal{D}_n\cap\mathcal{U}_n^{(i_j)}\cap\mathcal{V}_n^{(i_j)}\right\}=1\\ \text{for all } j=1,2,\ldots,t. \tag{32}$$

It is important to notice that for an arbitrarily fixed $j=1,2,\ldots,t$ we have

$$\frac{1}{n^\alpha}\log\frac{P_{W_n^{(i_j)}|W_n^{(I_j)}}(w_n^{(i_j)}|w_n^{(I_j)})}{P_{W_n^{(i_j)}|W_n^{(i_1,\ldots,i_{j-1})}}(w_n^{(i_j)}|w_n^{(i_1,\ldots,i_{j-1})})}\\ +\frac{1}{n^\alpha}\log\frac{1}{P_{W_n^{(i_j)}|X^nW_n^{(I_j)}}(w_n^{(i_j)}|x^n,w_n^{(I_j)})}\geq-\gamma\\ \text{for all } (x^n,w_n^{(i_1,\ldots,i_t)})\in\mathcal{V}_n^{(i_j)}. \tag{33}$$

In fact, this inequality follows because the first term is greater than or equal to $-\gamma$ from the definition of $\mathcal{V}_n^{(i_j)}$ and the second term is nonnegative. By using Bayes' formula

$$\begin{aligned} P_{X^n W_n^{(i_j)} | W_n^{(I_j)}} &= P_{X^n | W_n^{(I_j)}} P_{W_n^{(i_j)} | X^n W_n^{(I_j)}} \\ &= P_{W_n^{(i_j)} | W_n^{(I_j)}} P_{X^n | W_n^{(i_1,\ldots,i_t)}} \end{aligned}$$

we can rewrite (33) in the following form:

$$\begin{aligned} &\frac{1}{n^\alpha} \log \frac{P_{X^n}(x^n)}{P_{W_n^{(i_j)} | W_n^{(i_1,\ldots,i_{j-1})}}(w_n^{(i_j)} | w_n^{(i_1,\ldots,i_{j-1})})} \\ &\quad + \frac{1}{n^\alpha} \log \cdot \frac{P_{X^n | W_n^{(I_j)}}(x^n | w_n^{(I_j)})}{P_{X^n}(x^n)} \\ &\quad + \frac{1}{n^\alpha} \log \frac{1}{P_{X^n | W_n^{(i_1,\ldots,i_t)}}(x^n | w_n^{(i_1,\ldots,i_t)})} \geq -\gamma \\ &\qquad\qquad \text{for all } (x^n, w_n^{(i_1,\ldots,i_t)}) \in \mathcal{V}_n^{(i_j)} \end{aligned}$$

which means that

$$\begin{aligned} &\frac{1}{n^\alpha} \log \frac{P_{X^n}(x^n)}{P_{W_n^{(i_j)} | W_n^{(i_1,\ldots,i_{j-1})}}(w_n^{(i_j)} | w_n^{(i_1,\ldots,i_{j-1})})} \geq -3\gamma \\ &\qquad \text{for all } (x^n, w_n^{(i_1,\ldots,i_t)}) \in \mathcal{D}_n \cap \mathcal{U}_n^{(i_j)} \cap \mathcal{V}_n^{(i_j)}. \qquad (34) \end{aligned}$$

Since $\gamma > 0$ can be arbitrarily small, (32) and (34) lead to the claim of the lemma. □

Proof of Theorem 2: Fix $\{i_1, \ldots, i_t\} \subset \mathcal{P}$ and $j = 1, 2, \ldots, t$ arbitrarily. First, we prove (27). Due to the property (6) of the limit inferior in probability, the left side of (27) is lower bounded by

$$\begin{aligned} &\sum_{k=1}^{j} \operatorname*{p\text{-}\lim\inf}_{n\to\infty} \frac{1}{n^\alpha} \\ &\qquad \times \log \frac{P_{X^n}(X^n)}{P_{W_n^{(i_k)} | W_n^{(i_1,\ldots,i_{k-1})}}(W_n^{(i_k)} | W_n^{(i_i,\ldots,i_{k-1})})}. \qquad (35) \end{aligned}$$

Since Lemma 6 tells us that every term in (35) is nonnegative, we have (27).

Next, we prove (28). Since the encoder f_n is deterministic, we have

$$P_{W_n^{(1,\ldots,m)} | X^n}(w_n^{(1,\ldots,m)} | x^n) \geq \frac{1}{|\mathcal{E}_n|}$$

for all $(x^n, w_n^{(1,\ldots,m)}) \in \mathcal{X}^n \times \mathcal{W}_n^{(1,\ldots,m)}$ satisfying $P_{X^n W_n^{(1,\ldots,m)}}(x^n, w_n^{(1,\ldots,m)}) > 0$. That is, if $P_{X^n W_n^{(1,\ldots,m)}}(x^n, w_n^{(1,\ldots,m)}) > 0$, there exists at least one $e_n \in \mathcal{E}_n$ satisfying $w_n^{(1,\ldots,m)} = f_n(x^n, e_n)$. Recall here that E_n is uniformly distributed on $\mathcal{E}_n$. Thus, it holds that

$$P_{W_n^{(i_1,\ldots,i_{t-1})} | X^n}(w_n^{(i_1,\ldots,i_{t-1})} | x^n) \geq \frac{1}{|\mathcal{E}_n|}$$

for all $(x^n, w_n^{(i_1,\ldots,i_{t-1})}) \in \mathcal{X}^n \times \mathcal{W}_n^{(i_1,\ldots,i_{t-1})}$ satisfying $P_{X^n W_n^{(i_1,\ldots,i_{t-1})}}(x^n, w_n^{(i_1,\ldots,i_{t-1})}) > 0$. Hence, letting

$$\begin{aligned} \mathcal{F}_n = \Big\{ &(x^n, w_n^{(i_1,\ldots,i_{t-1})}) \in \mathcal{X}^n \times \mathcal{W}_n^{(i_1,\ldots,i_{t-1})} : \\ &\qquad P_{X^n W_n^{(i_1,\ldots,i_{t-1})}}(x^n, w_n^{(i_1,\ldots,i_{t-1})}) > 0 \Big\} \end{aligned}$$

we have

$$\begin{aligned} &\frac{1}{n^\alpha} \log \Big[|\mathcal{E}_n| P_{W_n^{(i_1,\ldots,i_{t-1})} | X^n}(w_n^{(i_1,\ldots,i_{t-1})} | x^n) \Big] \geq 0 \\ &\qquad\qquad \text{for all } (x^n, w_n^{(i_1,\ldots,i_{t-1})}) \in \mathcal{F}_n \qquad (36) \end{aligned}$$

and

$$\Pr\Big\{ (X^n, W_n^{(i_1,\ldots,i_{t-1})}) \in \mathcal{F}_n \Big\} = 1. \qquad (37)$$

Notice here that Bayes' formula tells us that (36) can be written as

$$\begin{aligned} &\frac{1}{n^\alpha} \log \Bigg[|\mathcal{E}_n| ((P_{X^n}(x^n))^{t-1} \frac{P_{W_n^{(i_1,\ldots,i_{t-1})}}(w_n^{(i_1,\ldots,i_{t-1})})}{(P_{X^n}(x^n))^{t-1}} \\ &\qquad \times \frac{P_{X^n | W_n^{(i_1,\ldots,i_{t-1})}}(x^n | w_n^{(i_1,\ldots,i_{t-1})})}{P_{X^n}(x^n)} \Bigg] \geq 0 \\ &\qquad\qquad \text{for all } (x^n, w_n^{(i_1,\ldots,i_{t-1})}) \in \mathcal{F}_n. \qquad (38) \end{aligned}$$

Now, define

$$\begin{aligned} \mathcal{Y}_n = \Bigg\{ &(x^n, w_n^{(i_1,\ldots,i_{t-1})}) \in \mathcal{X}^n \times \mathcal{W}_n^{(i_1,\ldots,i_{t-1})} : \\ &\qquad \frac{1}{n^\alpha} \log \frac{P_{W_n^{(i_1,\ldots,i_{t-1})}}(w_n^{(i_1,\ldots,i_{t-1})})}{(P_{X^n}(x^n))^{t-1}} \leq \gamma \Bigg\} \\ \tilde{\mathcal{U}}_n = \Bigg\{ &(x^n, w_n^{(i_1,\ldots,i_{t-1})}) \in \mathcal{X}^n \times \mathcal{W}_n^{(i_1,\ldots,i_{t-1})} : \\ &\qquad \frac{1}{n^\alpha} \log \frac{P_{X^n | W_n^{(i_1,\ldots,i_{t-1})}}(x^n | w_n^{(i_1,\ldots,i_{t-1})})}{P_{X^n}(x^n) \leq \gamma} \Bigg\} \end{aligned}$$

for an arbitrary constant $\gamma > 0$. Then, owing to (27) with $j = t - 1$ it holds that

$$\lim_{n\to\infty} \Pr\{ (X^n, W_n^{(i_1,\ldots,i_{t-1})}) \in \mathcal{Y}_n \} = 1. \qquad (39)$$

In addition, condition G2) tells us that we have

$$\lim_{n\to\infty} \Pr\{ (X^n, W_n^{(i_1,\ldots,i_{t-1})}) \in \tilde{\mathcal{U}}_n \} = 1. \qquad (40)$$

In view of (38) it easily follows that

$$\begin{aligned} &\frac{1}{n^\alpha} \log\big(|\mathcal{E}_n| (P_{X^n}(x^n))^{t-1} \big) \geq -2\gamma \\ &\qquad\qquad \text{for all } (x^n, w_n^{(i_1,\ldots,i_{t-1})}) \in \mathcal{Y}_n \cap \tilde{\mathcal{U}}_n \cap \mathcal{F}_n. \qquad (41) \end{aligned}$$

Since (37), (39) and (40) guarantee

$$\lim_{n\to\infty} \Pr\{(X^n, W_n^{(i_1,\ldots,i_{t-1})}) \in \mathcal{Y}_n \cap \tilde{\mathcal{U}}_n \cap \mathcal{F}_n\} = 1 \tag{42}$$

and $\gamma > 0$ is arbitrary, (41) and (42) yield (28). □

Theorem 2 yields the following corollary.

Corollary 1: Let $\{(f_n, g_n)\}_{n=1}^{\infty}$ be an arbitrary sequence of encoders and decoders realizing the (t, m)-threshold scheme for a general source $\boldsymbol{X}$. Then, it holds that

$$\liminf_{n\to\infty} \frac{1}{n} \log |\mathcal{E}_n| \geq (t-1)\underline{H}(\boldsymbol{X})$$

where $\underline{H}(\boldsymbol{X})$ is the spectrum inf-entropy rate of $\boldsymbol{X}$ defined in (3).

Corollary 1 immediately follows from the combination of (5), (7), (28) with $\alpha = 1$ and

$$\text{p-}\liminf_{n\to\infty} \frac{1}{n} \log |\mathcal{E}_n| = \liminf_{n\to\infty} \frac{1}{n} \log |\mathcal{E}_n|.$$

It is known that, if $\boldsymbol{X}$ is a stationary memoryless source with the entropy $H(X) < \infty$, $\underline{H}(\boldsymbol{X})$ in (3) coincides with $H(X)$. Hence, Corollary 1, together with $\underline{H}(\boldsymbol{X}) = H(X)$, leads to the same claim (11) in Theorem 1 under the different criterion G2) on the secrecy of X^n under arbitrary $t-1$ shares.

V. Direct Theorem for General Sources

In the preceding section we have developed the inequality (28) that characterizes $|\mathcal{E}_n|$ in the (t, m)-threshold scheme for a general source. In this section we investigate construction of the (t, m)-threshold scheme for a general source satisfying conditions G1) and G2) in Definition 2.

We have the following direct theorem that is valid for each $n \geq 1$.

Theorem 3: Let a general source $\boldsymbol{X} = \{X^n\}_{n=1}^{\infty}$ be given. Suppose that an arbitrary sequence $\{p_n\}_{n=1}^{\infty}$ of prime numbers satisfying $p_n > m$ for all $m \geq 1$ and

$$\tau_n \overset{\text{def}}{=} \Pr\{\log(p_n P_{X^n}(X^n)) \geq 0\} \in (0, 1] \tag{43}$$

for all $n \geq 1$ is given, where

$$\mathcal{S}_n = \{x^n \in \mathcal{X}^n : \log(p_n P_{X^n}(x^n)) \geq 0\}. \tag{44}$$

Define $\mathcal{E}_n = \{0, 1, \ldots, p_n - 1\}^{t-1}$ and let E_n be the uniformly distributed random variable on $\mathcal{E}_n$. Then, for each $n \geq 1$, there exist an encoder f_n and a decoder g_n satisfying

$$|\mathcal{W}_n^{(i)}| = p_n + 1 \quad \text{for all } i = 1, 2, \ldots, m \tag{45}$$

$$\Pr\left\{g_n^{(i_1,\ldots,i_t)}(f_n^{(i_1,\ldots,i_t)}(X^n, E_n)) = X^n \;\middle|\; X^n \in \mathcal{S}_n\right\} = 1 \quad \text{for all } \{i_1, \ldots, i_t\} \subset \mathcal{P} \tag{46}$$

and

$$\Pr\left\{\log \frac{P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(X^n|W_n^{(i_1,\ldots,i_{t-1})})}{P_{X^n}(X^n)} = \log \frac{1}{\tau_n} \;\middle|\; X^n \in \mathcal{S}_n\right\} = 1 \quad \text{for all } \{i_1, \ldots, i_{t\text{-}1}\} \subset \mathcal{P}. \tag{47}$$

In addition, the above encoder and decoder satisfy

$$\varepsilon_n^{(i_1,\ldots,i_t)} = 1 - \tau_n \quad \text{for all } \{i_1, \ldots, i_t\} \subset \mathcal{P} \tag{48}$$

Proof: Fix $n \geq 1$ and a prime number p_n satisfying $\tau_n \in (0, 1]$ arbitrarily. Since $P_{X^n}(x^n) \geq 1/p_n$ for all $x^n \in \mathcal{S}_n$, we have $|\mathcal{S}_n| \leq p_n$ by using the following argument:

$$1 = \sum_{x^n \in \mathcal{X}^n} P_{X^n}(x^n) \geq \sum_{x^n \in \mathcal{S}_n} P_{X^n}(x^n) \geq |\mathcal{S}_n| \frac{1}{p_n}. \tag{49}$$

Therefore, there exists a one-to-one mapping $\varphi_n : \mathcal{S}_n \to \boldsymbol{Z}_{p_n}$, where $\boldsymbol{Z}_{p_n} = \{0, 1, \ldots, p_n - 1\}$. We define additions, subtractions, multiplications and divisions of two elements of $\boldsymbol{Z}_{p_n}$ as the respective operations under modulo p_n. In addition, define $E_n^{(i)}, i = 1, 2, \ldots, t-1$, as the uniformly distributed random variables on $\boldsymbol{Z}_{p_n}$ and $E_n = (E_n^{(1)}, E_n^{(2)}, \ldots, E_n^{(t-1)})$. Clearly, E_n is uniformly distributed on $\mathcal{E}_n = \boldsymbol{Z}_{p_n}^{t-1}$. Furthermore, we arbitrarily choose distinct m elements α_i, $i = 1, 2, \ldots, m$, from $\boldsymbol{Z}_{p_n}$ all of which are not equal to zero. We can choose such α_i, $i = 1, 2, \ldots, m$, owing to the assumption of $p_n > m$. We do not explicitly write dependency of α_i on n for simplifying notations. Set $\mathcal{W}_n^{(i)} = \boldsymbol{Z}_{p_n} \cup \{p_n\}$ for all $i = 1, 2, \ldots, m$.

We use the following pair of an encoder and a decoder. Basic idea is application of Shamir's threshold scheme to the elements of $\mathcal{S}_n$.

Encoder $f_n : \mathcal{X}^n \times \mathcal{E}_n \to \mathcal{W}_n^{(1,\ldots,m)}$

Let $E_n = (E_n^{(1)}, E_n^{(2)}, \ldots, E_n^{(t-1)}) \in \mathcal{E}_n$ is given.

1) If $X^n \in \mathcal{S}_n$, the encoder outputs

$$(W_n^{(1)}, W_n^{(2)}, \ldots, W_n^{(m)}) = (h_n(\alpha_1), h_n(\alpha_2), \ldots, h_n(\alpha_m))$$

where

$$h_n(u) = \varphi_n(X^n) + E_n^{(1)}u + E_n^{(2)}u^2 + \cdots + E_n^{(t-1)}u^{t-1}$$

is a random polynomial with the degree at most $t-1$.

2) Otherwise, the encoder outputs

$$(W_n^{(1)}, W_n^{(2)}, \ldots, W_n^{(m)}) = (p_n, p_n, \ldots, p_n).$$

Decoder $g_n^{(i_1,\ldots,i_t)} : \mathcal{W}^{(i_1,\ldots,i_t)} \to \mathcal{X}^n$, $\{i_1, \ldots, i_t\} \subset \mathcal{P}$

Let $\{i_1, \ldots, i_t\} \subset \mathcal{P}$ and $W_n^{(i_1,\ldots,i_t)} \in \mathcal{W}_n^{(i_1,\ldots,i_t)}$ be given.

1) If $W_n^{(i_1,\ldots,i_t)} \in \boldsymbol{Z}_{p_n}^t$, the decoder computes $U_n \in \boldsymbol{Z}_{p_n}$ by solving the following system of linear equations:

$$\begin{bmatrix} W_n^{(i_1)} \\ W_n^{(i_2)} \\ \vdots \\ W_n^{(i_t)} \end{bmatrix} = \begin{bmatrix} 1 & \alpha_{i_1} & \cdots & \alpha_{i_1}^{t-1} \\ 1 & \alpha_{i_2} & \cdots & \alpha_{i_2}^{t-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_{i_t} & \cdots & \alpha_{i_t}^{t-1} \end{bmatrix} \begin{bmatrix} U_n \\ E_n^{(1)} \\ \vdots \\ E_n^{(t-1)} \end{bmatrix}.$$

Notice that the decoder can always compute U_n because the matrix on the right side is the Vandelmonde matrix and $\alpha_1, \alpha_2, \ldots, \alpha_m$ are assumed to be distinct. Then, the decoder outputs a unique $\hat{X}^n \in \mathcal{S}_n$ satisfying $U_n = \varphi_n(\hat{X}^n)$.

2) If $W_n^{(i_1,\ldots,i_t)} = (p_n, \ldots, p_n)$, the decoder outputs an element $x_0^n \notin \mathcal{S}_n$ determined in advance.

- *Evaluation of the Decoding Error Probability:*

From the above definition of f_n and g_n, if $X^n \in \mathcal{S}_n$, X^n is correctly decoded from $W_n^{(i_1,\ldots,i_t)}$ for any $\{i_1, \ldots, i_t\} \subset \mathcal{P}$. This fact guarantees (46). In addition, if $X^n \notin \mathcal{S}_n$, the encoder always outputs $(p_n, p_n, \ldots, p_n)$ and therefore the decoder always declares an error. Therefore, it holds that $\varepsilon_n^{(i_1,\ldots,i_t)} = \Pr\{X^n \notin \mathcal{S}_n\} = 1 - \tau_n$, which establishes (48).

- *Evaluation of Security of Shares:*

We first prove that

$$P_{W_n^{(i_1,\ldots,i_{t-1})}|X^n}(w_n^{(i_1,\ldots,i_{t-1})}|x^n) = p_n^{-(t-1)} \quad \text{for all } x^n \in \mathcal{S}_n \text{ and } w_n^{(i_1,\ldots,i_{t-1})} \in \boldsymbol{Z}_{p_n}^{t-1}. \tag{50}$$

Notice that, in view of the definition of the encoder, it holds that

$$w_n^{(i_j)} = \varphi_n(x^n) + e_n^{(1)}\alpha_{i_j} + e_n^{(2)}\alpha_{i_j}^2 + \cdots + e_n^{(t-1)}\alpha_{i_j}^{t-1}, \quad j = 1, 2, \ldots, t-1$$

for some $e_n = (e_n^{(1)}, \ldots, e_n^{(t-1)}) \in \mathcal{E}_n$. Then, since α_i, $i = 1, 2, \ldots, m$, are assumed to be distinct nonzero elements in $\boldsymbol{Z}_{p_n}$, such e_n is determined as a unique solution to the following system of linear equation:

$$\begin{bmatrix} w_n^{(i_1)} - \varphi_n(x^n) \\ w_n^{(i_2)} - \varphi_n(x^n) \\ \vdots \\ w_n^{(i_{t-1})} - \varphi_n(x^n) \end{bmatrix} = \begin{bmatrix} \alpha_{i_1} & \alpha_{i_1}^2 & \cdots & \alpha_{i_1}^{t-1} \\ \alpha_{i_2} & \alpha_{i_2}^2 & \cdots & \alpha_{i_2}^{t-1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{i_{t-1}} & \alpha_{i_{t-1}}^2 & \cdots & \alpha_{i_{t-1}}^{t-1} \end{bmatrix} \begin{bmatrix} e_n^{(1)} \\ e_n^{(2)} \\ \vdots \\ e_n^{(t-1)} \end{bmatrix}. \tag{51}$$

Here, we have used the fact that the determinant of the matrix on the right side of (51) is equal to $(\prod_{j=1}^{t-1} \alpha_{i_j})(\prod_{1 \le j < k \le t-1}(\alpha_{i_k} - \alpha_{i_j})) \neq 0$. Thus, we have (50) due to the uniformity of E_n.

Next, we evaluate $P_{W_n^{(i_1,\ldots,i_{t-1})}}(w_n^{(i_1,\ldots,i_{t-1})})$ for each $w_n^{(i_1,\ldots,i_{t-1})} \in \boldsymbol{Z}_{p_n}^{t-1}$. Since $f_n^{(i_1,\ldots,i_{t-1})}(x^n, e_n) = (p_n, \ldots, p_n) \notin \boldsymbol{Z}_{p_n}^{t-1}$ for all $x^n \notin \mathcal{S}_n$ and $e_n \in \mathcal{E}_n$, we have

$$P_{W_n^{(i_1,\ldots,i_{t-1})}|X^n}(w_n^{(i_1,\ldots,i_{t-1})}|x^n) = 0 \quad \text{for all } x^n \notin \mathcal{S}_n \text{ and } w_n^{(i_1,\ldots,i_{t-1})} \in \boldsymbol{Z}_{p_n}^{t-1}. \tag{52}$$

Therefore, for all $w_n^{(i_1,\ldots,i_{t-1})} \in \boldsymbol{Z}_{p_n}^{t-1}$ it holds that

$$\begin{aligned} P_{W_n^{(i_1,\ldots,i_{t-1})}}(w_n^{(i_1,\ldots,i_{t-1})}) &= \sum_{x^n \in \mathcal{X}^n} P_{X^n}(x^n) P_{W_n^{(i_1,\ldots,i_{t-1})}|X^n}(w_n^{(i_1,\ldots,i_{t-1})}|x^n) \\ &= \sum_{x^n \in \mathcal{S}_n} P_{X^n}(x^n)\, p_n^{-(t-1)} \\ &= \tau_n p_n^{-(t-1)} \end{aligned} \tag{53}$$

where the second equality follows from (50) and (52) and the third equality follows from (43).

We are ready to prove (47). We first note that (50) and (53) guarantee

$$\begin{aligned} &\log \frac{P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(x^n|w_n^{(i_1,\ldots,i_{t-1})})}{P_{X^n}(x^n)} \\ &= \log \frac{P_{W_n^{(i_1,\ldots,i_{t-1})}|X^n}(w_n^{(i_1,\ldots,i_{t-1})}|x^n)}{P_{W_n^{(i_1,\ldots,i_{t-1})}}(w_n^{(i_1,\ldots,i_{t-1})})} = \log \frac{1}{\tau_n} \\ &\qquad \text{for all } (x^n, w_n^{(i_1,\ldots,i_{t-1})}) \in \mathcal{S}_n \times \boldsymbol{Z}_{p_n}^{t-1} \end{aligned} \tag{54}$$

where the first equality in (54) follows from Bayes' formula. Notice that (54) holds for all $x^n \in \mathcal{S}_n$ and $w_n^{(i_1,\ldots,i_{t-1})} \in \boldsymbol{Z}_{p_n}^{t-1}$. In addition, in view of the definition of the encoder, $w_n^{(i_1,\ldots,i_{t-1})} \in \boldsymbol{Z}_{p_n}^{t-1}$ if and only if $x^n \in \mathcal{S}_n$. This argument establishes (47). □

Theorem 3 immediately yields the following corollary that corresponds to the direct counterpart of Theorem 2.

Corollary 2: Let a general source $\boldsymbol{X} = \{X^n\}_{n=1}^\infty$ be given. Assume that there exists a sequence of prime numbers $\{p_n\}_{n=1}^\infty$ satisfying $p_n > m$ for all $n \ge 1$ and

$$\lim_{n \to \infty} \Pr\{X^n \in \mathcal{S}_n\} = 1. \tag{55}$$

Define $\mathcal{E}_n = \{0, 1, \ldots, p_n - 1\}^{t-1}$ and let E_n be the uniformly distributed random variable on $\mathcal{E}_n$. If we construct f_n and g_n in Theorem 3 for all $n \ge 1$, then the sequence $\{(f_n, g_n)\}_{n=1}^\infty$ realizes the (t, m)-threshold scheme for $\boldsymbol{X}$.

Proof: Notice that the assumption (55) of this corollary guarantees that τ_n in (43) goes to one as $n \to \infty$. Hence, (48) and (55) guarantee that $\{(f_n, g_n)\}_{n=1}^\infty$ in Theorem 3 satisfies condition G1) in Definition 2. In addition, since $\tau_n \to 1$ implies $\log \frac{1}{\tau_n} \to 0$ as $n \to \infty$, for any constant $\gamma > 0$ $\log \frac{1}{\tau_n} \le \gamma$ if n is sufficiently large. Hence, (47) and (55) guarantee that

$$\begin{aligned} &\Pr\left\{ \log \frac{P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(X^n|W_n^{(i_1,\ldots,i_{t-1})})}{P_{X^n}(X^n)} \le \gamma \right\} \\ &\ge \Pr\{X^n \in \mathcal{S}_n\} \times \\ &\Pr\left\{ \log \frac{P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(X^n|W_n^{(i_1,\ldots,i_{t-1})})}{P_{X^n}(X^n)} \le \gamma \;\middle|\; X^n \in \mathcal{S}_n \right\} \\ &\to 1 \quad \text{as } n \to \infty \end{aligned}$$

which implies that $\{(f_n, g_n)\}_{n=1}^{\infty}$ in Theorem 3 satisfies condition G2) in Definition 2 □

In Theorem 3, it is immediate from (43) that for each $n \geq 1$, the greater p_n we choose, the greater τ_n becomes. This fact follows since

$$\{x^n \in \mathcal{X}^n : P_{X^n}(x^n) \geq 1/p_n\} \subset \{x^n \in \mathcal{X}^n : P_{X^n}(x^n) \geq 1/p'_n\}$$

for any $p_n < p'_n$. However, both Theorem 3 and Corollary 2 do not suggest how p_n should be large for meeting conditions G1) and G2) in Definition 2. The following corollary gives an intuition about the choice of p_n.

Corollary 3: Let a general source $\boldsymbol{X} = \{X^n\}_{n=1}^{\infty}$ be given. Assume that there exists a sequence $\{p_n\}_{n=1}^{\infty}$ of prime numbers satisfying $\frac{1}{n}\log p_n \to p^* > 0$ as $n \to \infty$ and $\overline{H}(\boldsymbol{X}) < p^*$. Then, there exists a sequence $\{(f_n, g_n)\}_{n=1}^{\infty}$ of encoders and decoders realizing the (t, m)-threshold scheme satisfying

$$\lim_{n\to\infty} \frac{1}{n} \log |\mathcal{W}_n^{(i)}| = p^* \quad \text{for all } i = 1, 2, \ldots, m \tag{56}$$

and

$$\lim_{n\to\infty} \frac{1}{n} \log |\mathcal{E}_n| = (t-1)p^*. \tag{57}$$

Proof: In the proof of Corollary 2, we have already seen that (55) is a sufficient condition that $\{(f_n, g_n)\}_{n=1}^{\infty}$ in Theorem 3 realizes the (t, m)-threshold scheme in the sense of Definition 2. Thus, we first establish (55) below.

By the assumption of the corollary, we can choose a constant $\gamma_0 > 0$ satisfying $\overline{H}(\boldsymbol{X}) \leq p^* - 2\gamma_0$. Since $\frac{1}{n}\log p_n \geq p^* - \gamma_0$ for all sufficiently large n, we have $\overline{H}(\boldsymbol{X}) \leq \frac{1}{n}\log p_n - \gamma_0$ for all sufficiently large n. It is also important to notice that we have

$$\Pr\left\{\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} \leq \overline{H}(\boldsymbol{X}) + \gamma_0\right\} \to 1 \quad \text{as } n \to \infty$$

owing to the definition of $\overline{H}(\boldsymbol{X})$. Therefore, it follows that

$$\Pr\left\{\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} \leq \frac{1}{n}\log p_n\right\} \to 1 \quad \text{as } n \to \infty,$$

which is equivalent to (55).

To complete the proof, we must prove $p_n > m$ for all sufficiently large n because this property is not assumed in the statement of the corollary. However, $p_n > m$ is trivial owing to the assumption of this corollary because p_n grows in exponentially in n while m is a constant. □

We can also evaluate the mutual information $I(X^n; W_n^{(i_1,\ldots,i_j)})$ of the encoding and decoding in the proof of Theorem 3.

Corollary 4: The sequence $\{(f_n, g_n)\}_{n=1}^{\infty}$ in Corollary 2 satisfies

$$\lim_{n\to\infty} I(X^n; W_n^{(i_1,\ldots,i_j)}) = 0$$

for all $j = 1, 2, \ldots, t-1$ and $\{i_1, \ldots, i_j\} \subset \mathcal{P}$.

Proof: Since the chain rule of the mutual information tells us that

$$\begin{aligned} &I(X^n; W_n^{(i_1,\ldots,i_{t-1})}) \\ &= I(X^n; W_n^{(i_1,\ldots,i_j)}) + I(X^n; W_n^{(i_{j+1},\ldots,i_{t-1})} | W_n^{(i_1,\ldots,i_j)}) \\ &\geq I(X^n; W_n^{(i_1,\ldots,i_j)}) \geq 0 \end{aligned}$$

for all $j = 1, 2, \ldots, t-2$, it suffices to prove that $I(X^n; W_n^{(i_1,\ldots,i_{t-1})}) \to 0$ as $n \to \infty$. Let $\mathcal{S}_n$ be the set defined in (44). We first note that $I(X^n; W_n^{(i_1,\ldots,i_{t-1})})$ can be written in the following form:

$$\begin{aligned} &I(X^n; W_n^{(i_1,\ldots,i_{t-1})}) \\ &= \sum_{x^n \in \mathcal{S}_n} \sum_{w_n^{(i_1,\ldots,i_{t-1})} \in \boldsymbol{Z}_{p_n}^{t-1}} P_{X^n W_n^{(i_1,\ldots,i_{t-1})}}(x^n, w_n^{(i_1,\ldots,i_{t-1})}) \\ &\quad \log\frac{P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(x^n | w_n^{(i_1,\ldots,i_{t-1})})}{P_{X^n}(x^n)} \\ &\quad + \sum_{x^n \notin \mathcal{S}_n} P_{X^n W_n^{(i_1,\ldots,i_{t-1})}}(x^n, \pi_n) \\ &\quad \log\frac{P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(x^n | \pi_n)}{P_{X^n}(x^n)} \end{aligned} \tag{58}$$

where $\pi_n = (p_n, p_n, \ldots, p_n) \in \mathcal{W}_n^{(i_1,\ldots,i_{t-1})}$ and we have used the facts that $P_{W_n^{(i_1,\ldots,i_{t-1})}|X^n}(w_n^{(i_1,\ldots,i_{t-1})} | x^n) = 0$ for all $x^n \in \mathcal{S}_n$ and $w_n^{(i_1,\ldots,i_{t-1})} \in \mathcal{W}_n^{(i_1,\ldots,i_{t-1})} \backslash \boldsymbol{Z}_{p_n}^{t-1}$ and $P_{W_n^{(i_1,\ldots,i_{t-1})}|X^n}(\pi_n | x^n) = 1$ for all $x^n \notin \mathcal{S}_n$.

We first evaluate the terms on the right side of (58) separately. In view of (54), the first term on the right side of (58) is evaluated in the following way:

$$\begin{aligned} &\sum_{x^n \in \mathcal{S}_n} \sum_{w_n^{(i_1,\ldots,i_{t-1})} \in \boldsymbol{Z}_{p_n}^{t-1}} P_{X^n W_n^{(i_1,\ldots,i_{t-1})}}(x^n, w_n^{(i_1,\ldots,i_{t-1})}) \\ &\qquad \times \log\frac{P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(x^n | w_n^{(i_1,\ldots,i_{t-1})})}{P_{X^n}(x^n)} \\ &\quad = \sum_{x^n \in \mathcal{S}_n} \sum_{w_n^{(i_1,\ldots,i_{t-1})} \in \boldsymbol{Z}_{p_n}^{t-1}} P_{X^n W_n^{(i_1,\ldots,i_{t-1})}}(x^n, w_n^{(i_1,\ldots,i_{t-1})}) \\ &\qquad \times \left(\log\frac{1}{\tau_n}\right) \\ &\quad = \sum_{x^n \in \mathcal{S}_n} P_{X^n}(x^n) \left(\log\frac{1}{\tau_n}\right) \\ &\quad = \tau_n \log\frac{1}{\tau_n} \end{aligned} \tag{59}$$

where τ_n is defined in (43) and the second equality holds because every $x^n \in \mathcal{S}_n$ is mapped to elements of $\boldsymbol{Z}_{p_n}$ with probability 1. On the other hand, by using Bayes' formula and $P_{W_n^{(i_1,\ldots,i_{t-1})}|X^n}(\pi_n | x^n) = 1$ for all $x^n \notin \mathcal{S}_n$, the second term on the right side of (58) can be evaluated as follows:

$$
\begin{aligned}
&\sum_{x^n \notin \mathcal{S}_n} P_{X^n W_n^{(i_1,\ldots,i_{t-1})}}(x^n,\pi_n)\log\frac{P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(x^n|\pi_n)}{P_{X^n}(x^n)}\\
&\quad=\sum_{x^n\notin\mathcal{S}_n}P_{X^n}(x^n)P_{W_n^{(i_1,\ldots,i_{t-1})}|X^n}(\pi_n|x^n)\\
&\qquad\times\log\frac{P_{W_n^{(i_1,\ldots,i_{t-1})}|X^n}(\pi_n|x^n)}{P_{W_n^{(i_1,\ldots,i_{t-1})}}(\pi_n)}\\
&\quad=\sum_{x^n\notin\mathcal{S}_n}P_{X^n}(x^n)\log\frac{1}{\Pr\{X^n\notin\mathcal{S}_n\}}\\
&\quad=(1-\tau_n)\log\frac{1}{1-\tau_n}
\end{aligned}
\tag{60}
$$

where the second and the third inequalities follow from $P_{W_n^{(i_1,\ldots,i_{t-1})}|X^n}(\pi_n|x^n) = 1$ for all $x^n \notin \mathcal{S}_n$ and $P_{W_n^{(i_1,\ldots,i_{t-1})}|X^n}(\pi_n|x^n) = 0$ for all $x^n \in \mathcal{S}_n$, respectively.

Therefore, the combination of (58), (59) and (60) yields

$$
I(X^n;W_n^{(i_1,\ldots,i_{t-1})}) = \tau_n\log\frac{1}{\tau_n}+(1-\tau_n)\log\frac{1}{1-\tau_n}. \tag{61}
$$

Since (55) guarantees that $\tau_n \to 1$ as $n\to\infty$, we have the claim of this corollary from (61). □

So far, we have constructed the (t,m)-threshold scheme satisfying conditions G1) and G2) by using Theorem 3. However, Theorem 3 is more involving. Instead of condition G1), we can also consider the case where the decoding error probability does not go to zero as $n\to\infty$ but is asymptotically upper bounded by some $\delta\in[0,1)$. We can expect that the sizes of shares are reduced in such a case. We can obtain the following corollary from Theorem 3 that is an extension of Corollary 2 and is related to ε-source coding in [6].

Corollary 5: Let a general source $\boldsymbol{X}=\{X^n\}_{n=1}^\infty$ be given. For an arbitrarily fixed $\delta\in[0,1)$ assume that there exists a sequence $\{p_n\}_{n=1}^\infty$ of prime numbers satisfying

$$
\liminf_{n\to\infty}\Pr\{\log(p_nP_{X^n}(X^n))\ge 0\}\ge 1-\delta. \tag{62}
$$

Define $\mathcal{E}_n=\{0,1,\ldots,p_n-1\}^{t-1}$ and let E_n be the uniformly distributed random variable on $\mathcal{E}_n$. Then, we can construct a sequence $\{(f_n,g_n)\}_{n=1}^\infty$ of encoders and decoders satisfying (45) and (46) for all $n\ge 1$ and

$$
\begin{aligned}
\lim_{n\to\infty}\Pr\Bigg\{&\frac{1}{n^\alpha}\log\frac{P_{X^n|W_n^{(i_1,\ldots,i_{t-1})}}(X^n|W_n^{(i_1,\ldots,i_{t-1})})}{P_{X^n}(X^n)}\le\gamma\\
&\Bigg| X^n\in\mathcal{S}_n\Bigg\}=1\quad\text{for all }\{i_1,\ldots,i_{t\text{-}1}\}\subset\mathcal{P}
\end{aligned}
\tag{63}
$$

for any constants $\gamma>0$ and $\alpha>0$. In addition, such $\{(f_n,g_n)\}_{n=1}^\infty$ satisfies

$$
\limsup_{n\to\infty}\varepsilon_n^{(i_1,\ldots,i_t)}\le\delta\quad\text{for all }\{i_1,\ldots,i_t\}\subset\mathcal{P}. \tag{64}
$$

Proof: Equation (63) easily follows from (47) because $\frac{1}{n^\alpha}\to 0$ as $n\to\infty$ and $\tau_n=\Pr\{X^n\in\mathcal{S}_n\}$ is bounded away from zero for all sufficiently large n owing to (62) and $\delta\in[0,1)$. In addition, (64) immediately follows from (48) and (62). □

We conclude this section by giving a complete answer to the question given at the beginning of Section IV.

Example 1: Consider the case where $\boldsymbol{X}=\{X^n\}_{n=1}^\infty$ is a stationary memoryless source with a countably infinite alphabet $\mathcal{X}$. Recall that both $\overline{H}(\boldsymbol{X})$ and $\underline{H}(\boldsymbol{X})$ coincide with the entropy $H(X)$ of the source if $H(X)<\infty$. Hereafter, assume that $H(X)<\infty$.

We can construct the (t,m)-threshold scheme in the sense of Definition 2 by using Theorem 3 and Corollary 2. Fix a small constant $\gamma_0>0$ arbitrarily and define the typical set $\mathcal{A}_n$ by (10), where we use γ_0 instead of γ. Then, it holds that $\Pr\{X^n\in\mathcal{A}_n\}\to 1$ as $n\to\infty$ [5]. In addition, we have $2^{-n(H(X)+\gamma_0)}\le P_{X^n}(x^n)\le 2^{-n(H(X)-\gamma_0)}$ for all $x^n\in\mathcal{A}_n$ and $n\ge 1$ from the definition of $\mathcal{A}_n$. In order to construct an encoder and a decoder of the (t,m)-threshold scheme, we arbitrarily choose a prime number p_n satisfying $p_n\ge 2^{n(H(X)+\gamma_0)}$. Then, we have (55) because it holds that

$$
\log(p_nP_{X^n}(x^n))\ge n(H(X)+\gamma_0)-n(H(X)+\gamma_0)=0
$$

for all $x^n\in\mathcal{A}_n$ and $\Pr\{X^n\in\mathcal{A}_n\}\to 1$ as $n\to\infty$. Hence, Corollary 2 guarantees that $\{(f_n,g_n)\}_{n=1}^\infty$ in Theorem 3 realizes the (t,m)-threshold scheme. In particular, if we can choose p_n, $n\ge 1$, satisfying $\frac{1}{n}\log p_n\to H(X)+\gamma_0$ as $n\to\infty$, Corollary 3 tells us that

$$
\begin{aligned}
&\lim_{n\to\infty}\frac{1}{n}\log|\mathcal{W}_n^{(i)}|=H(X)+\gamma_0\quad\text{for all }i=1,2,\ldots,m\\
&\text{and}\lim_{n\to\infty}\frac{1}{n}\log|\mathcal{E}_n|=(t-1)(H(X)+\gamma_0).
\end{aligned}
$$

The right sides can be arbitrarily close to $H(X)$ and $(t-1)H(X)$, respectively, because we can choose an arbitrarily small $\gamma_0>0$.

We can give an impossibility result on the rate of E_n by using Theorem 2. We can actually prove that if

$$
\liminf_{n\to\infty}\frac{1}{n}\log|\mathcal{E}_n|<(t-1)H(X) \tag{65}
$$

then there is no $\{(f_n,g_n)\}_{n=1}^\infty$ that realizes the (t,m)-threshold scheme satisfying conditions G1) and G2) in Definition 2. This fact is proved by a contradiction argument. First, notice that (65) guarantees the existence of a small constant $\gamma_0'>0$ and a subsequence $\{n_i\}_{i=1}^\infty$ such that $\frac{1}{n_i}\log|\mathcal{E}_{n_i}|\le(t-1)(H(X)-2\gamma_0')$ for all $i\ge 1$. Then, letting $\mathcal{A}_n$ be the typical set in (10) with $\gamma=\gamma_0'$, it holds that

$$
\begin{aligned}
&\frac{1}{n_i}\log(|\mathcal{E}_{n_i}|(P_{X^{n_i}}(x^{n_i}))^{t-1})\\
&\quad\le(t-1)(H(X)-2\gamma_0')-(t-1)(H(X)-\gamma_0')\\
&\quad=-(t-1)\gamma_0'<0\quad\text{for all }x^n\in\mathcal{A}_n\text{ and }i\ge 1.
\end{aligned}
\tag{66}
$$

In addition, we also have $\Pr\{X^n \in \mathcal{A}_n\} \to 1$ as $n \to \infty$. Thus, we can conclude

$$\operatorname*{p\text{-}lim\,inf}_{n\to\infty} \frac{1}{n} \log(|\mathcal{E}_n|(P_{X^n}(X^n))^{t-1}) < 0$$

because of the property of limit inferior in probability given in Fact 1. That is, we cannot prove $\Pr\{\frac{1}{n}\log(|\mathcal{E}_n|(P_{X^n}(X^n))^{t-1}) \geq -\gamma\} \to 1$ as $n \to \infty$ for all γ satisfying $\gamma < (t-1)\gamma_0$ owing to (66).

Example 2: Next, let us consider the case where $\boldsymbol{X} = \{X^n\}_{n=1}^\infty$ is a mixed source [6]. For each $i = 1, 2$ let X_i be a random variable on $\mathcal{X}$ and X_i^n be n i.i.d. copies of X_i. Denote by $P_{X_i^n}$ the probability distribution of X_i^n. We call $\boldsymbol{X} = \{X^n\}_{i=1}^\infty$ a mixed source if P_{X^n} is defined by

$$P_{X^n}(x^n) = (1-\eta)P_{X_1^n}(x^n) + \eta P_{X_2^n}(x^n)$$

for all $x^n \in \mathcal{X}^n$ and $n \geq 1$, where $\eta \in (0,1)$ is a constant. We assume that the entropies of X_1 and X_2 satisfy $H(X_1) < H(X_2) < \infty$. It is known that $\overline{H}(\boldsymbol{X}) = H(X_2)$ and $\underline{H}(\boldsymbol{X}) = H(X_1)$ for this mixed source [6].

Given a sequence $\{p_n\}_{n=1}^\infty$ of prime numbers we consider the same sequence $\{(f_n, g_n)\}_{n=1}^\infty$ of encoders and decoders as in Corollary 2. Assume that the limit of $\frac{1}{n}\log p_n$, $n \geq 1$, exists. Letting p^* denote the limit, we can obtain the following three facts:

Case 1) $H(X_2) < p^*$
Corollary 3 guarantees that $\{(f_n, g_n)\}_{n=1}^\infty$ realizes the (t,m)-threshold scheme satisfying G1), G2), (56) and (57).

Case 2) $H(X_1) < p^* < H(X_2)$
Corollary 5 guarantees that $\{(f_n, g_n)\}_{n=1}^\infty$ satisfies (46), (56), (57) and (63). Since the law of large numbers tells us that $\Pr\{X^n \in \mathcal{S}_n\} \to 1-\eta$ as $n \to \infty$, it holds that

$$\lim_{n\to\infty} \varepsilon_n^{(i_1,\ldots,i_t)} = \eta \quad \text{for all } \{i_1,\ldots,i_t\} \subset \mathcal{P} \tag{67}$$

where $\mathcal{S}_n$ is defined in (44).

Case 3) $p^* < H(X_1)$
Since $\Pr\{X^n \in \mathcal{S}_n\} \to 0$ as $n \to \infty$ for this case, it holds that $\varepsilon_n^{(i_1,\ldots,i_t)} \to 1$ as $n \to \infty$ for all $\{i_1,\ldots,i_t\} \subset \mathcal{P}$.

Notice that in Case 2), we can realize a scheme similar to the (t,m)-threshold scheme with smaller sizes of shares than the scheme in Case 1) and the decoding error probability close to η. This kind of scheme can be better if η is small enough and small decoding error is permissible.

VI. Application to Shannon's Cipher System

In this section we consider new coding theorems for Shannon's cipher system given in Fig. 2. In Fig. 2, for each $n \geq 1$ $X^n \in \mathcal{X}^n$ denotes n outputs from a general source, where $\mathcal{X}$ is a finite or a countably infinite alphabet. Let $E_n \in \mathcal{E}_n$ be a key shared by an encoder and a decoder in advance. Assume that E_n is independent of X^n for each $n \geq 1$. Given n source outputs X^n and a key E_n, an encoder generates a cryptogram $W_n \in \mathcal{W}_n$, where $\mathcal{W}_n$ is a set of cryptograms. In this section, we consider *stochastic* encoders. While a *deterministic* decoder is defined as a mapping $f_n : \mathcal{X}^n \times \mathcal{E}_n \to \mathcal{W}_n$, a stochastic encoder is identified as a conditional probability distribution $P_{W_n|X^nE_n}$. That is, given an $(x^n, e_n) \in \mathcal{X}^n \times \mathcal{E}_n$, the stochastic encoder outputs a cryptogram W_n randomly subject to $P_{W_n|X^nE_n}(\cdot|x^n, e_n)$. We can regard a deterministic encoder f_n as one of stochastic encoders that satisfies $P_{W_n|X^nE_n}(w_n|x^n, e_n) = 1$ if $w_n = f_n(w^n, e_n)$ and $P_{W_n|X^nE_n}(w_n|x^n, e_n) = 0$ otherwise for each $(x^n, e_n) \in \mathcal{X}^n \times \mathcal{E}_n$. Stochastic encoders are also denoted by f_n when there is no confusion. Throughout this section, the term "encoder" means a stochastic encoder unless we mention that an encoder is deterministic.

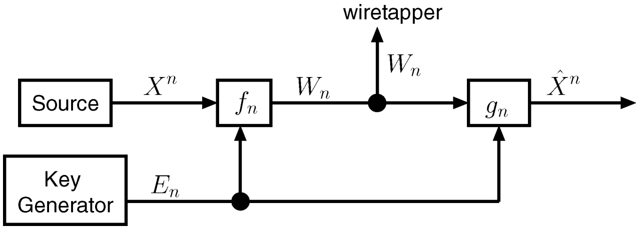

Fig. 2. Block diagram of Shannon's cipher system.

On the other hand, we consider only deterministic decoders. A decoder is defined as a mapping $g_n : \mathcal{W}_n \times \mathcal{E}_n \to \mathcal{X}^n$. When a cryptogram W_n is transmitted, a decoder decrypts W_n to $\hat{X}^n \in \mathcal{X}^n$ under a key E_n. Given an encoder f_n and a decoder g_n the decoding error probability ε_n is defined as

$$\varepsilon_n = \Pr\{\hat{X}^n \neq X^n\}$$

which can be written as

$$\varepsilon_n = \sum_{x^n \in \mathcal{X}^n} \sum_{e_n \in \mathcal{E}_n} \sum_{w_n \in \mathcal{W}_n} P_{X^n}(x^n) P_{E_n}(e_n) \times P_{W_n|X^nE_n}(w_n|x^n, e_n)\overline{\chi}(x^n, e_n, w_n) \tag{68}$$

where

$$\overline{\chi}(x^n, e_n, w_n) = \begin{cases} 1, & x^n \neq g_n(w_n, e_n) \\ 0, & \text{otherwise} \end{cases}$$

and $P_{W_n|X^nE_n}(w_n|x^n, e_n)$ in (68) is determined by the encoder f_n.

A wiretapper, who observes the cryptogram W_n but has no information on E_n, wants to know something about X^n. We should construct f_n and g_n so that the wiretapper can obtain almost no information on X^n from W_n while a receiver can decrypt a cryptogram with negligible decoding error probability.

We define encoders and decoders with the perfect secrecy as follows:

Definition 3: If a sequence $\{(f_n, g_n)\}_{n=1}^\infty$ of encoders and decoders satisfies the following S1) and S2), we say that $\{(f_n, g_n)\}_{n=1}^\infty$ realizes the perfect secrecy:

S1)

$$\lim_{n\to\infty} \varepsilon_n = 0$$

S2)

$$\operatorname*{p\text{-}lim\,sup}_{n\to\infty} \log \frac{P_{X^n|W_n}(X^n|W_n)}{P_{X^n}(X^n)} \leq 0.$$

Condition S1) requires that the decoding error probability goes to zero as $n \to \infty$. On the other hand, condition S2) guarantees the security of the system. That is, S2) guarantees that W_n is almost independent on X^n for all sufficiently large n in the same sense as G2) in Section IV.

We have the following converse theorem on Shannon's cipher system with the perfect secrecy. Notice in this theorem the uniformity of E_n is not assumed.

Theorem 4: Given a general source $\boldsymbol{X} = \{X^n\}_{n=1}^{\infty}$, let $\{(f_n, g_n)\}_{n=1}^{\infty}$ be any sequence of encoders and decoders realizing the perfect secrecy. Then, for any constant $\alpha > 0$ it holds that

$$\text{p-}\liminf_{n\to\infty} \frac{1}{n^\alpha} \log \frac{P_{X^n}(X^n)}{P_{W_n}(W_n)} \geq 0 \tag{69}$$

$$\text{p-}\liminf_{n\to\infty} \frac{1}{n^\alpha} \log \frac{P_{X^n}(X^n)}{P_{E_n}(E_n)} \geq 0. \tag{70}$$

Proof: In the proof we use the same methods as in the proof of Theorem 2. However, we give the proof of Theorem 4 here for not only establishing the new result (70) but also clear understanding of the main contributions, Theorem 2 and Theorem 4, of this paper.

First, by using an argument similar to the proof of Lemma 4, we can obtain

$$\text{p-}\limsup_{n\to\infty} \log \frac{1}{P_{X^n|W_nE_n}(X^n|W_n, E_n)} \leq 0 \tag{71}$$

for any $\{(f_n, g_n)\}_{n=1}^{\infty}$ satisfying S1). See Appendix D for the proof of (71). In addition, by applying the same method given in the proof of Lemma 5, we can easily obtain

$$\text{p-}\liminf_{n\to\infty} \frac{1}{n^\alpha} \log \frac{P_{W_n|E_n}(W_n|E_n)}{P_{W_n}(W_n)} \geq 0 \tag{72}$$

for any constant $\alpha > 0$. Letting $\alpha > 0$ and $\gamma > 0$ be arbitrary constants, define

$$\mathcal{D}_n = \left\{ (x^n, e_n, w_n) \in \mathcal{X}^n \times \mathcal{E}_n \times \mathcal{W}_n : \frac{1}{n^\alpha} \log \frac{1}{P_{X^n|W_nE_n}(x^n|w_n, e_n)} \leq \gamma \right\}$$

$$\mathcal{V}_n = \left\{ (x^n, e_n, w_n) \in \mathcal{X}^n \times \mathcal{E}_n \times \mathcal{W}_n : \frac{1}{n^\alpha} \log \frac{P_{W_n|E_n}(w_n|e_n)}{P_{W_n}(w_n)} \geq -\gamma \right\}.$$

Then, it follows from (71), (72) and Facts 1 and 2 that

$$\Pr\{(X^n, E_n, W_n) \in \mathcal{D}_n \cap \mathcal{V}_n\} \to 1 \quad \text{as } n \to \infty. \tag{73}$$

Since we have

$$\frac{1}{n^\alpha} \log \frac{P_{W_n|E_n}(w_n|e_n)}{P_{W_n}(w_n)} + \frac{1}{n^\alpha} \log \frac{1}{P_{W_n|X^nE_n}(w_n|x^n, e_n)} \geq -\gamma \quad \text{for all } (x^n, e_n, w_n) \in \mathcal{D}_n \cap \mathcal{V}_n \tag{74}$$

by taking Bayes' formula

$$P_{X^nW_n|E_n} = P_{X^n|E_n} P_{W_n|X^nE_n} = P_{W_n|E_n} P_{X^n|W_nE_n}$$

and independence of X^n and E_n into consideration, it follows from (74) that

$$\frac{1}{n^\alpha} \log \frac{P_{X^n}(x^n)}{P_{W_n}(w_n)} \geq -2\gamma \quad \text{for all } (x^n, e_n, w_n) \in \mathcal{D}_n \cap \mathcal{V}_n. \tag{75}$$

Since $\gamma > 0$ is arbitrary, (73) and (75) imply (69).

Next, we prove (70). Setting

$$\mathcal{U}_n = \left\{ (x^n, e_n, w_n) \in \mathcal{X}^n \times \mathcal{E}_n \times \mathcal{W}_n : \frac{1}{n^\alpha} \log \frac{P_{X^n|W_n}(x^n|w_n)}{P_{X^n}(x^n)} \leq \gamma \right\}$$

in view of condition S2) and (73) it holds that

$$\Pr\{(X^n, E_n, W_n) \in \mathcal{D}_n \cap \mathcal{U}_n \cap \mathcal{V}_n\} \to 1 \quad \text{as } n \to \infty. \tag{76}$$

In addition, since we can rewrite $\mathcal{V}_n$ as

$$\mathcal{V}_n = \left\{ (x^n, e_n, w_n) \in \mathcal{X}^n \times \mathcal{E}_n \times \mathcal{W}_n : \frac{1}{n^\alpha} \log \frac{P_{E_n|W_n}(e_n|w_n)}{P_{E_n}(e_n)} \geq -\gamma \right\}$$

it holds that

$$\frac{1}{n^\alpha} \log \frac{P_{E_n|W_n}(e_n|w_n)}{P_{E_n}(e_n)} + \frac{1}{n^\alpha} \log \frac{1}{P_{E_n|X^nW_n}(e_n|x^n, w_n)} \geq -\gamma \quad \text{for all } (x^n, e_n, w_n) \in \mathcal{D}_n \cap \mathcal{U}_n \cap \mathcal{V}_n. \tag{77}$$

Application of Bayes' formula

$$P_{X^nE_n|W_n} = P_{E_n|W_n} P_{X^n|W_nE_n} = P_{X^n|W_n} P_{E_n|X^nW_n}$$

to (77) leads to

$$\frac{1}{n^\alpha} \log \left(\frac{P_{X^n}(x^n)}{P_{E_n}(e_n)} \frac{P_{X^n|W_n}(x^n|w_n)}{P_{X^n}(x^n)} \frac{1}{P_{X^n|W_nE_n}(x^n|w_n, e_n)} \right) \geq -\gamma \quad \text{for all } (x^n, e_n, w_n) \in \mathcal{D}_n \cap \mathcal{U}_n \cap \mathcal{V}_n$$

which implies

$$\frac{1}{n^\alpha} \log \frac{P_{X^n}(x^n)}{P_{E_n}(e_n)} \geq -3\gamma \quad \text{for all } (x^n, e_n, w_n) \in \mathcal{D}_n \cap \mathcal{U}_n \cap \mathcal{V}_n. \tag{78}$$

Now, (70) follows from the combination of (76) and (78). □

We also have the direct theorem that corresponds to Theorem 3. In case of Shannon's cipher system the set $\mathcal{W}_n$ of cryptograms need not contain a finite field as its subset.

Theorem 5: Let a general source $\boldsymbol{X} = \{X^n\}_{n=1}^{\infty}$ be given. For each $n \geq 1$ let q_n be an arbitrary positive integer satisfying $\tau_n \in (0, 1]$, where

$$\tau_n = \Pr\left\{\log(q_n P_{X^n}(X^n)) \geq 0\right\}. \qquad (79)$$

Define $\mathcal{E}_n = \{0, 1, \ldots, q_n - 1\}$ and let E_n be the random variable subject to the uniform distribution on $\mathcal{E}_n$. Set $\mathcal{W}_n = \{0, 1, \ldots, q_n\}$. Then, there exists a deterministic encoder f_n and a decoder g_n satisfying $\varepsilon_n \leq 1 - \tau_n$ and

$$\Pr\left\{\log \frac{P_{X^n|W_n}(X^n|W_n)}{P_{X^n}(X^n)} = \log \frac{1}{\tau_n}\right\} \geq \tau_n. \qquad (80)$$

Proof: This theorem is proved by combination of an ordinary method (e.g., [16]) and the methods given in the proof of Theorem 3. Define

$$\mathcal{S}_n = \{x^n \in \mathcal{X}^n : \log(q_n P_{X^n}(x^n)) \geq 0\}.$$

Then, similarly to (49), we have $|S_n| \leq q_n$. Therefore, there exists a one-to-one mapping $\varphi_n : \mathcal{S}_n \to \tilde{\mathcal{W}}_n \overset{\text{def}}{=} \mathcal{W}_n \backslash \{q_n\}$. We define an encoder and a decoder in the following way:

Encoder $f_n : \mathcal{X}^n \times \mathcal{E}_n \to \mathcal{W}_n$

Given a source output X^n and a key E_n, the encoder outputs

$$W_n = \begin{cases} \varphi_n(X^n) \oplus E_n, & \text{if } X^n \in \mathcal{S}_n \\ q_n, & \text{otherwise} \end{cases}$$

where $\oplus$ denotes the addition of modulo q_n. Note that this encoder is deterministic.

Decoder $g_n : \mathcal{W}^n \times \mathcal{E}_n \to \mathcal{X}^n$

Given a cryptogram W_n and a key E_n, the decoder outputs

$$\hat{X}^n = \begin{cases} \varphi_n^{-1}(W_n \ominus E_n), & \text{if } W_n \in \{0, 1, \ldots, q_n - 1\} \\ x_0^n, & \text{otherwise} \end{cases}$$

where x_0^n is an arbitrary fixed element of $\mathcal{X}^n$, $\ominus$ denotes the subtraction of modulo q_n and $\varphi_n^{-1}(W_n \ominus E_n)$ means a unique element $\hat{X}^n \in \mathcal{S}_n$ satisfying $\varphi_n(\hat{X}^n) = W_n \ominus E_n$.

Since the above encoder and decoder can cause the decoding error if $X^n \notin \mathcal{S}_n$, it is immediate that $\varepsilon_n \leq 1 - \tau_n$.

Hereafter, we prove (80). From the definition of the encoder f_n, it holds that

$$P_{W_n|X^n}(w_n|x^n) = \frac{1}{|\mathcal{E}_n|} \quad \text{for all } (x^n, w_n) \in \mathcal{S}_n \times \tilde{\mathcal{W}}_n. \qquad (81)$$

In addition, since $f_n(x^n, e_n) \notin \tilde{\mathcal{W}}_n$ for all $x^n \notin \mathcal{S}_n$ and $e_n \in \mathcal{E}_n$, we have

$$P_{W_n}(w_n) = \sum_{x^n \in \mathcal{S}_n} P_{X^n}(x^n) \frac{1}{|\mathcal{E}_n|} = \frac{\tau_n}{|\mathcal{E}_n|} \quad \text{for all } w_n \in \tilde{\mathcal{W}}_n \qquad (82)$$

where the second equality follows from (79) and the definition of $\mathcal{S}_n$. Hence, the combination of (81) and (82) yields

$$\log \frac{P_{X^n|W_n}(x^n|w_n)}{P_{X^n}(x^n)} = \log \frac{P_{W_n|X^n}(w_n|x^n)}{P_{W_n}(w_n)} = \log \frac{1}{\tau_n} \quad \text{for all } (x^n, w_n) \in \mathcal{S}_n \times \tilde{\mathcal{W}}_n, \qquad (83)$$

where the first equality (83) follows from Bayes' formula. Since $W_n \in \tilde{\mathcal{W}}_n$ if and only if $X^n \in \mathcal{S}_n$ owing to the definition of the encoder, it is obvious that

$$\Pr\{(X^n, W_n) \in \mathcal{S}_n \times \tilde{\mathcal{W}}_n\} = \Pr\{X^n \in \mathcal{S}_n\} = \tau_n. \qquad (84)$$

Thus, we obtain (80). □

We have two corollaries to Theorem 5 that correspond to Corollary 2 and Corollary 3, respectively.

Corollary 6: Assume that there exists a sequence of positive integers $\{q_n\}_{n=1}^{\infty}$ satisfying

$$\lim_{n\to\infty} \Pr\{\log(q_n P_{X^n}(X^n)) \geq 0\} = 1. \qquad (85)$$

Define $\mathcal{E}_n = \{0, 1, \ldots, q_n - 1\}$ and let E_n be the uniformly distributed random variable on $\mathcal{E}_n$. If we construct f_n and g_n in Theorem 5 for all $n \geq 1$, then the sequence $\{(f_n, g_n)\}_{n=1}^{\infty}$ realizes the perfect secrecy.

Corollary 7: Assume that there exists a sequence $\{q_n\}_{n=1}^{\infty}$ satisfying $\frac{1}{n} \log q_n \to q^* > 0$ as $n \to \infty$ and $\overline{H}(\boldsymbol{X}) < q^*$. Then, there exists a sequence $\{(f_n, g_n)\}_{n=1}^{\infty}$ of encoders and decoders realizing the perfect secrecy and satisfying

$$\lim_{n\to\infty} \frac{1}{n} \log |\mathcal{W}_n| = q^* \text{ and } \lim_{n\to\infty} \frac{1}{n} \log |\mathcal{E}_n| = q^*.$$

It is also easily verified, similarly to Corollary 4, that the sequence $\{(f_n, g_n)\}_{n=1}^{\infty}$ in Corollary 6 satisfies $I(X^n; W_n) \to 0$ as $n \to \infty$.

VII. Application to Fixed-Length Source Coding

In this section, we focus on fixed-length coding of a general source in which the decoding error probability vanishes as $n \to \infty$ first discussed in [7]. In this section we consider a modified version of the problem in which we treat a wider class of encoders than in [7].

Suppose that a general source $\boldsymbol{X} = \{X^n\}_{n=1}^{\infty}$ is given, where $X^n \in \mathcal{X}^n$ for each $n \geq 1$ and $\mathcal{X}$ is a countably infinite alphabet. For each $n \geq 1$ a stochastic encoder encodes n source outputs X^n to a codeword $W_n \in \mathcal{W}_n \overset{\text{def}}{=} \{0, 1, \ldots, M_n - 1\}$, where $\mathcal{W}_n$ is a set of codewords. Here, a stochastic encoder is defined as a conditional probability distribution $P_{W_n|X^n}$. If $X^n = x^n$, then the stochastic encoder generates a codeword $W_n \in \mathcal{W}_n$ randomly subject to $P_{W_n|X^n}(\cdot|x^n)$. The stochastic encoder is denoted by f_n. In a special case where f_n is a mapping from $\mathcal{X}^n$ to $\mathcal{W}_n$, i.e., for every $x^n \in \mathcal{X}^n$ there exists a $w_n \in \mathcal{W}_n$ satisfying $P_{W_n|X^n}(w_n|x^n) = 1$, we call f_n a *deterministic* encoder. On the other hand, a codeword W_n is decoded to $\hat{X}^n \in \mathcal{X}^n$ by a decoder. The decoder is defined as a mapping $g_n : \mathcal{W}_n \to \mathcal{X}^n$. We consider only deterministic decoders. It is important to notice that the fixed-length coding of a general source is a special case of Shannon's cipher system with $|\mathcal{E}_n| = 1$.

Given a stochastic encoder f_n and a decoder g_n, the decoding error probability is defined as

$$\varepsilon_n = \Pr\{\hat{X}^n \neq X^n\}.$$

Note that ε_n can be written in the following form:

$$\varepsilon_n = \sum_{x^n \in \mathcal{X}^n} \sum_{w_n \in \mathcal{W}_n} P_{X^n}(x^n) P_{W_n|X^n}(w_n|x^n) \overline{\chi}(x^n, w_n)$$

where $P_{W_n|X^n}$ is determined by f_n and

$$\overline{\chi}(x^n, w_n) = \begin{cases} 1, & x^n \neq g_n(w_n) \\ 0, & \text{otherwise.} \end{cases}$$

Throughout this section encoders mean stochastic encoders unless we mention that an encoder is deterministic.

We have the following converse theorem that does not have the factor $1/n^\alpha$.

Theorem 6: For any sequence $\{(f_n, g_n)\}_{n=1}^\infty$ of encoders and decoders satisfying $\varepsilon_n \to 0$ as $n \to \infty$, it holds that

$$\text{p-}\liminf_{n\to\infty} \log \frac{P_{X^n}(X^n)}{P_{W_n}(W_n)} \geq 0.$$

Proof: Similarly to (71), we can easily prove

$$\text{p-}\limsup_{n\to\infty} \log \frac{1}{P_{X^n|W_n}(X^n|W_n)} \leq 0. \quad (86)$$

Application of Bayes' formula and (5) to (86) yields

$$\text{p-}\liminf_{n\to\infty} \log \frac{P_{X^n}(X^n) P_{W_n|X^n}(W_n|X^n)}{P_{W_n}(W_n)} \geq 0. \quad (87)$$

Note that (7) and (87) yield

$$\text{p-}\liminf_{n\to\infty} \log \frac{P_{X^n}(X^n)}{P_{W_n}(W_n)} + \text{p-}\limsup_{n\to\infty} \log P_{W_n|X^n}(W_n|X^n) \geq 0. \quad (88)$$

Then, the claim of this theorem follows because the second term on the left side of (88) is nonpositive (notice that $P_{W_n|X^n}(W_n|X^n) \leq 1$ holds with probability 1). □

We also have the direct theorem for fixed-length source coding.

Theorem 7: If there exists a sequence $\{M_n\}_{n=1}^\infty$ of positive integers satisfying

$$\Pr\{\log(M_n P_{X^n}(X^n)) \geq 0\} \to 1 \quad \text{as } n \to \infty$$

then there exists a sequence $\{(f_n, g_n)\}_{n=1}^\infty$ of deterministic encoders and decoders satisfying $|\mathcal{W}_n| = M_n$ for all $n \geq 1$ and $\varepsilon_n \to 0$ as $n \to \infty$.

Proof of Theorem 7 is easy. Setting

$$\mathcal{S}_n = \{x^n \in \mathcal{X}^n : \log(M_n P_{X^n}(X^n)) \geq 0\}$$

the same argument as in (49) leads to $|\mathcal{S}_n| \leq M_n$. Hence, there exists a one-to-one mapping from $\mathcal{S}_n$ to $\mathcal{W}_n$. We use such φ_n as an encoder. Clearly, there exists a decoder that decodes all the elements of $\mathcal{S}_n$ correctly. Though the elements of $\mathcal{X}^n$ not belonging to $\mathcal{S}_n$ may not be correctly decoded, such a probability of decoding error vanishes because $\Pr\{X^n \notin \mathcal{S}_n\} \to 0$ as $n \to \infty$ due to the assumption of the theorem.

We can also give a formula of the infimum-achievable coding rate. We define the infimum-achievable coding rate as follows.

Definition 4: A rate R is called achievable if there exists a sequence $\{(f_n, g_n)\}_{n=1}^\infty$ of encoders and decoders satisfying

$$\limsup_{n\to\infty} \frac{1}{n} \log M_n \leq R \quad (89)$$

$$\lim_{n\to\infty} \varepsilon_n = 0. \quad (90)$$

The infimum of the achievable rate R is called the infimum-achievable coding rate and is denoted by $R_s(\boldsymbol{X})$.

If in Definition 4 we require the encoder f_n, $n \geq 1$, to be deterministic, we have the ordinary definition of the infimum-achievable coding rate. Denote by $R_d(\boldsymbol{X})$ the infimum-achievable coding rate for such a case. Han and Verdú show that $R_d(\boldsymbol{X}) = \overline{H}(\boldsymbol{X})$ [7].

The following theorem gives a general formula of $R_s(\boldsymbol{X})$.

Theorem 8: $R_s(\boldsymbol{X}) = \overline{H}(\boldsymbol{X})$.

Since deterministic encoders can be regarded as stochastic encoders, it immediately follows that $R_s(\boldsymbol{X}) \leq R_d(\boldsymbol{X}) = \overline{H}(\boldsymbol{X})$. Hence, we have only to prove $R_s(\boldsymbol{X}) \geq \overline{H}(\boldsymbol{X})$ for establishing Theorem 8. We use the following lemma for establishing Theorem 8, which was first given in [10]. The proof of this Lemma 7 is given in Appendix E for readers' convenience.

Lemma 7: Let $\boldsymbol{Z} = \{Z_n\}_{n=1}^\infty$ be an arbitrary real-valued random variables satisfying $Z_n \in \mathcal{Z}_n$, $n \geq 1$, where we assume that $\mathcal{Z}_n$, $n \geq 1$, are finite sets. If

$$C \overset{\text{def}}{=} \limsup_{n\to\infty} \frac{1}{n} \log_2 |\mathcal{Z}_n| < \infty$$

then $\overline{H}(\boldsymbol{Z}) \leq C$.

Proof of Theorem 8: It suffices to prove that $R \geq \overline{H}(\boldsymbol{X})$ for any achievable rate R. If $R = \infty$, $R \geq \overline{H}(\boldsymbol{X})$ is trivial. Hereafter, we assume that $R < \infty$ is achievable and prove that $R \geq \overline{H}(\boldsymbol{X})$.

We first note that Theorem 6 and Fact 2 guarantee

$$\text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{P_{X^n}(X^n)}{P_{W_n}(W_n)} \geq 0$$

for any $\{(f_n, g_n)\}_{n=1}^\infty$ satisfying (90) Then, by using (5) and (7) we can obtain

$$\text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \leq \text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{1}{P_{W_n}(W_n)}$$

which means $\overline{H}(\boldsymbol{X}) \leq \overline{H}(\boldsymbol{W})$ for $\boldsymbol{W} = \{W_n\}_{n=1}^\infty$. We note that R satisfies (89) because R is assumed to be achievable. Thus, in view of Lemma 7, we have $\overline{H}(\boldsymbol{W}) \leq R$. Since $\overline{H}(\boldsymbol{X}) \leq \overline{H}(\boldsymbol{W})$ and R is arbitrary, the claim of the theorem follows. □

VIII. Conclusion

In this paper we have considered the (t, m)-threshold scheme where an n-tuple of secrets generated from a general source is encrypted to m shares. The (t, m)-threshold scheme is required to satisfy two conditions, one is on the decoding error probability and the other is on the security of X^n against arbitrarily collection of less than t shares. We have developed two inequalities including the limit inferior in probability one of which is closely related to the minimum length of the fair random bits needed to a dealer. In addition, we have given a construction of the (t, m)-threshold scheme meeting the two condition under a certain assumption. We can also take the same approach to the problems of Shannon's cipher system with the perfect secrecy and fixed-length source coding with vanishing decoding error probabilities. We have obtained the inequalities in both problems that are valid for stochastic encoders and lead to the converse coding theorems as easy consequences.

Appendix A
Proofs of Facts 1 and 2

Proof of Fact 1: We only prove (8) because (9) is proved similarly. For simplicity, set $\xi = \text{p-}\liminf_{n\to\infty} U_n$.

($\Rightarrow$) Suppose that $\xi \geq A$. Then, from the definition (1), it holds that $\Pr\{U_n \geq \xi - \gamma\} \to 1$ as $n \to \infty$ for any $\gamma > 0$. Since $\xi \geq A$ implies that $\Pr\{U_n \geq \xi - \gamma\} \leq \Pr\{U_n \geq A - \gamma\}$, we have $\Pr\{U_n \geq A - \gamma\} \to 1$ as $n \to \infty$.

($\Leftarrow$) Suppose that $\Pr\{U_n \geq A - \gamma\} \to 1$ as $n \to \infty$ holds for any $\gamma > 0$. Since ξ is defined as the supremum of β satisfying $\Pr\{U_n \geq \beta\} \to 1$ as $n \to \infty$, we have $A - \gamma \leq \xi$. Hence, we obtain $A \leq \xi$ because $\gamma > 0$ is arbitrary. □

Proof of Fact 2: We only prove the first claim because the second claim is proved similarly. In view of Fact 1, $\text{p-}\liminf_{n\to\infty} U_n \geq 0$ is equivalent to $\lim_{n\to\infty} \Pr\{U_n \geq -\gamma\} = 1$ for any $\gamma > 0$. Since $n^\alpha \geq 1$ for any $n \geq 1$ and $\alpha > 0$, it holds that $\Pr\{U_n \geq -\gamma\} \leq \Pr\{U_n \geq -n^\alpha \gamma\} = \Pr\{\frac{1}{n^\alpha} U_n \geq -\gamma\}$. This guarantees that $\text{p-}\liminf_{n\to\infty} \frac{1}{n^\alpha} U_n \geq 0$ due to Fact 1. □

Appendix B
Proof of Lemma 4

Proof: Fix a sequence $\{(f_n, g_n)\}_{n=1}^\infty$ satisfying condition G1) arbitrary. We first prove that

$$1 - \varepsilon_n^{(i_1,\ldots,i_t)} = \sum_{w_n^{(i_1,\ldots,i_t)} \in \mathcal{W}_n^{(i_1,\ldots,i_t)}} P_{X^n W_n^{(i_1,\ldots,i_t)}}(g_n^{(i_1,\ldots,i_t)}(w_n^{(i_1,\ldots,i_t)}), w_n^{(i_1,\ldots,i_t)}) \quad (91)$$

for any $\{i_1, \ldots, i_t\} \subset \mathcal{P}$, where $P_{X^n W_n^{(i_1,\ldots,i_t)}}$ denotes the joint probability of X^n and $W_n^{(i_1,\ldots,i_t)}$. To this end, fix $\{i_1, \ldots, i_t\} \subset \mathcal{P}$ arbitrary and define

$$\chi_0^{(i_1,\ldots,i_t)}(x^n, e_n) = \begin{cases} 1, & \text{if } g_n^{(i_1,\ldots,i_t)}\ (f_n^{(i_1,\ldots,i_t)}(x^n, e_n)) = x^n \\ 0, & \text{otherwise.} \end{cases} \quad (92)$$

Then, by using independence of X^n and E_n, it clearly holds that

$$1 - \varepsilon_n^{(i_1,\ldots,i_t)} = \sum_{x^n \in \mathcal{X}^n} \sum_{e_n \in \mathcal{E}_n} P_{X^n}(x^n) P_{E_n}(e_n) \times \chi_0^{(i_1,\ldots,i_t)}(x^n, e_n). \quad (93)$$

For each $w_n^{(i_1,\ldots,i_t)} \in \mathcal{W}_n^{(i_1,\ldots,i_t)}$ define

$$\mathcal{J}_n(w_n^{(i_1,\ldots,i_t)}) = \{(x^n, e_n) \in \mathcal{X}^n \times \mathcal{E}_n : f_n^{(i_1,\ldots,i_t)}(x^n, e_n) = w_n^{(i_1,\ldots,i_t)}\}.$$

Since $f_n^{(i_1,\ldots,i_t)}$ is deterministic, $\mathcal{J}_n(w_n^{(i_1,\ldots,i_t)})$, $w_n^{(i_1,\ldots,i_t)} \in \mathcal{W}_n^{(i_1,\ldots,i_t)}$, form a partition of $\mathcal{X}^n \times \mathcal{E}_n$. Therefore, (93) can be written as

$$\begin{aligned} 1-\varepsilon_n^{(i_1,\ldots,i_t)} &= \sum_{w_n^{(i_1,\ldots,i_t)} \in \mathcal{W}_n^{(i_1,\ldots,i_t)}} \sum_{(x^n,e_n) \in \mathcal{J}_n(w_n^{(i_1,\ldots,i_t)})} P_{X^n}(x^n) P_{E_n}(e_n) \\ &\quad \times \chi_0^{(i_1,\ldots,i_t)}(x^n, e_n) \\ &= \sum_{w_n^{(i_1,\ldots,i_t)} \in \mathcal{W}_n^{(i_1,\ldots,i_t)}} \sum_{(x^n,e_n) \in \mathcal{J}_n(w_n^{(i_1,\ldots,i_t)})} P_{X^n}(x^n) P_{E_n}(e_n) \\ &\quad \times \chi^{(i_1,\ldots,i_t)}(x^n, e_n, w_n^{(i_1,\ldots,i_t)}) \end{aligned} \quad (94)$$

where

$$\chi^{(i_1,\ldots,i_t)}(x^n, e_n, w_n^{(i_1,\ldots,i_t)}) = \begin{cases} 1, & \text{if } f_n^{(i_1,\ldots,i_t)}(x^n, e_n) = w_n^{(i_1,\ldots,i_t)} \\ & \text{and } g_n(w_n^{(i_1,\ldots,i_t)}) = x^n \\ 0, & \text{otherwise,} \end{cases}$$

and the last equality follows from the fact that $(x^n, e_n) \in \mathcal{J}_n(w_n^{(i_1,\ldots,i_t)})$ implies $f_n(x^n, e_n) = w_n^{(i_1,\ldots,i_t)}$ and therefore we have $\chi_0^{(i_1,\ldots,i_t)}(x^n, e_n) = \chi^{(i_1,\ldots,i_t)}(x^n, e_n, w_n^{(i_1,\ldots,i_t)})$ for all $w_n^{(i_1,\ldots,i_t)} \in \mathcal{W}_n^{(i_1,\ldots,i_t)}$ and $(x^n, e_n) \in \mathcal{J}_n(w_n^{(i_1,\ldots,i_t)})$.

Notice here that, since $f_n^{(i_1,\ldots,i_t)}$ is deterministic, we have

$$P_{W_n^{(i_1,\ldots,i_t)}|X^n E_n}(w_n^{(i_1,\ldots,i_t)}|x^n, e_n) = \begin{cases} 1, & \text{if } (x^n, e_n) \in \mathcal{J}_n(w_n^{(i_1,\ldots,i_t)}) \\ 0, & \text{otherwise.} \end{cases} \quad (95)$$

Hence, (94) and (95) yield

$$\begin{aligned} 1-\varepsilon_n^{(i_1,\ldots,i_t)} &= \sum_{w_n^{(i_1,\ldots,i_t)} \in \mathcal{W}_n^{(i_1,\ldots,i_t)}} \sum_{x^n \in \mathcal{X}^n} \sum_{e_n \in \mathcal{E}_n} \\ &\quad P_{X^n E_n W_n^{(i_1,\ldots,i_t)}}(x^n, e_n, w_n^{(i_1,\ldots,i_t)}) \\ &\quad \times \chi^{(i_1,\ldots,i_t)}(x^n, e_n, w_n^{(i_1,\ldots,i_t)}) \\ &= \sum_{w_n^{(i_1,\ldots,i_t)} \in \mathcal{W}_n^{(i_1,\ldots,i_t)}} P_{W_n^{(i_1,\ldots,i_t)}}(w_n^{(i_1,\ldots,i_t)}) \\ &\quad \times \sum_{x^n \in \mathcal{X}^n} \sum_{e_n \in \mathcal{E}_n} P_{X^n E_n | W_n^{(i_1,\ldots,i_t)}}(x^n, e_n | w_n^{(i_1,\ldots,i_t)}) \\ &\quad \times \chi^{(i_1,\ldots,i_t)}(x^n, e_n, w_n^{(i_1,\ldots,i_t)}). \end{aligned} \quad (96)$$

Note that in (96) we have $\chi^{(i_1,\ldots,i_t)}(g_n(w_n^{(i_1,\ldots,i_t)}), e_n, w_n) = 1$ and $\chi^{(i_1,\cdots,i_t)}(x^n, e_n, w_n) = 0$ for all $x^n \neq g_n(w_n^{(i_1,\ldots,i_t)})$.

Therefore, by taking the sum with respect to x^n in (96) it follows that

$$\begin{aligned}
1-&\varepsilon_n^{(i_1,\ldots,i_t)}\\
&= \sum_{w_n^{(i_1,\ldots,i_t)}\in\mathcal{W}_n^{(i_1,\ldots,i_t)}} P_{W_n^{(i_1,\ldots,i_t)}}(w_n^{(i_1,\ldots,i_t)})\\
&\quad \times \sum_{e_n\in\mathcal{E}_n} P_{X^nE_n|W_n^{(i_1,\ldots,i_t)}}(g_n(w_n^{(i_1,\ldots,i_t)}), e_n|w_n^{(i_1,\ldots,i_t)})\\
&= \sum_{w_n^{(i_1,\ldots,i_t)}\in\mathcal{W}_n^{(i_1,\ldots,i_t)}} P_{W_n^{(i_1,\ldots,i_t)}}(w_n^{(i_1,\ldots,i_t)})\\
&\quad \times P_{X^n|W_n^{(i_1,\ldots,i_t)}}(g_n(w_n^{(i_1,\ldots,i_t)})|w_n^{(i_1,\ldots,i_t)})
\end{aligned} \tag{97}$$

which establishes (91).

Now, letting $\gamma > 0$ be an arbitrary constant, define

$$\begin{aligned}
\mathcal{G}_n =& \Big\{w_n^{(i_1,\ldots,i_t)} \in \mathcal{W}_n^{(i_1,\ldots,i_t)} :\\
&\quad P_{X^n|W_n^{(i_1,\ldots,i_t)}}(g_n(w_n^{(i_1,\ldots,i_t)})|w_n^{(i_1,\ldots,i_t)}) \geq 2^{-\gamma}\Big\},\\
\mathcal{B}_n =& \Big\{w_n^{(i_1,\ldots,i_t)} \in \mathcal{W}_n^{(i_1,\ldots,i_t)} :\\
&\quad P_{X^n|W_n^{(i_1,\ldots,i_t)}}(g_n(w_n^{(i_1,\ldots,i_t)})|w_n^{(i_1,\ldots,i_t)}) < 2^{-\gamma}\Big\}.
\end{aligned}$$

Clearly, $(\mathcal{G}_n, \mathcal{B}_n)$ gives a partition of $\mathcal{W}_n^{(i_1,\ldots,i_t)}$. Thus, in view of (97), we have

$$\begin{aligned}
1-&\varepsilon_n^{(i_1,\ldots,i_t)}\\
&\leq \sum_{w_n^{(i_1,\ldots,i_t)}\in\mathcal{G}_n} P_{W_n^{(i_1,\ldots,i_t)}}(w_n^{(i_1,\ldots,i_t)})\\
&\quad \times P_{X^n|W_n^{(i_1,\ldots,i_t)}}(g_n(w_n^{(i_1,\ldots,i_t)})|w_n^{(i_1,\ldots,i_t)})\\
&\quad + \sum_{w_n^{(i_1,\ldots,i_t)}\in\mathcal{B}_n} P_{W_n^{(i_1,\ldots,i_t)}}(w_n^{(i_1,\ldots,i_t)})\\
&\quad \times P_{X^n|W_n^{(i_1,\ldots,i_t)}}(g_n(w_n^{(i_1,\ldots,i_t)})|w_n^{(i_1,\ldots,i_t)})\\
&< \sum_{w_n^{(i_1,\ldots,i_t)}\in\mathcal{G}_n} P_{W_n^{(i_1,\ldots,i_t)}}(w_n^{(i_1,\ldots,i_t)})\\
&\quad + 2^{-\gamma} \sum_{w_n^{(i_1,\ldots,i_t)}\in\mathcal{B}_n} P_{W_n^{(i_1,\ldots,i_t)}}(w_n^{(i_1,\ldots,i_t)})\\
&= 1-(1-2^{-\gamma})\Pr\{W_n^{(i_1,\ldots,i_t)} \in \mathcal{B}_n\}.
\end{aligned} \tag{98}$$

Thus, the combination of (98) with condition G1) in Definition 2 yields

$$\Pr\{W_n^{(i_1,\ldots,i_t)} \in \mathcal{B}_n\} < \frac{\varepsilon_n^{(i_1,\ldots,i_t)}}{1-2^{-\gamma}} \to 0 \quad \text{as } n \to \infty,$$

which immediately implies

$$\Pr\{W_n^{(i_1,\ldots,i_t)} \in \mathcal{G}_n\} \to 1 \quad \text{as } n \to \infty. \tag{99}$$

In order to complete the proof of this lemma, we define

$$\begin{aligned}
\mathcal{U}_n =& \big\{(x^n, w_n^{(i_1,\ldots,i_t)}) \in \mathcal{X}^n \times \mathcal{W}_n^{(i_1,\ldots,i_t)} :\\
&\quad P_{X^n|W_n^{(i_1,\ldots,i_t)}}(x^n|w_n^{(i_1,\ldots,i_t)}) \geq 2^{-\gamma}\big\}\\
\tilde{\mathcal{G}}_n =& \big\{(g_n(w_n^{(i_1,\ldots,i_t)}), w_n^{(i_1,\ldots,i_t)}) \in \mathcal{X}^n \times \mathcal{W}_n^{(i_1,\ldots,i_t)} :\\
&\quad w_n^{(i_1,\ldots,i_t)} \in \mathcal{G}_n\big\}.
\end{aligned}$$

Since $\tilde{\mathcal{G}}_n \subset \mathcal{U}_n$, it follows from (99) that

$$\begin{aligned}
\Pr\{(X^n, W_n^{(i_1,\ldots,i_t)}) \in \mathcal{U}_n\} &\geq \Pr\{(X^n, W_n^{(i_1,\ldots,i_t)}) \in \tilde{\mathcal{G}}_n\}\\
&= \Pr\{W_n^{(i_1,\ldots,i_t)} \in \mathcal{G}_n\}\\
&\to 1 \quad \text{as } n \to \infty
\end{aligned}$$

where the equality follows from the definition of $\tilde{\mathcal{G}}_n$. Since $\gamma > 0$ is arbitrary, we obtain the claim of this lemma. □

Appendix C
Proof of Lemma 5

Proof: Letting $\gamma > 0$ be an arbitrary constant, it suffices to prove that

$$\begin{aligned}
\Pr\left\{\frac{1}{n^\alpha}\log\frac{P_{W_n^{(i_j)}|W_n^{(I_j)}}(W_n^{(i_j)}|W_n^{(I_j)})}{P_{W_n^{(i_j)}|W_n^{(i_1,\ldots,i_{j-1})}}(W_n^{(i_j)}|W_n^{(i_1,\ldots,i_{j-1})})} \leq -\gamma\right\}&\\
\leq 2^{-n^\alpha\gamma} \to 0 \quad \text{as } n \to \infty.&
\end{aligned} \tag{100}$$

To this end, define

$$\mathcal{B}_n^{(i_j)} = \left\{w^{(i_1,\ldots,i_t)} \in \mathcal{W}_n^{(i_1,\ldots,i_t)} : \frac{1}{n^\alpha}\log\frac{P_{W_n^{(i_j)}|W_n^{(I_j)}}(w_n^{(i_j)}|w_n^{(I_j)})}{P_{W_n^{(i_j)}|W_n^{(i_1,\ldots,i_{j-1})}}(w_n^{(i_j)}|w_n^{(i_1,\ldots,i_{j-1})})} \leq -\gamma\right\}.$$

Note that for all $w^{(i_1,\ldots,i_t)} \in \mathcal{B}_n^{(i_j)}$ it holds that

$$\begin{aligned}
P&_{W_n^{(i_j)}|W_n^{(I_j)}}(w_n^{(i_j)}|w_n^{(I_j)})\\
&\leq 2^{-n^\alpha\gamma} P_{W_n^{(i_j)}|W_n^{(i_1,\ldots,i_{j-1})}}(w_n^{(i_j)}|w_n^{(i_1,\ldots,i_{j-1})})
\end{aligned} \tag{101}$$

for $j = 1, 2, \ldots, t$. By multiplying

$$\begin{aligned}
P&_{W_n^{(I_j)}}(w_n^{(I_j)})\\
&= P_{W_n^{(i_1,\ldots,i_{j-1},i_{j+1},\ldots,i_t)}}(w_n^{(i_1,\ldots,i_{j-1},i_{j+1},\ldots,i_t)})\\
&= P_{W_n^{(i_1,\ldots,i_{j-1})}}(w_n^{(i_1,\ldots,i_{j-1})})\\
&\quad \times P_{W_n^{(i_{j+1},\ldots,i_t)}|W_n^{(i_1,\ldots,i_{j-1})}}(w_n^{(i_{j+1},\ldots,i_t)}|w_n^{(i_1,\ldots,i_{j-1})})
\end{aligned}$$

to both sides of (101), we obtain

$$\begin{aligned}
&P_{W_n^{(i_1,\ldots,i_t)}}(w_n^{(i_1,\ldots,i_t)})\\
&\leq 2^{-n^\alpha\gamma} P_{W_n^{(i_1,\ldots,i_j)}}(w_n^{(i_1,\ldots,i_j)})\\
&\quad \times P_{W_n^{(i_{j+1},\ldots,i_t)}|W_n^{(i_1,\ldots,i_{j-1})}}(w_n^{(i_{j+1},\ldots,i_t)}|w_n^{(i_1,\ldots,i_{j-1})}).
\end{aligned} \tag{102}$$

Then, it follows that

$$\begin{aligned}
&\Pr\{W_n^{(i_1,\ldots,i_t)} \in \mathcal{B}_n^{(i_j)}\}\\
&= \sum_{w_n^{(i_1,\ldots,i_t)} \in \mathcal{B}_n^{(i_j)}} P_{W_n^{(i_1,\ldots,i_t)}}(w_n^{(i_1,\ldots,i_t)})\\
&\le 2^{-n^\alpha \gamma} \sum_{w_n^{(i_1,\ldots,i_t)} \in \mathcal{B}_n^{(i_j)}} P_{W_n^{(i_1,\ldots,i_j)}}(w_n^{(i_1,\ldots,i_j)})\\
&\quad \times P_{W_n^{(i_{j+1},\ldots,i_t)}|W_n^{(i_1,\ldots,i_{j-1})}}(w_n^{(i_{j+1},\ldots,i_t)}|w_n^{(i_1,\ldots,i_{j-1})})\\
&\le 2^{-n^\alpha \gamma} \sum_{w_n^{(i_1,\ldots,i_j)} \in \mathcal{W}_n^{(i_1,\ldots,i_j)}} P_{W_n^{(i_1,\ldots,i_j)}}(w_n^{(i_1,\ldots,i_j)})\\
&\quad \times \sum_{w_n^{(i_{j+1},\ldots,i_t)} \in \mathcal{W}_n^{(i_{j+1},\ldots,i_t)}}\\
&\quad P_{W_n^{(i_{j+1},\ldots,i_t)}|W_n^{(i_1,\ldots,i_{j-1})}}(w_n^{(i_{j+1},\ldots,i_t)}|w_n^{(i_1,\ldots,i_{j-1})})\\
&= 2^{-n^\alpha \gamma} \to 0 \quad \text{as } n \to \infty
\end{aligned} \tag{103}$$

where the first inequality in (103) follows from (102). This establishes (100). □

Appendix D
Proof of (71)

Proof: Fix a sequence $\{(f_n, g_n)\}_{n=1}^\infty$ of encoders and decoders. From the definition of the decoding error probability, it holds that

$$\begin{aligned}
1 - \varepsilon_n = &\sum_{x^n \in \mathcal{X}^n} \sum_{e_n \in \mathcal{E}_n} \sum_{w_n \in \mathcal{W}_n} P_{X^n}(x^n) P_{E_n}(e_n)\\
&\times P_{W_n|X^nE_n}(w_n|x^n, e_n)\chi(x^n, e_n, w_n)
\end{aligned} \tag{104}$$

where

$$\chi(x^n, e_n, w_n) = \begin{cases} 1, & x^n = g_n(w_n, e_n)\\ 0, & \text{otherwise}\end{cases}$$

and $P_{W_n|X^nE_n}$ in (104) is the conditional probability distribution corresponding to the stochastic encoder f_n. Notice in (104) that we can take the sum with respect to $x^n \in \mathcal{X}^n$. That is, in view of the definition of $\chi(x^n, e_n, w_n)$, it holds that

$$\begin{aligned}
1 - \varepsilon_n &= \sum_{w_n \in \mathcal{W}_n} \sum_{e_n \in \mathcal{E}_n} P_{X^n}(g_n(w_n, e_n)) P_{E_n}(e_n)\\
&\quad \times P_{W_n|X^nE_n}(g_n(w_n, e_n)|x^n, e_n)\\
&= \sum_{w_n \in \mathcal{W}_n} \sum_{e_n \in \mathcal{E}_n} P_{W_nE_n}(w_n, e_n)\\
&\quad \times P_{X^n|W_nE_n}(g_n(w_n, e_n)|w_n, e_n)
\end{aligned} \tag{105}$$

where the second equality follows from Bayes' formula.

Hereafter, we repeat the argument given in the proof of Lemma 4 in Appendix B. Define

$$\begin{aligned}
\mathcal{G}_n &= \big\{(w_n, e_n) \in \mathcal{W}_n \times \mathcal{E}_n :\\
&\qquad P_{X^n|W_nE_n}(g_n(w_n, e_n)|w_n, e_n) \ge 2^{-\gamma}\big\}\\
\mathcal{B}_n &= \big\{(w_n, e_n) \in \mathcal{W}_n \times \mathcal{E}_n :\\
&\qquad P_{X^n|W_nE_n}(g_n(w_n, e_n)|w_n, e_n) < 2^{-\gamma}\big\}.
\end{aligned}$$

Then, in view of (105), we have

$$\begin{aligned}
1 - \varepsilon_n = &\sum_{(w_n, e_n) \in \mathcal{G}_n} P_{W_nE_n}(w_n, e_n)\\
&\times P_{X^n|W_nE_n}(g_n(w_n, e_n)|w_n, e_n)\\
&+ \sum_{(w_n, e_n) \in \mathcal{B}_n} P_{W_nE_n}(w_n, e_n)\\
&\times P_{X^n|W_nE_n}(g_n(w_n, e_n)|w_n, e_n)\\
< &\sum_{(w_n, e_n) \in \mathcal{G}_n} P_{W_nE_n}(w_n, e_n)\\
&+ 2^{-\gamma} \sum_{(w_n, e_n) \in \mathcal{B}_n} P_{W_nE_n}(w_n, e_n)\\
= &1 - (1 - 2^{-\gamma}) \Pr\{(W_n, E_n) \in \mathcal{B}_n\}
\end{aligned}$$

which, together with S1) in Definition 3, yields

$$\Pr\{(W_n, E_n) \in \mathcal{B}_n\} < \frac{\varepsilon_n}{1 - 2^{-\gamma}} \to 0 \quad \text{as } n \to \infty.$$

Hence, we obtain

$$\Pr\{(W_n, E_n) \in \mathcal{G}_n\} \to 1 \quad \text{as } n \to \infty. \tag{106}$$

In view of (106), we can establish (71) from the following argument:

$$\begin{aligned}
\Pr\{(X^n, E_n, W_n) \in \mathcal{U}_n\} &\ge \Pr\{(X^n, E_n, W_n) \in \tilde{\mathcal{G}}_n\}\\
&= \Pr\{(W_n, E_n) \in \mathcal{G}_n\}\\
&\to 1 \quad \text{as } n \to \infty,
\end{aligned} \tag{107}$$

where $\mathcal{U}_n$ and $\tilde{\mathcal{G}}_n$ are defined by

$$\begin{aligned}
\mathcal{U}_n &= \{(x^n, e_n, w_n) \in \mathcal{X}^n \times \mathcal{E}_n \times \mathcal{W}_n :\\
&\qquad P_{X^n|W_nE_n}(x^n|w_n, e_n) \ge 2^{-\gamma}\}\\
\tilde{\mathcal{G}}_n &= \{(g_n(w_n, e_n), e_n, w_n) \in \mathcal{X}^n \times \mathcal{E}_n \times \mathcal{W}_n :\\
&\qquad (w_n, e_n) \in \mathcal{G}_n\}
\end{aligned}$$

and the inequality in (107) follows from $\tilde{\mathcal{G}}_n \subset \mathcal{U}_n$. □

Appendix E
Proof of Lemma 7

Proof: We prove Lemma 7 by a contradiction argument. Assume that $\overline{H}(\boldsymbol{Z}) > C$. Then, there exists a constant $\varepsilon_0 > 0$ satisfying $\overline{H}(\boldsymbol{Z}) - \varepsilon_0 \ge C + 2\varepsilon_0$. Define $\mathcal{Z}_n'$ by

$$\mathcal{Z}_n' = \left\{ z_n \in \mathcal{Z}_n : \frac{1}{n} \log_2 \frac{1}{P_{Z_n}(z_n)} \ge C + 2\varepsilon_0 \right\}.$$

Then, due to the definition of ε_0, it follows that

$$\begin{aligned}
&\Pr\left\{ \frac{1}{n} \log_2 \frac{1}{P_{Z_n}(Z_n)} \ge \overline{H}(\boldsymbol{Z}) - \varepsilon_0 \right\}\\
&\le \Pr\left\{ \frac{1}{n} \log_2 \frac{1}{P_{Z_n}(Z_n)} \ge C + 2\varepsilon_0 \right\}\\
&= \Pr\{Z_n \in \mathcal{Z}_n'\}.
\end{aligned} \tag{108}$$

Note that, owing to the definition of $\overline{H}(\boldsymbol{Z})$, the left side of (108) is positive for infinitely many n. That is, there exists a constant $\gamma_0 > 0$ satisfying

$$\Pr\{Z_n \in \mathcal{Z}_n'\} \geq \gamma_0 \quad \text{infinitely often.} \tag{109}$$

On the other hand, since $P_{Z_n}(z_n) \leq 2^{-n(C+2\varepsilon_0)}$ for all $n \geq 1$ and $z_n \in \mathcal{Z}_n'$, it follows that

$$\begin{aligned}\Pr\{Z_n \in \mathcal{Z}_n'\} &= \sum_{z_n \in \mathcal{Z}_n'} P_{Z_n}(z_n)\\ &\leq \sum_{z_n \in \mathcal{Z}_n'} 2^{-n(C+2\varepsilon_0)}\\ &= |\mathcal{Z}_n'|\, 2^{-n(C+2\varepsilon_0)} \quad \text{for all } n \geq 1.\end{aligned} \tag{110}$$

Therefore, the combination of (109) and (110) yields

$$\gamma_0 \leq |\mathcal{Z}_n'|\, 2^{-n(C+2\varepsilon_0)} \quad \text{infinitely often}$$

that is,

$$\frac{1}{n}\log_2 |\mathcal{Z}_n'| \geq C + 2\varepsilon_0 + \frac{1}{n}\log_2 \gamma_0 \quad \text{infinitely often.} \tag{111}$$

Notice here that, since $\varepsilon_0 + \frac{1}{n}\log_2 \gamma_0 \geq 0$ for all sufficiently large n, (111) guarantees the existence of a subsequence $\{n_j\}_{j=1}^{\infty}$ satisfying

$$\frac{1}{n_j}\log_2 |\mathcal{Z}_{n_j}'| \geq C + \varepsilon_0 \quad \text{for all } j \geq 1. \tag{112}$$

Hence, in view of (112) we obtain

$$\begin{aligned}C + \varepsilon_0 &\leq \limsup_{j\to\infty} \frac{1}{n_j}\log_2 |\mathcal{Z}_{n_j}'|\\ &\leq \limsup_{n\to\infty} \frac{1}{n}\log_2 |\mathcal{Z}_n'|\\ &\leq \limsup_{n\to\infty} \frac{1}{n}\log_2 |\mathcal{Z}_n| = C\end{aligned} \tag{113}$$

where the second inequality follows from a property of the limit superior and the last inequality follows from $\mathcal{Z}_n' \subset \mathcal{Z}_n$. Equation (113) and the assumption of $C < \infty$ imply that $\varepsilon_0 \leq 0$. However, this contradicts the assumption that ε_0 is a positive constant. This completes the proof of $\overline{H}(\boldsymbol{Z}) \leq C$. □

Acknowledgment

A part of the results in this paper was obtained when the author visited ETH, Zurich. The author wishes to thank Professor U. Maurer and the members of his research group for fruitful discussions. The author also wishes to thank Japan Society of Promotion of Science and the Swiss National Foundation of Science for supporting his visit to ETH. Thanks also go to Professor Y. Desmedt of University College London, U.K., for pointing out a relationship between [4] and this study.

References

[1] G. R. Blakley, “Safeguarding cryptographic keys,” in *Proc. AFIPS 1979 National Comp. Conf.*, 1979, vol. 48, pp. 313–317.

[2] C. Blundo, A. De Santis, and U. Vaccaro, “Randomness in distribution protocols,” *Inf. Computat.*, vol. 131, pp. 111–139, 1996.

[3] R. M. Capocelli, A. De Santis, L. Gargano, and U. Vaccaro, “On the size of shares for secret sharing schemes,” in *Proc. Crypto’91*, 1991, vol. 573, LNCS, pp. 101–113.

[4] B. Chor and E. Kushilevitz, “Secret sharing over infinite domain,” in *Proc. Crypto’ 89*, 1990, vol. 435, LNCS, pp. 299–306.

[5] T. M. Cover and J. A. Thomas*, Elements of Information Theory*. New York: Wiley, 1991.

[6] T. S. Han, *Information-Spectrum Methods in Information Theory*. New York: Springer-Verlag, 2003.

[7] T. S. Han and S. Verdú, “Approximation theory of output statistics,” *IEEE Trans. Inf. Theory*, vol. IT-39, pp. 752–772, 1993.

[8] E. D. Karnin, J. W. Greene, and M. E. Hellman, “On secret sharing systems,” *IEEE Trans. Inf. Theory*, vol. IT29, pp. 35–41, 1983.

[9] H. Koga, “A coding theorem on the Shannon cipher system with a general source using fixed-length homophonic coding,” in *Proc. 2003 IEEE ISIT*, Yokohama, Japan, 2003, p. 225.

[10] H. Koga and T. Ooishi, “Coding theorems on secret sharing schemes for a general source with the strong converse property,” in *Proc. ISITA 2004*, Parma, Italy, 2004, pp. 1104–1109.

[11] A. Shamir, “How to share a secret,” *Commun. ACM*, vol. 22, pp. 612–613, 1979.

[12] C. E. Shannon, “Communication theory of secret system,” *Bell Syst. Tech. J.*, vol. 27, pp. 656–715, 1949.

[13] S. Vembu and S. Verdú, “General random bits from an arbitrary source: Fundamental limits,” *IEEE Trans. Inf. Theory*, vol. IT-41, pp. 1322–1332, 1995.

[14] S. Verdú and T. S. Han, “A general formula for channel capacity,” *IEEE Trans. Inf. Theory*, vol. IT40, pp. 1147–1157, 1994.

[15] H. Yamamoto, “On secret sharing communication systems with two or three channels,” *IEEE Trans. Inf. Theory*, vol. IT32, pp. 387–393, 1986.

[16] H. Yamamoto, “Information theory in Cryptology,” *IEICE Trans. Fundamentals*, vol. E74, no. 9, pp. 2456–2464, 1991.

[17] H. Yamamoto, “Coding theorems for Shannon’s cipher system with correlated source outputs, and common information,” *IEEE Trans. Inf. Theory*, vol. 40, pp. 85–95, 1994.