A Characterization of Optimal FF Coding Rate Using a New Optimistically Optimal Code

Mitsuharu ARIMURA[†a]**, Hiroki KOGA**[††b]**, *Senior Members, and* Ken-ichi IWATA**[†††c]**, *Member***

SUMMARY   In this letter, we first introduce a stronger notion of the optimistic achievable coding rate and discuss a coding theorem. Next, we give a necessary and sufficient condition under which the coding rates of all the optimal FF codes asymptotically converge to a constant.
key words: information spectrum, general source, FF code, optimistic achievable coding rate

1.   Introduction

In information spectrum methods [4], the performance of a fixed-to-fixed length (FF) source code is measured by the limit superior of the coding rate and the error probability of the code. Since there is a trade-off between the coding rate and the error probability, the optimal coding rate is defined as the infimum of the limit superior of the coding rate subject to the condition that the error probability converges to zero. It is proved that this optimal coding rate coincides with the spectral sup-entropy rate $\overline{H}(X)$ [4]. Under this criterion, however, the coding rate does not converge to a constant in general.

On the other hand, let us consider the situation where we construct an FF code with vanishing error probability for a stationary and memoryless source. Since almost all the typical sequences should be correctly decoded in such a code [3, Theorem 3.3.1], its coding rate is asymptotically lower-bounded by $H(X) - \gamma$, where $H(X)$ is the entropy of the source and $\gamma > 0$ is an arbitrarily small constant. Hence, if we require that the FF code is optimal, i.e., its coding rate is upper-bounded by $H(X) + \gamma$ for sufficiently large block-length, the coding rate is almost equal to the entropy. That is, the coding rate of *any* optimal code is arbitrarily close to the entropy.

The objective of this paper is clarifying the condition under which the coding rate of any optimal code asymptotically attaining $\overline{H}(X)$ converge to a constant that is independent of the optimal code. To this end, we define a

new optimality of the source codes which is closely related to the optimistic source coding discussed in [2], [6], [8], [9] and obtain a general formula of the infimum attainable coding rate. Then, by using this notion of the optimality, we obtain a necessary and sufficient condition under which the coding rate converges to $\overline{H}(X)$.

This letter is organized as follows. Section 2 is devoted to definitions of a general source and an FF source code. In Sect. 3 we define the strongly optimistic achievable coding rate, and show that its infimum coincides with the quantity $\overline{H}^*(X)$ introduced in [2], [7]. The necessary and sufficient condition is established in Sect. 4.

2.   FF Coding of General Sources

In this section, we define FF coding of general sources.

Let $\mathbb{N}$ be the set of all the positive integers. For each $n \in \mathbb{N}$ let $X^n = X_1 X_2 \cdots X_n$ be a random variable representing n outputs from a source, where each X_i takes values in a finite or countable set $\mathcal{X}$. The probability distribution of X^n is denoted by P_{X^n}. The probability that $X^n = x^n$ is expressed as $P_{X^n}(x^n)$. We call $X = \{X^n\}_{n \in \mathbb{N}}$ a general source [5]. We do not impose the consistency condition on $P_{X^n}, n \in \mathbb{N}$.

For a general source, we define two limits concerning the probability distribution of the self information.

Definition 2.1 (Han-Verdú [5]):

$$\overline{H}(X) = \inf\left\{\alpha : \liminf_{n\to\infty} \Pr\left\{\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} \le \alpha\right\} = 1\right\}. \quad (1)$$

Definition 2.2 (Chen-Alajaji [2]):

$$\overline{H}^*(X) = \inf\left\{\alpha : \limsup_{n\to\infty} \Pr\left\{\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} \le \alpha\right\} = 1\right\}. \quad (2)$$

Throughout this letter, the bases of logarithmic and exponential functions are assumed to be 2. It is known that $\overline{H}^*(X) \le \overline{H}(X)$ for any general source X [2], [7].

Next, we define an FF code for a general source. A source output x^n of length n is encoded to a codeword $\varphi_n(x^n)$ by an encoder

$$\varphi_n : \mathcal{X}^n \to \mathcal{M}_n \overset{\text{def}}{=} \{1, 2, \ldots, M_n\}$$

and the codeword $\varphi_n(x^n)$ is decoded to $\psi_n(\varphi_n(x^n))$ by a decoder

$$\psi_n : \mathcal{M}_n \to \mathcal{X}^n.$$

We call $C = \{(\varphi_n, \psi_n)\}_{n \in \mathbb{N}}$ a code. The set of all x^n which can be encoded and decoded correctly is denoted by

$$\mathcal{D}_n = \{x^n \in \mathcal{X}^n : x^n = \psi_n(\varphi_n(x^n))\}.$$

Then the probability of decoding error (error probability) is written as

$$\varepsilon_n = \Pr\{X^n \notin \mathcal{D}_n\}.$$

Han and Verdú formulated the infimum achievable coding rate $R(X)$ and obtained a general formula for $R(X)$.

Definition 2.3 (Han-Verdú [5]): A rate R is called an achievable coding rate if there exists a code C satisfying

$$\limsup_{n \to \infty} \frac{1}{n} \log M_n \le R \quad \text{and} \quad \lim_{n \to \infty} \varepsilon_n = 0. \quad (3)$$

The infimum of the achievable coding rate R is called the infimum achievable coding rate and is denoted by $R(X)$.

Theorem 2.1 (Han-Verdú [5]): $R(X) = \overline{H}(X)$.

Based on this theorem, we introduce the optimality of a code C as follows.

Definition 2.4 (Arimura-Koga-Iwata [1]): A code C is said to be $\overline{H}$-optimal for X if C satisfies (3) with $R = \overline{H}(X)$.

It is shown in [1] that there exists a $\overline{H}$-optimal code for any general source X. We discuss a basic property on the $\overline{H}$-optimal codes in Sect. 4.

3. Strongly Optimistic Achievable Coding Rate

In this section we introduce a new notion of optimality on the FF coding. First we review the following results on the optimistic FF coding.

Definition 3.1 (Optimistic Achievable Coding Rate [9]):
A rate R is called an *optimistic achievable coding rate* for X if for any $\delta > 0$ and $\tau > 0$ there exists a code C satisfying both

$$\frac{1}{n} \log M_n \le R + \delta \quad \text{and} \quad \varepsilon_n \le \tau \quad (4)$$

for infinitely many n. The infimum of the optimistic achievable coding rate for X is called the *optimistic infimum achievable coding rate* and is denoted by $\underline{T}(X)$.

Theorem 3.1 (Chen-Alajaji [2]): $\underline{T}(X) = \overline{H}^*(X)$.

Definition 3.1 means that, if R is optimistic achievable, then for arbitrarily given $\delta > 0$ and $\tau > 0$ a code C satisfies (4) with some subsequence $\{n_i\}_{i \in \mathbb{N}}$. It is important to note that this subsequence $\{n_i\}_{i \in \mathbb{N}}$ depends on δ and τ. The following definition actually requires that C has a subsequence $\{n_i\}_{i \in \mathbb{N}}$ that is independent on δ and τ.

Definition 3.2 (Strongly Optimistic Achievable Coding Rate):

A rate R is called a *strongly optimistic achievable coding rate* for X if there exists a code $C = \{(\varphi_n, \psi_n)\}_{n \in \mathbb{N}}$ satisfying

$$\limsup_{i \to \infty} \frac{1}{n_i} \log M_{n_i} \le R, \quad (5)$$

$$\lim_{i \to \infty} \varepsilon_{n_i} = 0 \quad (6)$$

for some subsequence $\{n_i\}_{i \in \mathbb{N}}$. The infimum of the strongly optimistic achievable coding rate for X is called the *strongly optimistic infimum achievable coding rate* and is denoted by $\underline{T}^*(X)$.

We have the following theorem for the strongly optimistic infimum achievable coding rate.

Theorem 3.2: $\underline{T}^*(X) = \overline{H}^*(X)$.

Proof of the direct part of Theorem 3.2 is similar to [1, Theorem 4.4]. The scheme used here can be regarded as another version of the diagonal line argument [4]. Before giving the proof, we give lemmas used in the proof.

From the definition of $\overline{H}^*(X)$, for any constant $\gamma > 0$ and $\tau \in (0, 1)$ it holds that

$$\Pr\left\{\frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \ge \overline{H}^*(X) + \gamma\right\} \le \tau \quad \text{infinitely often.} \quad (7)$$

We arbitrarily fix a constant $\gamma > 0$ and a sequence $\{\tau_i\}_{i \in \mathbb{N}}$ satisfying

$$1 > \tau_1, \ \tau_i > 0, \tau_i > \tau_{i+1}, \forall i \in \mathbb{N}, \ \text{and} \ \lim_{i \to \infty} \tau_i = 0. \quad (8)$$

Define

$$\mathcal{N}_i = \left\{n \in \mathbb{N} : \Pr\left\{\frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \ge \overline{H}^*(X) + \gamma\right\} \le \tau_i\right\}.$$

Then, it follows from (7) that the cardinality of $\mathcal{N}_i$ is countably infinite for each i.

Lemma 3.1: $\mathcal{N}_{i+1} \subseteq \mathcal{N}_i$ for any $i \in \mathbb{N}$.

Proof: This claim immediately follows from the definitions of $\mathcal{N}_i$ and $\{\tau_i\}_{i \in \mathbb{N}}$. Q.E.D.

Lemma 3.2: There exists a strictly monotone increasing sequence $\{n_i\}_{i \in \mathbb{N}}$ of positive integers satisfying $\{n_j : j \ge i\} \subseteq \mathcal{N}_i$ for all $i \in \mathbb{N}$.

Proof: Define $n_0 = 0$ and

$$n_i = \min\{n \in \mathbb{N} : n \in \mathcal{N}_i \text{ and } n > n_{i-1}\} \text{ for } i \ge 1. \quad (9)$$

Note that, since $\mathcal{N}_i$ is a countably infinite set, for each $i \ge 1$ n_i is well-defined. That is, if there is no $n \in \mathcal{N}_i$ satisfying $n > n_{i-1}$, $\mathcal{N}_i$ turns out to be a finite set, which is a contradiction. Therefore, the claim of this lemma follows because $n_i \in \mathcal{N}_i$ and $\mathcal{N}_{i+1} \subseteq \mathcal{N}_i$ for all $i \in \mathbb{N}$. Q.E.D.

Lemma 3.3: The sequence $\{n_i\}_{i \in \mathbb{N}}$ defined by (9) satisfies

$$\lim_{i \to \infty} \Pr\left\{\frac{1}{n_i} \log \frac{1}{P_{X^{n_i}}(X^{n_i})} \geq \overline{H}^*(X) + \gamma\right\} = 0. \qquad (10)$$

Proof: Letting $\tau \in (0, 1)$ be an arbitrary constant, we prove

$$\Pr\left\{\frac{1}{n_i} \log \frac{1}{P_{X^{n_i}}(X^{n_i})} \geq \overline{H}^*(X) + \gamma\right\} \leq \tau$$

for all sufficiently large i.

In view of the definition of $\{\tau_i\}_{i \in \mathbb{N}}$, for any $\tau \in (0, 1)$ we can define i_0 as the minimum integer $i \geq 1$ satisfying $\tau_i \leq \tau$. Then, it obviously holds that

$$\Pr\left\{\frac{1}{n_i} \log \frac{1}{P_{X^{n_i}}(X^{n_i})} \geq \overline{H}^*(X) + \gamma\right\} \leq \tau_i \leq \tau_{i_0} \leq \tau$$

for all $i \geq i_0$. Q.E.D.

We also use the following lemma for proving the converse part of Theorem 3.2.

Lemma 3.4: For any code C satisfying

$$\liminf_{n \to \infty} \varepsilon_n = 0, \qquad (11)$$

it holds that

$$\liminf_{n \to \infty} \frac{1}{n} \log M_n \geq \overline{H}^*(X). \qquad (12)$$

Proof: Assume that (12) does not hold. Then, there exists a real number $\gamma_0 > 0$ satisfying

$$\frac{1}{n} \log M_n \leq \overline{H}^*(X) - 2\gamma_0 \quad \text{for infinitely many } n. \quad (13)$$

Recall here that any code C satisfies

$$\varepsilon_n \geq \Pr\left\{\frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \geq \frac{1}{n} \log M_n + \gamma_0\right\} - \exp(-n\gamma_0) \qquad (14)$$

for all $n \geq 1$ [4, Lemma 1.3.2]. Thus, the combination of (13) and (14) yields

$$\varepsilon_n \geq \Pr\left\{\frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \geq \overline{H}^*(X) - \gamma_0\right\} - \exp(-n\gamma_0)$$

for infinitely many n. (15)

It is important to notice that

$$\liminf_{n \to \infty} \Pr\left\{\frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \geq \overline{H}^*(X) - \gamma_0\right\} > 0 \qquad (16)$$

because the definition of $\overline{H}^*(X)$ implies that

$$\limsup_{n \to \infty} \Pr\left\{\frac{1}{n} \log \frac{1}{P_{X^n}(X^n)} \leq \overline{H}^*(X) - \gamma_0\right\} < 1.$$

Therefore, (16) tells us that the right hand side of (15) is

positive for all sufficiently large n. This contradicts (11). Q.E.D.

Now, we are ready to prove Theorem 3.2.
Proof of Theorem 3.2: Let $\gamma > 0$ be an arbitrary constant. First we prove $\underline{T}^*(X) \leq \overline{H}^*(X)$ by showing that $R = \overline{H}^*(X) + \gamma$ is a strongly optimistic achievable coding rate. To this end, fix $\{\tau_i\}_{i \in \mathbb{N}}$ satisfying (8) arbitrarily. Then, Lemma 3.2 guarantees the existence of a subsequence $\{n_i\}_{i \in \mathbb{N}}$ satisfying (10). Define $\mathcal{N} = \{n_i : i \geq 1\}$.

In order to prove the direct part, it suffices to construct a code only for $n_i \in \mathcal{N}$. Define M_{n_i} as $M_{n_i} = \exp(\lceil n_i(\overline{H}^*(X) + \gamma) \rceil)$. Then, it clearly holds that

$$\limsup_{i \to \infty} \frac{1}{n_i} \log M_{n_i} \leq \overline{H}^*(X) + \gamma. \qquad (17)$$

In addition, it follows from [4, Lemma 1.3.1] and Lemma 3.3 that

$$\varepsilon_{n_i} \leq \Pr\left\{\frac{1}{n_i} \log \frac{1}{P_{X^{n_i}}(X^{n_i})} \geq \frac{1}{n_i} \log M_{n_i}\right\}$$

$$\leq \Pr\left\{\frac{1}{n_i} \log \frac{1}{P_{X^{n_i}}(X^{n_i})} \geq \overline{H}^*(X) + \gamma\right\} \to 0 \text{ as } i \to \infty,$$

where the second inequality follows from the definition of M_{n_i}. This argument establishes that $R = \overline{H}^*(X) + \gamma$ is a strongly optimistic achievable coding rate.

Next, we prove $\underline{T}^*(X) \geq \overline{H}^*(X)$. Let C be an arbitrary code satisfying

$$\lim_{i \to \infty} \varepsilon_{n_i} = 0$$

for some subsequence $\{n_i\}_{i \in \mathbb{N}}$. Since it holds that

$$\liminf_{n \to \infty} \varepsilon_n \leq \liminf_{i \to \infty} \varepsilon_{n_i} = \lim_{i \to \infty} \varepsilon_{n_i} = 0,$$

Lemma 3.4 guarantees that

$$\liminf_{i \to \infty} \frac{1}{n_i} \log M_{n_i} \geq \liminf_{n \to \infty} \frac{1}{n} \log M_n \geq \overline{H}^*(X).$$

This establishes $\underline{T}^*(X) \geq \overline{H}^*(X)$. Q.E.D.

Note that $\underline{T}^*(X) \geq \overline{H}^*(X)$ can be proved by $\underline{T}^*(X) \geq \underline{T}(X)$ and Theorem 3.1. But we use Lemma 3.4 because this lemma is used also in the following section.

4. A New Characterization of Optimal FF Coding Rate

Using the result of the preceding section, we can characterize the asymptotic behavior of the coding rate of $\overline{H}$-optimal codes. The following theorem shows that $\overline{H}(X) = \overline{H}^*(X)$ is a necessary and sufficient condition such that coding rates of all the $\overline{H}$-optimal FF codes converge to $\overline{H}(X)$.

Theorem 4.1: All the $\overline{H}$-optimal FF codes for X satisfy

$$\lim_{n \to \infty} \frac{1}{n} \log M_n = \overline{H}(X) \qquad (18)$$

if and only if $\overline{H}(X) = \overline{H}^*(X)$.

Proof: First we show that if a source X satisfies $\overline{H}(X) = \overline{H}^*(X)$, then any $\overline{H}$-optimal code for X satisfies (18). From the definition of the $\overline{H}$-optimality, any $\overline{H}$-optimal code for X satisfies

$$\limsup_{n\to\infty} \frac{1}{n} \log M_n \le \overline{H}(X) \tag{19}$$

and

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

Then, Lemma 3.4 guarantees

$$\liminf_{n\to\infty} \frac{1}{n} \log M_n \ge \overline{H}^*(X). \tag{20}$$

Since $\overline{H}(X) = \overline{H}^*(X)$ by assumption, (19) and (20) yield

$$\lim_{n\to\infty} \frac{1}{n} \log M_n = \overline{H}(X) = \overline{H}^*(X).$$

Next, we prove that, if (18) holds for any $\overline{H}$-optimal code for X, then X satisfies $\overline{H}(X) = \overline{H}^*(X)$. We prove this by a contradiction argument. Suppose that $\overline{H}^*(X) < \overline{H}(X)$. Then, there exists a real number $\gamma_0 > 0$ satisfying $\overline{H}^*(X) + 2\gamma_0 \le \overline{H}(X)$. Hereinafter, we construct an $\overline{H}$-optimal code $\widetilde{C}$ by combining an arbitrary $\overline{H}$-optimal code C with a code satisfying (5) and (6) with $R = \overline{H}^*(X) + \gamma_0$.

To be more precise, let $C = \{(\varphi_n, \psi_n)\}_{n\in\mathbb{N}}$ be an arbitrary $\overline{H}$-optimal FF code. From the definition of $\overline{H}$-optimality, C satisfies

$$\limsup_{n\to\infty} \frac{1}{n} \log M_n = \overline{H}(X) \tag{21}$$

and

$$\lim_{n\to\infty} \varepsilon_n = 0. \tag{22}$$

On the other hand, Theorem 3.2 tells us that there exists a code $C' = \{(\varphi_n', \psi_n')\}_{n\in\mathbb{N}}$ satisfying

$$\limsup_{i\to\infty} \frac{1}{n_i} \log M_{n_i} \le \overline{H}^*(X) + \gamma_0 \tag{23}$$

and

$$\lim_{i\to\infty} \varepsilon_{n_i} = 0 \tag{24}$$

for some subsequence $\{n_i\}_{i\in\mathbb{N}}$. Setting $\mathcal{N} = \{n_i : i \ge 1\}$, we define a new code $\widetilde{C} = \{(\tilde{\varphi}_n, \tilde{\psi}_n)\}_{n\in\mathbb{N}}$ from C and C' as follows:

$$\tilde{\varphi}_n = \begin{cases} \varphi_n & \text{if } n \notin \mathcal{N}, \\ \varphi_n' & \text{if } n \in \mathcal{N}, \end{cases} \qquad \tilde{\psi}_n = \begin{cases} \psi_n & \text{if } n \notin \mathcal{N}, \\ \psi_n' & \text{if } n \in \mathcal{N}. \end{cases}$$

From (22) and (24), the error probability of $\widetilde{C}$ satisfies

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

In addition, from (21), (23) and $\overline{H}^*(X) + \gamma_0 \le \overline{H}(X) - \gamma_0$, the coding rate of $\widetilde{C}$ satisfies

$$\limsup_{n\to\infty} \frac{1}{n} \log M_n = \overline{H}(X).$$

Therefore, $\widetilde{C}$ turns out to be $\overline{H}$-optimal.

However, by using (18) and the properties of the limits inferior and superior, it follows that

$$\overline{H}(X) = \liminf_{n\to\infty} \frac{1}{n} \log M_n \le \liminf_{i\to\infty} \frac{1}{n_i} \log M_{n_i}$$

$$\le \limsup_{i\to\infty} \frac{1}{n_i} \log M_{n_i} \le \overline{H}^*(X) + \gamma_0,$$

which contradicts $\overline{H}^*(X) + 2\gamma_0 \le \overline{H}(X)$ and $\gamma_0 > 0$. Hence, $\overline{H}^*(X) = \overline{H}(X)$ must hold. Q.E.D.

5. Concluding Remarks

In this letter, we have defined a stronger notion of the optimistic optimality of FF codes. It is shown that the infimum achievable strongly optimistic coding rates coincides with $\overline{H}^*(X)$ as in the ordinary case. By using this property, we have proved that the rates of all the optimal FF codes converge to $\overline{H}(X)$ if and only if the source X satisfies $\overline{H}(X) = \overline{H}^*(X)$, which means that the right endpoint of the entropy spectrum converges.

Acknowledgments

References

[1] M. Arimura, H. Koga, and K. Iwata, "Redundancy-optimal FF codes for a general source and its relationships to the rate-optimal FF codes," IEICE Trans. Fundamentals, vol.E96-A, no.12, pp.2332–2342, Dec. 2013.

[2] P.-N. Chen and F. Alajaji, "Optimistic Shannon coding theorem for arbitrary single-user systems," IEEE Trans. Inf. Theory, vol.45, no.7, pp.2623–2629, Nov. 1999.

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

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

[5] T.S. Han and S. Verdú, "Approximation theory of output statistics," IEEE Trans. Inf. Theory, vol.39, no.3, pp.752–772, May 1993.

[6] M. Hayashi, "Second-order asymptotics in fixed-length source coding and intrinsic randomness," IEEE Trans. Inf. Theory, vol.54, no.10, pp.4619–4637, Oct. 2008.

[7] H. Koga, "Four limits in probability and their roles in source coding," IEICE Trans. Fundamentals, vol.E94-A, no.11, pp.2073–2082, Nov. 2011.

[8] N. Sato and H. Koga, "New results on optimistic source coding," IEICE Trans. Fundamentals, vol.E87-A, no.10, pp.2577–2580, Oct. 2004.

[9] S. Vembu, S. Verdú, and Y. Steinberg, "The source-channel separation theorem revisited," IEEE Trans. Inf. Theory, vol.41, no.1, pp.44–54, Jan. 1995.