数理解析研究所講究録
第 1852 巻 2013 年 127-138

Error Probability in Information Transmission over Gaussian Channels with Feedback

Shunsuke Ihara
Nagoya University

1 Introduction

In this paper we discuss on the error probabilities in information transmission over Gaussian channels (GC's) with feedback. The GC is a communication channel with an additive Gaussian noise. A GC is called a white Gaussian channel (WGC), if the noise is a white Gaussian noise. The GC is one of most important communication channels not only from the theoretical point of view but also from the view point of applications. In information transmission over GC's, it is known that the minimum error probability, under the average power constraint, converges to zero quite rapidly. In 1966, Shalkwijk and Kailath [1] proposed a coding scheme and demonstrated that the resulting error probability converges to zero double exponentially fast. Then it has been shown that, for any positive integer K, there exists a coding scheme under which the error probability $P_e(T)$ at time T decreases more rapidly than the exponential of order K (cf. [1]–[7]). Recently Gallager and Nakiboğlu [7] proposed a new coding scheme for the discrete-time WGC and proved a stronger result on the multiple-exponential decay of the error probability.

The first aim of this paper is to prove a stronger result on the asymptotic behavior of the error probability for the continuous-time WGC. The second one is to generalize the result due to Gallager and Nakiboğlu [7] to the discrete-time GC, where the additive noise is not necessarily white.

The continuous-time WGC is presented by

$$\dot{Y}(t) = X(t) + \dot{B}(t), \quad t > 0,$$

or equivalently

$$Y(t) = \int_0^t X(u)du + B(t), \quad t > 0, \tag{1.1}$$

where $X(t)$ and $Y(t)$ are the input signal and the output signal, respectively, at time t, and the noise $\{B(t)\}$ is a Brownian motion ($\{\dot{B}(t)\}$ is a Gaussian white noise). The discrete-time GC is presented by

$$Y(n) = X(n) + Z(n), \quad n = 0, 1, 2, ..., \tag{1.2}$$

where $X(n)$ and $Y(n)$ are the input signal and the output signal, respectively, and the additive noise $\{Z(n)\}$ is a Gaussian process. The GC is said to be a stationary Gaussian channel (SGC) if $\{Z(n)\}$ is a stationary Gaussian process, and a WGC if $\{Z(n)\}$ is independent random sequence with identical Gaussian distribution.

We assume in (1.1) and (1.2) that the input signal satisfies the average power constraint

$$E[X(t)^2] \leq P, \quad \forall t, \tag{1.3}$$

where $P > 0$ is a constant. We also assume that the feedback link is noiseless and without time-lag. Let T be the terminal time of the information transmission. Then the message $U_0 \equiv U_0(T)$ to be transmitted is a random variable such that

$$\Pr(U_0 = m) = \frac{1}{M_T}, \quad m \in \mathcal{M}_T \equiv \{1, 2, ..., M_T\}, \tag{1.4}$$

where U_0 is independent of the noise, $M \equiv M_T = \lfloor e^{RT} \rfloor$ ($\lfloor x \rfloor$ is the maximum integer not greater than x) and $R > 0$ is a constant. R is said to be the rate of the message U_0. Given a coding scheme (encoding and decoding), we can reproduce the decoding message $\widehat{U}(t)$ as a function of the output $Y_0^t \equiv \{Y(u); u \leq t\}$. We denote by

$$P_e(t) \triangleq \Pr(\widehat{U}(t) \neq U_0), \quad t \leq T, \tag{1.5}$$

the error probability at time t. The rate R is said to be achievable, if there exists a coding scheme such that the resulting error probability satisfies $\lim_{T\to\infty} P_e(T) = 0$. The capacity C of the channel is the supremum of achievable rates.

We introduce the notation

$$\exp_n(x) \triangleq \exp\{\exp_{n-1}(x)\}, \quad n = 1, 2, ..., \tag{1.6}$$

to denote the exponential function of order n, where $\exp_0(x) = x$. Then the results obtained in ([1]–[7]) can be stated as follows. For any positive integer K, there exist a coding scheme under which the error probability $P_e(T)$ decreases more rapidly than the exponential of order K, i.e.

$$P_e(T) = o\left(\frac{1}{\exp_K(T)}\right), \quad T \to \infty. \tag{1.7}$$

For the discrete-time WGC, Gallager and Nakiboğlu [7] proposed a coding scheme and successfully demonstrated the multiple-exponential decay (1.7) of the resulting error probability at all rates below capacity. In addition, they pointed out that the error probability decreases with an exponential order which is linearly increasing with block length T, i.e. for *some* positive constant α, there exists a coding scheme under which the error probability $P_e(T)$ decreases as

$$P_e(T) = o\left(\frac{1}{\exp_{\lfloor \alpha T \rfloor}(T)}\right), \quad T \to \infty. \tag{1.8}$$

Needless to say, the known result (1.7) is an easy consequence of (1.8). We can generalize the results (1.7) and (1.8) by Gallager and Nakiboğlu [7] to discrete-time SGC's ([8]).

In this paper we also treat the continuous-time WGC with feedback and prove a stronger result on the multiple-exponential decay of the error probability. More precisely, we shall show that, for *any* positive constant α, there exists a coding scheme under which the error probability $P_e(T)$ decreases as (1.8). It should be emphasized that the order $\lfloor \alpha T \rfloor$ of the exponent in (1.8) is linear in T and the coefficient α may be taken arbitrarily large.

The continuous-time WGC is treated in Section 2. We propose our coding scheme in §2.1, and give the formula to calculate the error probability in Theorem 1. The asymptotic behavior of the error probability is evaluated in §2.2, and one of main result (1.8) is given in Theorem 2. The discrete-time SGC is treated in Section 3. The Gallager-Nakiboğlu scheme for the SGC is explained in §3.1. The multiple exponential decay of the error probability for the discrete-time SGC is shown in §3.2 (see Theorem 4).

2 Continuous-Time White Gaussian Channel

2.1 Coding Scheme and Error Probability

In this section, we treat the continuous-time WGC (1.1) with feedback. We assume that the average power constraint (1.3) is imposed on the input signals. The constraint (1.3) may be replaced by

$$\frac{1}{T}\int_0^T E[X(t)^2]dt \leq P, \tag{2.1}$$

or

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T E[X(t)^2]dt \leq P. \tag{2.2}$$

It is well known that the capacity C of the WGC subject to (1.3) (or (2.1), (2.2)) is not increased by feedback and is equal to $C = P/2$.

Let $U_0 \equiv U_0(T)$ be the message given by (1.4). The input signal $X(t)$ and the decoding message $\widehat{U}(t)$ at t are given in the forms

$$X(t) = \varphi_T(t, U_0, Y_0^{t-}) \quad \text{and} \quad \widehat{U}(t) = \psi_T(t, Y_0^t),$$

respectively, where $\varphi_T : (0,T] \times \mathcal{M}_T \times \mathrm{R}^{(0,T)} \to \mathrm{R}$, $\psi_T : (0,T] \times \mathrm{R}^{(0,T]} \to \mathcal{M}_T$ are measurable functions, $Y_0^{t-} \equiv \{Y(u); u < t\}$ and $Y_0^t \equiv \{Y(u); u \leq t\}$.

Let us define our coding scheme. To define the encoding scheme and decoding scheme, the schemes investigated in [1, 3, 6, 7] are helpful and useful. We divide the time axis $(0,\infty)$ into sub-intervals $\mathbf{T}_k = (T_{k-1}, T_k]$, $k = 1, 2, \ldots$, $(0 = T_0 < T_1 < T_2 < \cdots)$ and denote by $|\mathbf{T}_k| = T_k - T_{k-1}$ the length of the interval $\mathbf{T}_k$. Although, the length $|\mathbf{T}_1| = T_1$ of the first sub-interval $\mathbf{T}_1$ should be long enough, the lengths of other intervals $\mathbf{T}_k$, $k \geq 2$, may be chosen arbitrary. For example, we may define as $T_k = T_{k-1} + \Delta$, $k \geq 2$, where $\Delta > 0$ is an arbitrary constant. On each interval $\mathbf{T}_k$, the WGC (1.1) can be rewritten in the form

$$Y_k(t) = \int_0^t X_k(u)du + B_k(t), \quad 0 < t \leq |\mathbf{T}_k|, \tag{2.3}$$

where $X_k(t) = X(t + T_{k-1})$, $Y_k(t) = Y(t + T_{k-1}) - Y(T_{k-1})$ and $B_k(t) = B(t + T_{k-1}) - B(T_{k-1})$. Note that each $B_k \equiv \{B_k(t); 0 < t \leq |\mathbf{T}_k|\}$ is a Brownian motion and $B_1, B_2, \ldots$, are mutually independent. We also note that, for each k, (2.3) presents a WGC. The receiver decodes at time T_{k-1} and denotes by $\widehat{U}(T_{k-1})$ the decoding message. The definition of $\widehat{U}(T_{k-1})$ will be given by (2.9).

On the interval $\mathbf{T}_k$, the transmitter inputs the signal $X_k(t)$ given by

$$X_k(t) = \sqrt{P}a_k \exp(Pt/2)(W_k - \widehat{W}_k(t)), \quad 0 < t \leq |\mathbf{T}_k|, \tag{2.4}$$

where

$$W_k \triangleq U_0 - \widehat{U}(T_{k-1}) \tag{2.5}$$

is the decoding error at time T_{k-1},

$$\widehat{W}_k(t) \triangleq \sqrt{P}\, a_k^{-1} \int_0^t \exp(-Pu/2) dY_k(u), \tag{2.6}$$

and $a_k > 0$ is a constant satisfying

$$a_k^2 E[W_k^2] = a_k^2 E[|U_0 - \widehat{U}(T_{k-1})|^2] = 1. \tag{2.7}$$

Intuitively speaking, the transmitter sends linearly scaled versions of the decoding error with help of linear feedback. If the message U_0 is correctly decoded at T_{k-1}, i.e. $\widehat{U}(T_{k-1}) = U_0$, then $W_k = 0$, meaning that no signals are input on the interval $\mathbf{T}_k$. This is one of basic ideas to define an optimal coding scheme under the average power constraint (cf. [3, 6, 7]). Having received the output signal

$$Y_k(t) = \sqrt{P} a_k \int_0^t \exp(Pu/2)(W_k - \widehat{W}_k(u))du + B_k(t), \quad 0 < t \leq |\mathbf{T}_k|, \tag{2.8}$$

the receiver reproduces at time T_k a decoded message $\widehat{U}(T_k)$ defined by

$$\widehat{U}(T_k) \triangleq \widehat{U}(T_{k-1}) + \widetilde{W}_k, \quad k = 1, 2, ..., \tag{2.9}$$

where

$$\widetilde{W}_k = \left\lceil \frac{\widehat{W}_k(|\mathbf{T}_k|)}{a_k(1 - \exp(-P|\mathbf{T}_k|))} - \frac{1}{2} \right\rceil \tag{2.10}$$

($\lceil x \rceil$ denotes the minimum integer not less than x) and $\widehat{U}(T_0) \triangleq E[U_0] = (M+1)/2$. We can show that W_k is independent of $B_k = \{B_k(t); 0 < t \leq |\mathbf{T}_k|\}$, $E[W_k] = 0$ and that $E[X_k(t)^2] = P$, so that the constraint (1.3) is satisfied with equality. It is clear from (2.5) and (2.9) that $\{W_k\}$ satisfies

$$W_{k+1} = W_k - \widetilde{W}_k. \tag{2.11}$$

So far, we have defined $\widehat{U}(t)$ only for $t = T_k$. We may define $\widehat{U}(t)$ simply by

$$\widehat{U}(t) \triangleq \widehat{U}(T_k), \quad T_k \leq t < T_{k+1}.$$

Corresponding to the feedback channel (2.8), the channel without feedback is presented by

$$\widetilde{Y}_k(t) = \sqrt{P} a_k \int_0^t \exp(Pu/2) W_k du + B_k(t) = \int_0^t \xi_k(u) du + B_k(t), \quad 0 < t \leq |\mathbf{T}_k|, \tag{2.12}$$

where

$$\xi_k(t) \triangleq \sqrt{P} a_k \exp(Pt/2) W_k, \quad 0 < t \leq |\mathbf{T}_k|. \tag{2.13}$$

The covariance function of $\{\xi_k(t)\}$ is given by

$$R(t,u) = E[\xi_k(t)\xi_k(u)] = P \exp\{P(t+u)/2\}. \tag{2.14}$$

Applying the linear filtering theory (see [9, 10, 11]), one can show that the linear subspace $\mathcal{L}_t(Y_k)$ spanned by $Y_k(s)$, $s \leq t$, coincides with $\mathcal{L}_t(\widetilde{Y}_k)$ and that the orthogonal projection $\widehat{\xi}_k(t)$ of $\xi_k(t)$ on $\mathcal{L}_t(Y_k) = \mathcal{L}_t(\widetilde{Y}_k)$ is given by

$$\widehat{\xi}_k(t) = P \exp(Pt/2) \int_0^t \exp(-Pu/2) dY_k(u) = P \exp(-Pt/2) \int_0^t \exp(Pu/2) d\widetilde{Y}_k(u). \tag{2.15}$$

Hence we know from (2.6), (2.13) and (2.15) that $\widehat{W}_k(t)$ is nothing but the orthogonal projection of W_k on $\mathcal{L}_t(Y_k) = \mathcal{L}_t(\widetilde{Y}_k)$ and satisfies

$$a_k \widehat{W}_k(t) = \sqrt{P} \exp(-Pt) \int_0^t \exp(Pu/2) d\widetilde{Y}_k(u). \tag{2.16}$$

A random variable $\widehat{B}_k(t)$ defined by

$$\widehat{B}_k(t) = \frac{\sqrt{P} \exp(-Pt)}{1 - \exp(-Pt)} \int_0^t \exp(Pu/2)\, dB_k(u) \tag{2.17}$$

will play important roles to evaluate the error probability. Since $\{B_k(t)\}$ is a Brownian motion, $\widehat{B}_k(t)$ is a Gaussian random variable with expectation 0 and variance

$$E[\widehat{B}_k(t)^2] = \frac{P \exp(-2Pt)}{(1 - \exp(-Pt))^2} \int_0^t \exp(Pu)\, du = \frac{1}{\exp(Pt) - 1}. \tag{2.18}$$

It is easy to see from (2.12), (2.16) and (2.17) that

$$a_k \widehat{W}_k(t) = (1 - \exp(-Pt)) \left(a_k W_k + \widehat{B}_k(t) \right). \tag{2.19}$$

Using (2.10), (2.11) and (2.19), we can easily show that the decoding error $W_{k+1} = U_0 - \widehat{U}(T_k)$ at time T_k is presented by

$$W_{k+1} = U_0 - \widehat{U}(T_k) = -\left\lceil \frac{\widehat{B}_k(|\mathbf{T}_k|)}{a_k} - \frac{1}{2} \right\rceil. \tag{2.20}$$

A formula for the error probability $Pr(\widehat{U}(T_k) \neq U_0)$ is given in the following theorem.

Theorem 1. Under the coding scheme proposed in this section, the error probability $P_e(T_k) = Pr(\widehat{U}(T_k) \neq U_0)$ is given by

$$P_e(T_k) = 2Q(b_k a_k), \quad k \geq 1, \tag{2.21}$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-y^2/2\right) dy \tag{2.22}$$

is the complementary distribution function of $N(0,1)$ and

$$b_k = \frac{\sqrt{\exp(P|\mathbf{T}_k|) - 1}}{2}. \tag{2.23}$$

Proof. It is clear from (2.20) that

$$P_e(T_k) = Pr\left(\widehat{U}(T_k) \neq U_0\right) = Pr(W_{k+1} \neq 0) = Pr\left(|\widehat{B}_k(|\mathbf{T}_k|)| \geq a_k/2\right). \tag{2.24}$$

Since $\widehat{B}_k(|\mathbf{T}_k|) \sim N(0, (2b_k)^{-2})$ (see (2.18)), (2.21) follows from (2.24). □

2.2 Asymptotic Behavior of Error Probability

Let us evaluate the asymptotic behavior of the error probability and prove our main result (1.8) in the following theorem.

Theorem 2. Assume that the WGC (1.1) with feedback is subject to the average power constraint (1.3) and that the rate R is less the capacity C. Then, for any constant $\alpha > 0$, there exists a coding scheme such that the resulting error probability $P_e(T) = Pr\left(\widehat{U}(T) \neq U_0\right)$ satisfies

$$\lim_{T\to\infty} P_e(T)\ \exp_{\lfloor \alpha T \rfloor}(T) = 0. \tag{2.25}$$

The following lemma is useful to evaluate the asymptotic behavior of the error probability (see Gallager and Nakiboğlu [7] for the proof).

Lemma 1 ([7]). Let η be a Gaussian random variable with distribution $N(0, \sigma^2)$. Then the second moment $E[\tilde{\eta}^2]$ of a random variable $\tilde{\eta} \triangleq \left\lceil \frac{\eta}{c} - \frac{1}{2} \right\rceil$ is upper-bounded by

$$E[\tilde{\eta}^2] \leq \frac{1.6\sigma}{c} \exp\left(-\frac{c^2}{8\sigma^2}\right), \tag{2.26}$$

where $c \geq 4\sigma$ is a constant.

Remark. In [7], the random variable $\tilde{\eta}$ is called the c-quantization of η.

We proceed to prove Theorem 2.

Proof of Theorem 2. Let δ, D_0 and Δ be positive constants such that

$$R < (1-\delta)C, \quad D_0 < (1-\delta)C - R, \quad \Delta < \delta/\alpha. \tag{2.27}$$

Let T be the terminal time of the information transmission. We use the coding scheme proposed in §2.1, where T_k is determined by

$$T_k = (1-\delta)T + (k-1)\Delta, \quad k = 1, 2, \ldots, \tag{2.28}$$

and denote by $Pe(T_k) = Pr(\widehat{U}(T_k) \neq U_0)$ the resulting error probability. Note that $|\mathbf{T}_k| = \Delta$ and $b_k = 2^{-1}\sqrt{\exp(P\Delta) - 1}$ $(k \geq 2)$ do not depend on k. Since the error probability is given by $P_e(T_k) = 2Q(b_k a_k)$ (Theorem 1), to evaluate the asymptotic behavior of the error probability $P_e(T_k)$, it is sufficient to examine the asymptotic behavior of $\{a_k\}$. Note the asymptotic behavior

$$Q(x) \sim \frac{1}{\sqrt{2\pi}x} \exp\left(-x^2/2\right), \quad x \to \infty, \tag{2.29}$$

where $f(x) \sim g(x)$ $(x \to \infty)$ means $\lim_{x\to\infty} f(x)/g(x) = 1$. We also note that

$$2Q(x) < \exp\left(-x^2/2\right), \quad \forall x \geq a_0, \tag{2.30}$$

where $a_0 > 0$ is a constant. Then, it is clear from (2.21) and (2.30) that

$$P_e(T_k) = 2Q(b_k a_k) < \exp\left(-\frac{b_k^2 a_k^2}{2}\right), \quad k \geq 1. \tag{2.31}$$

Since $\widehat{B}_k(|\mathbf{T}_k|) \sim N(0,(2b_k)^{-2})$, applying Lemma 1 to (2.20), we obtain

$$E[|U_0 - \widehat{U}(T_k)|^2] \le \frac{1.6}{2b_k a_k} \exp\left(-\frac{b_k^2 a_k^2}{2}\right) < \frac{1}{b_k a_k} \exp\left(-\frac{b_k^2 a_k^2}{2}\right). \tag{2.32}$$

Then, noting (2.7), we have the key inequality

$$a_{k+1}^2 \ge b_k a_k \exp\left(\frac{b_k^2 a_k^2}{2}\right), \quad k \ge 1. \tag{2.33}$$

Since $|\mathbf{T}_1| = (1-\delta)T$, we can easily see that $b_1^2 \sim \exp\{2(1-\delta)CT\}/4$. Then, noting (2.27), we have

$$b_1^2 a_1^2 > \exp(2D_0 T), \quad T \to \infty. \tag{2.34}$$

Then, using (2.33), we can prove the inequality

$$b_k^2 a_k^2 > b_k^2 b_{k-1} a_{k-1} \exp_{k-1}\left\{\frac{\exp(2D_0T)}{2}\right\}, \quad k \ge 2, \tag{2.35}$$

by induction. We may assume $b_k^2 b_{k-1} a_{k-1} > 2$, because a_{k-1} is large enough. Then, combining (2.31), (2.34) and (2.35), we can obtain the inequality

$$P_e(T_k) \exp_k\left\{\frac{\exp(2D_0T)}{2}\right\} < 1, \quad k \ge 1. \tag{2.36}$$

Since $\exp_k(T) = o\left(\exp_k\{\exp(2D_0T)/2\}\right)$, it is clear from (2.36) that

$$P_e(T_k)\exp_k(T) = o(1). \tag{2.37}$$

Since $\Delta < \delta/\alpha$,

$$T_{\lfloor \alpha T \rfloor} < (1-\delta)T + \lfloor \alpha T \rfloor \Delta < (1-\delta)T + \delta T = T$$

and

$$P_e(T) \le P_e(T_{\lfloor \alpha T \rfloor}).$$

Then, putting $k = \lfloor \alpha T \rfloor$ in (2.37), we obtain the desired result (2.25). □

3 Discrete-Time Stationary Gaussian Channel

3.1 Coding Scheme and Error Probability

In this section, we treat the discrete-time SGC (1.2) with feedback and subject to the average power constraint (1.3), where the additive noise is a regular stationary Gaussian process. We assume without loss of generality that the distribution of $Z(n)$ is $N(0,1)$ ($Z(n) \sim N(0,1)$). If the Gaussian noise $\{Z(n)\}$ is an i.i.d., the channel (1.2) is memoryless, that is, a WGC. The terminal time N of the channel uses and the rate $R > 0$ are fixed, unless otherwise mentioned. Let $U_0 \equiv U_0(N)$ be the message given by (1.4). Then the input signal $X(n)$ is some function $\varphi_N(U_0, Y_0^{n-1})$ of the message U_0 and previous outputs $Y_0^{n-1} \equiv (Y(0), Y(1), ..., Y(n-1))$. Let $\widehat{U}(n)$ be a decoded message, at time n, which is a function $\psi_N(Y_0^n)$ of Y_0^n.

Let us define the GN scheme (Gallager and Nakiboğlu scheme) for the SGC (1.2) with feedback. We divide the time duration $\mathbf{N} = \{0, 1, 2, ..., N\}$ of the communication into three stages, $\mathbf{N}_0 = \{0\}$, $\mathbf{N}_1 = \{1, 2, ..., n_1\}$ and $\mathbf{N}_2 = \{n_1 + 1, n_1 + 2, ..., N\}$. The input signal $X(0)$ of the channel (1.2) at time 0 is given by

$$X(0) = \sqrt{P}\beta_0 (U_0 - E[U_0]) = \sqrt{P}\beta_0 \left(U_0 - \frac{M+1}{2}\right), \tag{3.1}$$

where $\beta_0 = \sqrt{12/(M^2 - 1)}$ is a normalizing constant. Since $E[U_0] = (M + 1)/2$ and $V[U_0] = E[(U_0 - E[U_0])^2] = (M^2 - 1)/12$, we have $E[X(0)] = 0$ and $V[X(0)] = E[X(0)^2] = P$. Note that, since $M = e^{RN}$,

$$\beta_0 \sim \sqrt{12}\exp(-RN), \quad N \to \infty.$$

The received signal $Y(0) = X(0) + Z(0)$ is fed back to the transmitter, which, knowing $X(0)$, determines $Z(0)$. Therefore, to transmit the message U_0, it is enough to send the Gaussian message $Z(0) \sim N(0, 1)$.

The channel use at time 0 is rather auxiliary and distinct from the others. Essentially speaking, the GN scheme is a combination of the Shalkwijk and Kailath scheme (SK scheme) on the time duration $\mathbf{N}_1$ and the high signal-to-noise ratio (high-SNR) scheme on $\mathbf{N}_2$. The terminal time n_1 of $\mathbf{N}_1$ will be specified later (see Theorem 4). The SK scheme on $\mathbf{N}_1$ is a linear scheme, where the input signal $X(n)$ at time n is defined by

$$X(n) = \beta_n(Z(0) - \hat{Z}(n-1)), \quad n \in \mathbf{N}_1, \tag{3.2}$$

here $\hat{Z}(n-1) \triangleq E[Z(0)|Y_1^{n-1}]$ $(n \geq 2)$ is the conditional expectation of $Z(0)$ given Y_1^{n-1}, $\hat{Z}(0) = 0$ and β_n is a constant satisfying

$$\beta_n^2 = \frac{P}{E[(Z(0) - \hat{Z}(n-1))^2]}. \tag{3.3}$$

Clearly the constraint (1.3) is satisfied with equality. Noting that the conditional expectation $E[Z(n)|Z(0) - \hat{Z}(n-1)]$ is written in the form

$$E[Z(n)|Z(0) - \hat{Z}(n-1)] = c_n(Z(0) - \hat{Z}(n-1))$$

(c_n is a constant), we define the sign of β_n as $\mathrm{sgn}(\beta_n) = \mathrm{sgn}(c_n)$ $(c_n \neq 0)$ and $\mathrm{sgn}(\beta_n) > 0$ $(c_n = 0)$. The decoded message $\hat{U}(n)$ $(n \in \mathbf{N}_1)$ is defined as

$$\hat{U}(n) = \left\lceil \frac{Y(0) - \hat{Z}(n)}{\sqrt{P}\beta_0} + \frac{M}{2} \right\rceil, \quad n \in \mathbf{N}_1. \tag{3.4}$$

We can easily show that the decoding error is given by

$$\hat{U}(n) - U_0 = \left\lceil \frac{Z(0) - \hat{Z}(n)}{\sqrt{P}\beta_0} - \frac{1}{2} \right\rceil. \tag{3.5}$$

Let us define the high-SNR scheme on $\mathbf{N}_2$. The input signal at time n is defined by

$$X(n) = \alpha_n(U_0 - \hat{U}(n-1)), \quad n \in \mathbf{N}_2, \tag{3.6}$$

where $\widehat{U}(n-1)$ is the decoded message at the preceding time $n-1$ and $\alpha_n > 0$ is a constant given by

$$\alpha_n^2 = \frac{P}{E[(U_0 - \widehat{U}(n-1))^2]}. \tag{3.7}$$

Clearly the constraint (1.3) is satisfied with equality. The decoded message $\widehat{U}(n)$ ($n \in \mathbf{N}_2$) is defined by

$$\widehat{U}(n) = \widehat{U}(n-1) + \frac{\widetilde{Y}(n)}{\alpha_n}, \quad n \in \mathbf{N}_2, \tag{3.8}$$

where $\widetilde{Y}(n)$ is the α_n-quantization of the output $Y(n)$.

Theorem 3. Under the GN scheme defined above, the error probability $P_e(n) = \Pr(\widehat{U}(n) \neq U_0)$ is given by

$$P_e(n) = \begin{cases} 2Q\left(\beta_0|\beta_{n+1}|/2\right), & n \in \mathbf{N}_1, \\ 2Q(\alpha_n/2), & n \in \mathbf{N}_2, \end{cases} \tag{3.9}$$

where $Q(x)$ is the complementary distribution function of $N(0,1)$ (see (2.22)).

Proof. Let $n \in \mathbf{N}_1$. Then, since $\widehat{U}(n) - U_0$ is the $\sqrt{P}\beta_0$-quantization of $Z(0) - \widehat{Z}(n) \sim N(0, P\beta_{n+1}^{-2})$ (see (3.5)), we can easily show

$$\Pr(\widehat{U}(n) - U_0 = j) = \frac{1}{\sqrt{2\pi}} \int_{(j-1/2)\beta_0|\beta_{n+1}|}^{(j+1/2)\beta_0|\beta_{n+1}|} \exp\left(-x^2/2\right) dx, \quad j = 0, \pm 1, \pm 2, ..., \tag{3.10}$$

and

$$P_e(n) = \Pr(\widehat{U}(n) - U_0 \neq 0) = \frac{2}{\sqrt{2\pi}} \int_{\beta_0|\beta_{n+1}|/2}^{\infty} \exp\left(-x^2/2\right) dx = 2Q(\beta_0|\beta_{n+1}|/2).$$

Thus, we have obtained the first equation of (3.9). Let $n \in \mathbf{N}_2$. Then, it is clear from (3.6) and (3.8) that

$$U_0 - \widehat{U}(n) = U_0 - \widehat{U}(n-1) - \frac{\widetilde{Y}(n)}{\alpha_n} = U_0 - \widehat{U}(n-1) - \frac{X(n) + \widetilde{Z}(n)}{\alpha_n} = -\frac{\widetilde{Z}(n)}{\alpha_n}, \tag{3.11}$$

where $\widetilde{Z}(n)$ is the α_n-quantization of $Z(n) \sim N(0,1)$. Therefore,

$$P_e(n) = \Pr(\widetilde{Z}(n) \neq 0) = \Pr(|Z(n)| \geq \alpha_n/2) = 2Q(\alpha_n/2).$$

Thus, we have obtained the second euation of (3.9). □

3.2 Asymptotic Behavior of Error Probability

A characterization of the feedback capacity C of the SGC has been given by Kim [12] as the solution to a variational problem. In particular, a closed-form expression of the capacity is given for the first-order autoregressive moving-average (ARMA) SGC ([12, 13]). However, in general, no explicit formulas are available for the capacity. In this paper, instead of the capacity C, we need to introduce a constant C^* by

$$C^* \triangleq \liminf_{n \to \infty} \frac{1}{n} \log |\beta_n|, \tag{3.12}$$

where β_n is the constant given by (3.2). It will be shown that $C^* \leq C$ (see Corollary 1). Then, we can prove the multiple-exponential decay of the error probability at all rates below C^*.

Let us evaluate the error probability under the coding scheme proposed in §3.1. We can prove the multiple-exponential decay of the error probability in the following theorem.

Theorem 4. Assume that the SGC (1.2) with feedback is subject to the average power constraint (1.3) and that the rate R is less than C^*. For any positive integer K, let D_k, $k = 0, 1, ..., K$, and $\delta > 0$ be constants such that

$$0 < D_K < \cdots < D_1 < D_0 < (1-\delta)C^* - R. \tag{3.13}$$

Under the GN scheme, the error probability $P_e(n)$, $n \in \mathbf{N}_2$, is upper-bounded by

$$P_e(n_1 + k) < \exp\{-\exp_{k+1}(2D_{k-1}N)\}, \quad k = 1, ..., K, \tag{3.14}$$

if $n_1 > (1-\delta)N$ and N is large enough, where n_1 is the terminal time of $\mathbf{N}_1$. Moreover, for any constant α such that $0 < \alpha < (C^* - R)/C^*$, we have

$$\lim_{N\to\infty} P_e(N) \exp_{\lfloor \alpha N \rfloor}(N) = 0. \tag{3.15}$$

Proof. Let $\widetilde{D}_k$, $k = 0, 1, ..., K$, be constants such that $0 < \widetilde{D}_{k+1} < D_k < \widetilde{D}_k < D_0 < \widetilde{D}_0 < (1-\delta)C^* - R$. Since $\widehat{U}(n) - U_0$ is the $\sqrt{P}\beta_0$-quantization of $Z(0) - \widehat{Z}(n) \sim N(0, P\beta_{n+1}^{-2})$, applying Lemma 1 and using (3.7) and (3.12), we can show that $\alpha_{n_1+1}^2$ is lower-bounded by

$$\alpha_{n_1+1}^2 > \exp_2(2\widetilde{D}_0 N). \tag{3.16}$$

Since $\widetilde{Z}(n)$ ($n \in \mathbf{N}_2$) is the α_n-quantization of $Z(n) \sim N(0,1)$, using Lemma 1, we have

$$E[\widetilde{Z}(n)^2] \leq \frac{1.6}{\alpha_n} \exp\left(-\frac{\alpha_n^2}{8}\right), \quad n \in \mathbf{N}_2.$$

Then it follows from (3.7) and (3.11) that

$$\alpha_{n+1}^2 = \frac{P\alpha_n^2}{E[\widetilde{Z}(n)^2]} \geq \frac{P\alpha_n^3}{1.6} \exp\left(\frac{\alpha_n^2}{8}\right) \geq \exp\left(\frac{\alpha_n^2}{8}\right), \quad n \in \mathbf{N}_2. \tag{3.17}$$

Using (3.16) and (3.17), by induction, we can prove the inequality

$$\alpha_{n_1+k}^2 > \exp_{k+1}(2\widetilde{D}_{k-1}N), \quad k \geq 1. \tag{3.18}$$

Since $D_{k-1} < \widetilde{D}_{k-1}$, using (3.9), (2.29) and (3.18), one can easily show (3.14):

$$P_e(n_1 + k) = 2Q(\alpha_{n_1+k}/2) < \exp\left\{-\exp_{k+1}(2D_{k-1}N)\right\}, \quad k \geq 1.$$

Let α be a constant such that $0 < \delta < \alpha < (C^* - R)/C^*$, and let $(1-\delta)N < n_1 < n_1 + K \leq N$. Then, since $P_e(N) \leq P_e(n_1 + K)$, the inequality

$$\limsup_{N\to\infty} P_e(N) \exp_{K+2}(2D_{K-1}N) \leq 1 \tag{3.19}$$

follows from (3.14). Since $K+2 \leq N-n_1+2 < \delta N+2$ and $\delta < \alpha$, we see $K+2 < \lfloor \alpha N \rfloor$ and $\exp_{\lfloor \alpha N \rfloor}(N) = o(\exp_{K+2}(2D_{K-1}N))$ for sufficiently large N. Therefore, (3.15) is an easy consequence of (3.19). □

We now briefly discuss on the capacity of the SGC.

Corollary 1. Let C be the feedback capacity of the SGC (1.2) subject to (1.3). Then

$$C^* \leq C, \tag{3.20}$$

where C^* is the constant given by (3.12).

Proof. Theorem 4 tells us that all rates below C^* are achievable. Therefore (3.20) is true. □

Example 1 (WGC). Let the channel (1.2) be the WGC subject to the constraint (1.3). Then we can easily see that C^* coincides with the capacity C and is given by

$$C^* = C = \frac{1}{2}\log(1+P).$$

Example 2 (ARMA(1,1) SGC). Let the Gaussian noise $\{Z(n)\}$ be an ARMA(1,1) process. In this case, $\{Z(n)\}$ has the representation

$$Z(n) + \beta Z(n-1) = W(n) + \alpha W(n-1),$$

where $\{W(n)\}$ is an i.i.d. with distribution $N(0,\sigma^2)$, $W(n)$ is independent of $\{Z(j); j \leq n-1\}$, $\alpha \in [-1,1]$, $\beta \in (-1,1)$ and $\sigma^2 = (1-\beta^2)/\{(1-\beta^2)+(\alpha-\beta)^2\}$. It is known (Kim [12]) that

$$C = C^* = \lim_{n\to\infty} \frac{1}{n}\log|\alpha_n| = -\log x_0,$$

where x_0 is a unique positive root of

$$Px^2 = \frac{(1-\beta^2)(1-x^2)(1+s\alpha x)^2}{((1-\beta^2)+(\alpha-\beta)^2)(1+s\beta x)^2}$$

and $s = \mathrm{sgn}(\beta-\alpha)$ ($s=0$ if $\alpha=\beta$).

4 Concluding Remarks

For the continuous-time WGC we have shown the multiple-exponential decay (1.8) of the error probability (Theorem 2), where the coefficient α may be taken arbitrarily large. We have seen that, to realize a large α, we need to take the length $\Delta = |\mathbf{T}_k|$ ($< \delta/\alpha$) of subintervals small enough. This is possible for the continuous-time GC, and the situation is different in the discrete-time GC, where the coefficient α is upper bounded by $\alpha < (C^*-R)/C^*$ (Theorem 4).

In the case of continuous-time channels, although we we have shown the multiple-exponential decay (1.8) only for WGC, it is expected that we may show (1.8) in a wide class of GC's.

References

[1] J. P. M. Schalkwijk and T. Kailath: A coding scheme for additive noise channels with feedback, Part I: No bandwidth constraints. *IEEE Trans. Inform. Theory*, IT-12 (1966), 172–182.

[2] M. S. Pinsker: The probability of error in block transmission in a memoryless Gaussian channel with feedback. *Problems Inform. Transmission*, 4 (1968), 1–14.

[3] A. J. Kramer: Improving communication reliability by use of an intermittent feedback channel. *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 52–60, Jan. 1969.

[4] K. Sh. Zigangirov: Upper bound for the error probability for Gaussian channels with feedback. *Problems Inform. Transmission*, 6 (1970), 159–163.

[5] S. Ihara and L. G. Xin: *Error exponents for discrete-time white Gaussian channels with feedback.* Proc. Meijo Winter School 2003, Quantum Information and Complexity, World Scientific, Singapore, pp. 227–243, 2004.

[6] S. Ihara: *Probability of error in information transmission over continuous-time Gaussian channel with feedback. Proc. Int. Symp. Inf. Theory*, Toronto, Canada, pp. 895–899, 2008.

[7] R. G. Gallager and B. Nakiboğlu: Variations on a theme by Schalkwijk and Kailath. *IEEE Trans. Inform. Theory*, vol. IT-56, no. 1, pp. 6–18, Jan. 2010.

[8] S. Ihara: *Upper bounds of error probabilities for stationary Gaussian channels with feedback. IEEE Trans. Inform. Theory*, vol. IT-58, no. 12, pp. 7068–7072, Dec. 2012.

[9] T. Kailath: An innovations approach to least-squares estimation, Part I: Linear filtering in additive white noise. *IEEE Trans. Auto. Control*, AC-13 (1968), 646–655.

[10] R. S. Liptser and A. N. Shiryayev: *Statistics of Random Processes I, II.* Springer Verlag, New York, 1977.

[11] S. Ihara: *Information Theory for Continuous Systems.* World Scientific, Singapore, 1993.

[12] Y.-H. Kim: Feedback capacity of stationary Gaussian channels. *IEEE Trans. Inform. Theory*, vol. IT-56, pp. 57–85, Jan. 2010.

[13] Y.-H. Kim: Feedback capacity of the first-order moving average Gaussian channel. *IEEE Trans. Inform. Theory*, vol. IT-52, pp. 3063–3079, July 2006.

Nagoya University
Nagoya, JAPAN
E-mail address: ihara@zm.commufa.jp

名古屋大学　井原 俊輔