1 Introduction

The following model for the discrete time Gaussian channel with feedback is considered:

\[ Y_n = S_n + Z_n, \quad n = 1, 2, \ldots \]

where \( Z = \{Z_n; n = 1, 2, \ldots\} \) is a non-degenerate, zero mean Gaussian process representing the noise and \( S = \{S_n; n = 1, 2, \ldots\} \) and \( Y = \{Y_n; n = 1, 2, \ldots\} \) are stochastic processes representing input signals and output signals, respectively. The channel is with noiseless feedback, so \( S_n \) is a function of a message to be transmitted and the output signals \( Y_1, \ldots, Y_{n-1} \). The signal is subject to an expected power constraint

\[ \frac{1}{n} \sum_{i=1}^{n} E[S_i^2] \leq P, \]

and the feedback is causal, i.e., \( S_i \) is independent of \( Z_i, Z_{i+1}, \ldots \) for \( i = 1, 2, \ldots, n \). Similarly, when there is no feedback, \( S_i \) is independent of \( Z^n = (Z_1, Z_2, \ldots, Z_n) \). Let \( |A| \) and \( Tr[A] \) be the determinant and trace of \( A \), respectively. Let \( R_X^{(n)} \) be the covariance matrix of \( Z^n \). It is well known that a finite block length capacity is given by

\[ C_{n,FB}(P) = \max \frac{1}{2n} \log_2 \left| R_X^{(n)} + R_Z^{(n)} \right| / \left| R_Z^{(n)} \right|, \]

where the maximum is on \( R_X^{(n)} n \times n \) nonnegative definite and \( B n \times n \) strictly lower triangular, such that

\[ Tr[(I + B)R_X^{(n)}(I + B^t) + BR_Z^{(n)}B^t] \leq nP. \]

\[ \text{supp} \] supported in part by the Grant-in-Aid for Scientific Research No.09640195, the Ministry of Education, Science and Culture, Japan
Similarly, let $C_n(P)$ be the maximal value when $B = 0$, i.e. when there is no feedback. When block length $n$ is fixed, $C_n(P)$ is given exactly (Cover and Thomas [2]). But $C_{n,FB}(P)$ has not been given exactly. So we are interested in the upper bounds to $C_{n,FB}(P)$. Several upper bounds are already given (Ebert [4], Pinsker [5], Cover and Pombra [1], Dembo [3], Yanagi [6], [7]). In particular in [7] we gave the following strongest upper bound in those obtained before when $P$ is sufficiently large.

**Proposition 1 (Yanagi [7])**

$$C_n(P) \leq C_{n,FB}(P) \leq C_n(P^*),$$

where

$$P^* = P + \sqrt{\frac{P}{n}} \left\{ \sum_{k=2}^{n} r_k + \sqrt{\text{Tr}[R_Z^{(n)}] - |R_Z^{11}(2)| - \frac{|R_Z^{11}(3)|}{|R_Z^{11}(2)|} - \cdots - \frac{|R_Z^{(n)}|}{|R_Z^{11}(n)|} } \right\},$$

$0 < r_1 \leq r_2 \leq \cdots \leq r_n$ are eigenvalues of $R_Z^{(n)}$, $R_Z^{11}(k)$ is the $(k-1) \times (k-1)$ submatrix of $R_Z^{(n)}$ generated by $1, \ldots, k-1$ rows and $1, \ldots, k-1$ columns.

When $P$ is sufficiently small, $C_n(P^*)$ is not a strong upper bound. Because $P^*$ has term $\sqrt{P}$. So we need the another upper bounds which are strong for small $P$. In this paper we obtain a tight upper bound of $C_{n,FB}(P)$, which is represented by $C_n(P_1)$, which is useful for small $P$. In Section 2 we obtain the exact representation $P_1$. For the sake of simplicity, we use $R_X, R_Z, \cdots$ instead of $R_X^{(n)}, R_Z^{(n)}, \cdots$.

## 2 Power Constraint $P_1$

Let $P_1 = \max \left\{ \frac{1}{n} \text{Tr}[R_X] \right\}$ under contraint (1). Then we can have $C_n(P_1)$ as an upper bound of $C_{n,FB}(P)$. So we minimize

$$\text{Tr}[B(R_X + R_Z)B^t + BR_X + R_X B^t]$$

under any strictly lower triangular matrices $B$. Let $B = \{b_{ij}; b_{ij} = 0 (i \leq j)\}$, $R_X = \{x_{ij}\}$, $R_Z = \{z_{ij}\}$. Since

$$\text{Tr}[B(R_X + R_Z)B^t + BR_X + R_X B^t]$$

$$= \sum_{k=2}^{n} \left\{ \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} (x_{ji} + z_{ji})b_{ki}b_{kj} + 2 \sum_{j=1}^{k-1} x_{jk}b_{kj} \right\},$$
we minimize
\[ \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} (x_{ji} + z_{ji})b_{ki}b_{kj} + 2 \sum_{j=1}^{k-1} x_{jk}b_{kj}, \]  
for \( 2 \leq k \leq n \). We partially differentiate (2) at \( b_{k\ell} \) \( (1 \leq \ell \leq k-1) \) and we get
\[ b_{k\ell} = -\frac{1}{|R_{X+Z}(1, \ldots, k-1)|} \sum_{i=1}^{k-1} y_{ki}x_{ik}, \]
where \( R_{X+Z}(1, \ldots, k-1) \) is the submatrix of \( R_X + R_Z \) generated by 1, \ldots, \( k-1 \) rows and 1, \ldots, \( k-1 \) columns, \( y_{ki} \) is the cofactor of \((\ell, i)\) component in \( R_{X+Z}(1, \ldots, k-1) \). Hence the minimal value of (2) is given by
\[ -\left<R_{X+Z}(1, \ldots, k-1)^{-1} \begin{pmatrix} x_{1k} \\ x_{2k} \\ \vdots \\ x_{k-1k} \\ x_{k-1k} \end{pmatrix}, \begin{pmatrix} x_{1k} \\ x_{2k} \\ \vdots \\ x_{k-1k} \\ x_{k-1k} \end{pmatrix} \right> \]
\[ = -x_k^tR_{X+Z}(1, \ldots, k-1)^{-1}x_k, \]
where \( x_k = \begin{pmatrix} x_{1k} & x_{2k} & \cdots & x_{k-1k} \end{pmatrix}^t \). Then the constraint (1) is given by
\[ x_{11} + \sum_{k=2}^{n} \{x_{kk} - x_k^tR_{X+Z}(1, \ldots, k-1)^{-1}x_k\} \leq nP. \]
By Lemma 2 in Appendix, we have
\[ x_{11} + \sum_{k=2}^{n} \frac{|R_Z(1, \ldots, k-1)|x_{kk}}{|R_Z(1, \ldots, k-1)| + \frac{x_k^t x_k}{x_{kk}}} \leq nP. \]  
Under (3) we maximize
\[ \frac{1}{n} Tr[R_X] = \frac{x_{11} + x_{22} + \cdots + x_{nn}}{n}. \]
In order to maximize \( \frac{1}{n} Tr[R_X] \) we maximize
\[ \frac{|R_Z(1, \ldots, k-1)|x_{kk}}{|R_Z(1, \ldots, k-1)| + \frac{x_k^t x_k}{x_{kk}}}, \]  
for \( 2 \leq k \leq n \). Since \( \frac{x_k^t x_k}{x_{kk}} \) has only one nonzero eigenvalue \( Tr[\frac{x_k^t x_k}{x_{kk}}] \), the maximal value of (4) is given by
\[ 1 + \frac{Tr[\frac{x_k^t x_k}{x_{kk}}]}{\lambda_{k-1}}, \]
3
where $\lambda_{k-1}$ is the smallest eigenvalue of $R_Z(1, \ldots, k-1)$. It follows from Lemma 1 in Appendix that
\[
Tr\left[\frac{x_k x_k^t}{x_k k}\right] \leq Tr[R_X(1, \ldots, k-1)] = x_{11} + x_{22} + \cdots + x_{k-1k-1}.
\]
Then we maximize $\frac{1}{n} Tr[R_X]$ under the constraint
\[
x_{11} + \frac{x_{22}}{1 + \frac{x_{11}}{\lambda_1}} + \frac{x_{33}}{1 + \frac{x_{11} + x_{22}}{\lambda_2}} + \cdots + \frac{x_{nn}}{1 + \frac{x_{11} + x_{22} + \cdots + x_{n-1n-1}}{\lambda_{n-1}}} \leq nP.
\]
It follows from Lemma 3 in Appendix that
\[
x_{11} + \frac{x_{22}}{1 + \frac{x_{11}}{\lambda_{n-1}}} + \frac{x_{33}}{1 + \frac{x_{11} + x_{22}}{\lambda_{n-1}}} + \cdots + \frac{x_{nn}}{1 + \frac{x_{11} + x_{22} + \cdots + x_{n-1n-1}}{\lambda_{n-1}}} \leq nP. \tag{5}
\]
Let $y_k = \frac{x_{kk}}{\lambda_{n-1}}$ for $1 \leq k \leq n$. Then (5) is represented by
\[
y_1 + \frac{y_2}{1 + y_1} + \frac{y_3}{1 + y_1 + y_2} + \cdots + \frac{y_n}{1 + y_1 + \cdots + y_{n-1}} \leq \frac{nP}{\lambda_{n-1}}. \tag{6}
\]
It follows from Lemma 4 in Appendix that the maximal value of
\[
\frac{x_{11} + x_{22} + \cdots + x_{nn}}{n} = \frac{\lambda_{n-1}}{n} (y_1 + y_2 + \cdots + y_n)
\]
under $y_1, y_2, \ldots, y_n \geq 0$ satisfying the condition (6) is given by
\[
\frac{\lambda_{n-1}}{n} \left\{(1 + \frac{P}{\lambda_{n-1}})^n - 1\right\}.
\]
The maximum is attained by
\[
y_k = \frac{P}{\lambda_{n-1}} (1 + \frac{P}{\lambda_{n-1}})^{k-1}, \ k = 1, 2, \ldots, n.
\]
Now we can state the second theorem.

**Theorem 1**

\[
C_n(P) \leq C_{n,FB}(P) \leq C_n(P_1),
\]

where
\[
P_1 = \frac{\lambda_{n-1}}{n} \left\{(1 + \frac{P}{\lambda_{n-1}})^n - 1\right\}.
\]
3 Appendix

Lemma 1

\[ R_X(1, \ldots, k-1) \geq \frac{x_k x_k^t}{x_{kk}}. \]

Proof. We denote the inverse matrix of \( R_X(1, \ldots, k) \) by

\[ R_X(1, \ldots, k)^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \]

where \( B_{11} \) is the submatrix of \( R_X(1, \ldots, k)^{-1} \) generated by \( 1, \ldots, k-1 \) rows and \( 1, \ldots, k-1 \) columns, \( B_{12} \) is the submatrix of \( R_X(1, \ldots, k)^{-1} \) generated by \( 1, \ldots, k-1 \) rows and \( k \) column, \( B_{21} = B_{12}^t \) and \( B_{22} \) is the submatrix of \( R_X(1, \ldots, k)^{-1} \) generated by \((k, k)\) component of \( R_X(1, \ldots, k)^{-1} \). Since \( B_{11} \geq 0 \), we get

\[ B_{11}^{-1} = R_X(1, \ldots, k-1) - \frac{x_k x_k^t}{x_{kk}} \geq 0. \]

Then we have the result. \(\square\)

Lemma 2

\[ x_{kk} - x_k^t R_{X+Z}(1, \ldots, k-1)^{-1} x_k \geq \frac{|R_Z(1, \ldots, k-1)| x_{kk}}{|R_Z(1, \ldots, k-1) + x_k x_k^t x_{kk}}. \]

Proof. By Lemma 1, we have

\[ R_{X+Z}(1, \ldots, k-1) \geq \frac{x_k x_k^t}{x_{kk}} + R_Z(1, \ldots, k-1) > 0. \]

Then

\[ -R_{X+Z}(1, \ldots, k-1)^{-1} \geq -\left(\frac{x_k x_k^t}{x_{kk}} + R_Z(1, \ldots, k-1)\right)^{-1}. \]

We denote

\[ R = \frac{x_k x_k^t}{x_{kk}} + R_Z(1, \ldots, k-1). \]
We give the following
\[
x_{kk} - x_k^t R X + Z (1, \ldots, k - 1)^{-1} x_k \\
\geq x_{kk} - x_k^t R^{-1} x_k \\
= \frac{|R| \{ x_{kk} - x_k^t R^{-1} x_k \}}{|R|} \\
= \frac{\begin{bmatrix} R & x_k \\ x_k^t & x_{kk} \end{bmatrix}}{|R|} \\
= \frac{|R - x_k x_k^t| x_{kk}}{|R|} \\
= \frac{|R Z (1, \ldots, k - 1) x_{kk}}{|R|}.
\]
\[
\]
\[
\]

Lemma 3 Let
\[
A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}
\]
be a positive definite symmetric matrix. Let \( \lambda \) be the minimal eigenvalue of \( A \) and let \( \mu \) be the minimal eigenvalue of \( A_{11} \). Then \( \lambda \leq \mu \).

Proof. Let
\[
A^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}
\]
be the inverse matrix of \( A \). Since
\[
A_{11}^{-1} = B_{11} - B_{12} B_{22}^{-1} B_{21} \leq B_{11},
\]
we have
\[
\frac{1}{\lambda} = \| A_{11}^{-1} \| \leq \| B_{11} \|.
\]
On the other hand
\[
\frac{1}{\lambda} = \| A^{-1} \| = \sup \{ < \begin{pmatrix} B_{11} \\ B_{21} \end{pmatrix}, f >; \| f \| = 1 \}
\geq \sup \{ < \begin{pmatrix} B_{11} \\ B_{21} \end{pmatrix}, \begin{pmatrix} x \\ 0 \end{pmatrix}, \begin{pmatrix} x \\ 0 \end{pmatrix} >; \| x \| = 1 \}
= \sup \{ < B_{11} x, x >; \| x \| = 1 \}
= \| B_{11} \|. 
\]
Then
\[ \frac{1}{\mu} \leq ||B_{11}|| \leq \frac{1}{\lambda}, \]
and we have the result. \qed

**Lemma 4**

\[
\max\{y_1 + y_2 + \cdots + y_n; \ y_1, y_2, \ldots, y_n \geq 0 \text{ satisfying (6)}\}
\]

\[
= \left(1 + \frac{P}{\lambda_{n-1}}\right)^n - 1.
\]

**Proof.** For simplicity we denote \( \lambda_{n-1} \) by \( \lambda \). By (6)

\[
y_n \leq \left(\frac{nP}{\lambda} - y_1 - \frac{y_2}{1+y_1} - \cdots - \frac{y_{n-1}}{1+y_1+\cdots+y_{n-2}}\right)(1 + y_1 + y_2 + \cdots + y_{n-1}).
\]

Then \( y_1 + y_2 + \cdots + y_n \) is maximized by

\[
y_{n-1} = \frac{1}{2}(1 + y_1 + y_2 + \cdots + y_{n-2})\left(\frac{nP}{\lambda} - y_1 - \cdots - \frac{y_{n-2}}{1+y_1+\cdots+y_{n-3}}\right).
\]

The maximal value is

\[
\frac{1}{4}(1 + y_1 + y_2 + \cdots + y_{n-2})\left(\frac{nP}{\lambda} - y_1 - \cdots - \frac{y_{n-2}}{1+y_1+\cdots+y_{n-3}} + 2\right)^2 - 1. \tag{7}
\]

Next (7) is maximized by

\[
y_{n-2} = \frac{1}{3}(1 + y_1 + \cdots + y_{n-3})\left(\frac{nP}{\lambda} - y_1 - \cdots - \frac{y_{n-3}}{1+y_1+\cdots+y_{n-4}}\right).
\]

The maximal value is

\[
\frac{1}{27}(1 + y_1 + \cdots + y_{n-3})\left(\frac{nP}{\lambda} - y_1 - \cdots - \frac{y_{n-3}}{1+y_1+\cdots+y_{n-4}} + 3\right)^3 - 1.
\]

We continue this procedure and get the following result. The maximal value of \( y_1 + y_2 + \cdots + y_n \) is \( \left(1 + \frac{P}{\lambda}\right)^n - 1 \), which is attained by

\[
y_1 = \frac{P}{\lambda},
\]
\[
y_2 = \frac{P}{\lambda}(1 + \frac{P}{\lambda}),
\]
\[ \vdots \]
\[
y_n = \frac{P}{\lambda}(1 + \frac{P}{\lambda})^{n-1}.
\]

\[ \square \]
References


Department of Applied Science,
Faculty of Engineering
Yamaguchi University,
Ube 755, Japan