Linear Joint Source-Channel Coding for Gaussian Sources through Fading Channels

Jin-Jun Xiao, Zhi-Quan Luo and Nihar Jindal
Dept. of Electrical and Computer Engineering, University of Minnesota
200 Union Street SE, Minneapolis, MN 55455
Emails: {xiao, luozq, nihar}@ece.umn.edu.

***Abstract*— We consider the linear coding of a discrete memoryless Gaussian source transmitted through a discrete memoryless fading channel with additive white Gaussian noise (AWGN). The goal is to minimize the mean squared error (MSE) of the source reconstruction at the destination subject to an average power constraint imposed on the channel input symbols. We show that among all single-letter (or symbol-by-symbol) codes, linear coding achieves the smallest MSE, and is thus optimal. But when block length increases, the linear coding still shares the same performance with the single-letter coding, and thus can not approach the Shannon's bound. In spite of the suboptimality, the performance loss of linear coding compared to the optimal coding can be quantitively bounded in terms of the variance of the fading gain and the average transmit power. We also show that for linear coding, when there is no transmitter channel state information (CSI), uniform power allocation is optimal, and in the presence of transmitter CSI, the optimal power allocation can be analytically solved in terms of the channel fading gains and the average power budget.**

I. Introduction

Shannon has shown in [1, Theorem 21] that in a point-to-point link, when a discrete memoryless source is transmitted through a discrete memoryless channel, the optimal tradeoff between (channel input) cost and (source reconstruction) distortion can be achieved by separate source and channel coding. Despite its conceptual beauty, in practice, to approach the optimal pair of cost and distortion, the separate source and channel coding leads to high complexity and long delay when block length increases.

Although joint source-channel coding does not have the separation property, it sometimes can lead to simple but optimal coding strategy. A well-known example is when a memoryless Gaussian source transmitted through an AWGN channel, an amplify and forward transmission strategy achieves the optimal power-distortion tradeoff [2], [3]. The perfect match between the source and channel leads to a very simple but optimal coding strategy which is both theoretically and practically appealing. Unfortunately, when source and channel do not come up with such a natural match, the simple but optimal coding is not easy to find. In this work, we study the case when the source is still Gaussian but there is fading in the channel. In such a source-channel communication system, we analyze the performance of a class of linear coding where the encoder simply maps the source symbols into the channel symbols by a linear mapping. We focus on the following questions: How optimal is the linear coding? Can linear coding achieve the Shannon's bound when block length increases? If not, how can the performance loss be bounded?

Specifically, we consider a memoryless Gaussian source $\{S(t) : t \in \mathbb{Z}^+\}$ which has instantaneously distribution $N(0, \sigma_S^2)$, and is transmitted through a discrete memoryless fading channel

$$Y(i) = h(i)X(i) + W(i), \quad i = 1, 2, \ldots,$$

where

i. $W(i)$ are AWGN with unitary variance,
ii. $h(i)$ are i.i.d. fading with known distribution h.

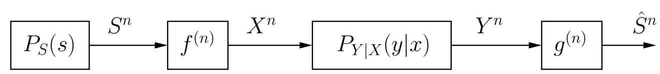

Fig. 1. A source-channel coding system where $f^{(n)}$ and $g^{(n)}$ are encoder and decoder respectively.

A general source-channel coding scheme of block length n is depicted in Fig. 1, where we adopt the notation $U^n \stackrel{\text{def}}{=} (U(1), U(2), \ldots, U(n))$ for any random vector $\{U(i) : i \in \mathbb{Z}^+\}$. The encoder $f^{(n)}$ maps source symbols S^n to channel input symbols X^n. In this problem, we consider the transmit power consumption and define $P(i) \stackrel{\text{def}}{=} E(|X(i)|^2)$. The decoder $g^{(n)}$ generates estimates of the source symbol based on the received signals Y^n, where the mean square error (MSE) is adopted to be the performance criterion. Specifically, we define the instantaneous MSE $D(i) \stackrel{\text{def}}{=} E(|S(i) - \hat{S}(i)|^2)$.

For any power-distortion pair (P, D), we say (P, D) is admissible if there are coding schemes $(f^{(n)}, g^{(n)})$ satisfying the average power constraint

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} P_i \le P, \tag{1}$$

and the achieved average MSE

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} D_i \le D. \tag{2}$$

When we limit $f^{(n)}$ to the class of linear encoding, the encoder is then an $n \times n$ matrix which maps the source symbols S^n to channel input symbols X^n. The optimal decoder is the mean square error estimator (MMSE) estimating S^n based on Y^n. In next section, we analyze the performance of linear

The work of J.-J. Xiao and Z.-Q. Luo is supported in part by the National Science Foundation, Grant No. DMS-0312416, and by the USDOD ARMY, grant number W911NF-05-1-0567.

coding and answer the three questions we proposed at the end of the second paragraph.

II. Linear Coding

In this section we study the linear coding for the source-channel problem in Fig. 1. The analysis is divided into two cases: $n = 1$ or $n \geq 2$. We assume that there is receiver channel state information (CSI), i.e., the realization of $h(i)$ is known by the decoder.

$$S(i) \longrightarrow \boxed{f(i)} \longrightarrow X(i) \longrightarrow \boxed{\text{channel}} \longrightarrow Y(i) \longrightarrow \boxed{g(i)} \longrightarrow \hat{S}(i)$$

Fig. 2. A linear coding system with block length $n = 1$.

A. Linear Coding with Block Length $n = 1$

The source-channel coding system of block length $n = 1$ is given in Fig. 2. In this case, the encoders are simply scalar multipliers $f(i)$ which scale $S(i)$ to satisfy the power constraint. Since $P(i) = E(|X(i)|^2)$, the encoder is given as

$$X(i) = \sqrt{\frac{P(i)}{\sigma_s^2}} S(i). \tag{3}$$

The decoder is the MMSE estimator

$$\hat{S}(i) = E(S(i)|Y^i) = E(S(i)|Y(i)) = \frac{P(i)|h(i)|^2}{P(i)|h(i)|^2 + 1} Y(i), \tag{4}$$

which is also linear. It is easy to calculate that with power constraint $P(i)$, the instantaneous MSE

$$D(i) \stackrel{\text{def}}{=} E(|S(i) - \hat{S}(i)|^2) = \sigma_S^2 E_{h(i)} \left\{ \frac{1}{1 + |h(i)|^2 P(i)} \right\}$$
$$= \sigma_S^2 E_h \left\{ \frac{1}{1 + |h|^2 P(i)} \right\}. \tag{5}$$

Without limiting $\{f(i), g(i)\}$ to the class of linear functions, $\{f(i), g(i)\}$ can be any functions that map $S(i) \to X(i)$ and $Y(i) \to \hat{S}(i)$ respectively. However, we have the following theorem for the optimality of linear encoding among all single-letter codes.

Theorem 1: Assume the power allocation $P(i)$ is given for all i. Then among all single-letter (or symbol-by-symbol) codes, the linear coding given in (3)-(4) is optimal.

Before proving the above theorem, we need the following lemma.

Lemma 2: Let S be a Gaussian random variable with variance σ_S^2, and $\hat{S}$ be any random variable jointly distributed with S. Then

$$\frac{E(|S - \hat{S}|^2)}{\sigma_S^2} \geq \exp\left(-2I(S; \hat{S})\right).$$

Proof: We have the following chain of inequalities:

$$\begin{aligned} I(S; \hat{S}) &= h(S) - h(S|\hat{S}) \\ &= h(S) - h(S - \hat{S}|\hat{S}) \\ &\geq h(S) - h(S - \hat{S}) \\ &\geq h(S) - \frac{1}{2}\log\left(2\pi e E(|S - \hat{S}|^2)\right) \\ &= \frac{1}{2}\log(2\pi e \sigma_S^2) - \frac{1}{2}\log\left(2\pi e E(|S - \hat{S}|^2)\right) \\ &= \frac{1}{2}\log\frac{\sigma_S^2}{E(|S - \hat{S}|^2)}. \end{aligned}$$

The proof is thus complete. ■

Proof of Theorem 1: For any single-letter codes $\{f(i), g(i)\}$ with $X(i) = f(S(i))$ and $\hat{S}(i) = g(Y(i))$, we have the following Markov chain for any i:

$$S(i) \longrightarrow X(i) \longrightarrow Y(i) \longrightarrow \hat{S}(i).$$

It follows from the data processing inequality that

$$I(S(i); \hat{S}(i)) \leq I(X(i); Y(i)) \leq \frac{1}{2}\log(1 + |h(i)|^2 P(i)). \tag{6}$$

Combing (6) and Lemma 2, we obtain that

$$\begin{aligned} E\left(|S(i) - \hat{S}(i)|^2 \middle| h(i)\right) &\geq \sigma_S^2 \exp\left(-2I(S(i); \hat{S}(i))\right) \\ &\geq \frac{\sigma_S^2}{1 + |h(i)|^2 P(i)}. \end{aligned}$$

Therefore,

$$\begin{aligned} E(|S(i) - \hat{S}(i)|^2) &= E_{h(i)}\left\{E\left(|S(i) - \hat{S}(i)|^2 \middle| h(i)\right)\right\} \\ &\geq \sigma_S^2 E_{h(i)}\left\{\frac{1}{1 + |h(i)|^2 P(i)}\right\}. \end{aligned}$$

It is easy to see that the equality is obtained by linear coding (c.f. (5)). Therefore among all single-letter codes, linear coding is optimal. ■

We have shown the optimality of linear codes among all single-letter codes when the power allocation $\{P(i) : i \in \mathbb{Z}^+\}$ is given. Since only an average power constraint is imposed on the channel input symbols, we can potentially perform power allocation (along the time index i) to optimize the performance. We have the following theorem for the optimal power allocation.

Theorem 3: When there is no transmitter CSI, uniform power allocation is optimal. The achieved MSE

$$D_l(P; \text{no Tx CSI}) = \sigma_S^2 E_h \left\{ \frac{1}{1 + |h|^2 P} \right\}. \tag{7}$$

When there is transmitter CSI, i.e., the realization of $h(i)$ is known by the encoder, then the optimal power allocation is given by

$$P(i)^* = \frac{1}{|h(i)|}\left(\mu - \frac{1}{|h(i)|}\right)^+,$$

where μ is a common threshold for all $h(i)$, and is decided by the average power constraint P and the statistics of h (c.f. (15)). The corresponding achieved MSE is given in (16).

The proof of Theorem 3 is given in the Appendix.

B. Linear Coding of Finite Block Length

In this section we consider the linear coding with block length $n \geq 2$. The encoder is given by a $n \times n$ matrix F, and the decoder is the MMSE decoder (see Fig. 3). Let $\Omega_S \stackrel{\text{def}}{=} E(S^{(n)} {S^{(n)}}^T)$, and $H \stackrel{\text{def}}{=} \text{diag}(h^{(n)})$, then we have

$$X^{(n)} = F S^{(n)},$$

Fig. 3. A linear coding system with block length $n \geq 2$.

$$
\begin{aligned}
Y^{(n)} &= X^{(n)} + W^{(n)} = HFS^{(n)} + W^{(n)}, \\
\hat{S}^{(n)} &= \left(HF\Omega_S F^T H^T + I\right)^{-1} HF\Omega_S F^T H^T Y^{(n)}.
\end{aligned}
$$

The achieved MSE, in terms of the encoding matrix F and channel matrix H, can be expressed as:

$$
\begin{aligned}
D(F,H) &\stackrel{\text{def}}{=} \frac{1}{n}\text{tr}\left\{E\left((S^{(n)} - \hat{S}^{(n)})(S^{(n)} - \hat{S}^{(n)})^T\right)\right\} \\
&= \frac{1}{n}\text{tr}\left\{\left(HF\Omega_S F^T H^T + I\right)^{-1}\Omega_S\right\}. \qquad (8)
\end{aligned}
$$

The power constraint implies

$$
\text{tr}(F\Omega_S F^T) \leq nP.
$$

Thus, introducing $Q = FF^T$, and noticing $\Omega_S = \sigma_S^2 I$, we can solve the following problem to obtain the optimal Q^*, and get the optimal encoding matrix F^*:

$$
\begin{aligned}
\min \quad & E_H\left\{\text{tr}\left(HQH^T + \sigma_S^{-2}I\right)^{-1}\right\} \\
\text{s. t.} \quad & \text{tr}(Q) \leq \frac{nP}{\sigma_S^2}.
\end{aligned} \qquad (9)
$$

To solve (9), we quote the following two lemmas in matrix algebra [4] without proof.

Lemma 4: For any square matrix $R \succ 0$, it holds that $\text{tr}(R^{-1}) \geq \sum_{i=1}^{n} R_{ii}^{-1}$, and equality holds iff R is diagonal.

Lemma 5: For any square matrices A and B, it holds that $\text{tr}(I + AB)^{-1} = \text{tr}(I + BA)^{-1}$.

We have the following theorem regarding the performance of linear coding when block length n increases.

Theorem 6: For the source-channel coding of a Gaussian source transmitted through an AWGN fading channel, any linear coding with finite block length can be performed in single-letter form without performance loss.

Proof: To solve (9), we first apply Lemma 4 and obtain $\text{tr}\left(HQH^T + \sigma_S^{-2}I\right)^{-1} = \text{tr}\left(QH^TH + \sigma_S^{-2}I\right)^{-1}$ for any H. Then by Lemma 5, we obtain

$$
\text{tr}\left(QH^TH + \sigma_S^{-2}I\right)^{-1} \geq \sum_{i=1}^{n} \frac{1}{Q_{ii}|h(i)|^2 + \sigma_S^{-2}}, \qquad (10)
$$

where equality holds iff Q is diagonal. Therefore, the optimal solution gives diagonal $Q^* = FF^T$. Thus, any $F^* = \sqrt{Q^*}U$ where U is unitary is an optimal solution. Specifically, if we take $U = I$, we can obtain a diagonal F^*. So any linear coding can be achieved in a single-letter form without performance loss. ■

C. Comparison of the Performance of Linear Coding with the Shannon's Bound

In this section we compare the performance of linear coding with the Shannon's bound, which is the theoretical benchmark. According to the separation theorem, the Shannon's bound can be obtained by combining the rate-distortion function and channel capacity. In the rest of this paper we consider the case when there is receiver CSI only. The analysis for the case with transmitter CSI can be done analogously.

The rate-distortion function of a memoryless Gaussian source with variance σ_S^2 is

$$
R(D) = \frac{1}{2}\log^+ \frac{\sigma_S^2}{D}. \qquad (11)
$$

Combining it with the channel capacity (when there is receiver CSI only and the average power constraint is P)

$$
C(P) = E_h\left\{\frac{1}{2}\log(1 + |h|^2 P)\right\},
$$

we obtain that the best achievable distortion in terms of P is

$$
D^*(P; \text{no Tx CSI}) = \sigma_S^2 \exp\left(E_h\left\{\log\frac{1}{1 + |h|^2 P}\right\}\right). \qquad (12)
$$

Recalling (7), it is easy to see that $D_l(P; \text{no Tx CSI}) \geq D^*(P; \text{no Tx CSI})$ from concavity of the log-function. The equality holds iff

$$
E_h\left\{\log\frac{1}{1 + |h|^2 P}\right\} = \log\left(E_h\left\{\frac{1}{1 + |h|^2 P}\right\}\right),
$$

which is equivalent to claiming that $1/(1 + |h|^2 P)$, or $|h|^2$ is a constant.

Therefore, we see that linear coding is optimal only if $h \equiv \pm c$ when h is real, or h is distributed on a circle when h is complex. For all other cases, linear coding is suboptimal, and can not achieve Shannon's bound. In what follows, we bound the performance gap in terms of P and the statistic of h.

Theorem 7: When there is no transmitter CSI, the MSE performance of linear coding can be bounded away from the theoretically best achievable MSE (i.e., the Shannon's bound) as follows:

$$
0 \leq \gamma(P, h) \leq P\sqrt{\text{Var}(|h|^2)},
$$

where

$$
\gamma(P, h) \stackrel{\text{def}}{=} \frac{D_l(P) - D^*(P)}{D^*(P)},
$$

and for simplicity, $D_l(P) := D_l(P; \text{no Tx CSI})$; $D^*(P) := D^*(P; \text{no Tx CSI})$.

Proof: Introducing $h_0 = \sqrt{E(|h|^2)}$, then from (12) we obtain

$$
\begin{aligned}
D^*(P) &= \sigma_S^2 \exp\left(E_h\left\{\log\frac{1}{1 + |h|^2 P}\right\}\right) \\
&\geq \sigma_S^2 \frac{1}{1 + h_0^2 P} \stackrel{\text{def}}{=} D_0(P).
\end{aligned}
$$

Notice that the right term in the above formula is the best achievable performance when channel is AWGN with the same average path gain, thus it is a natural performance bound for the fading case. Recalling (7), we obtain

$$
\begin{aligned}
0 \leq \gamma(h, P) &= \frac{D_l(P) - D^*(P)}{D^*(P)} \leq \frac{D_l(P) - D_0(P)}{D_0(P)} \\
&= E_h\left\{\frac{(h_0^2 - |h|^2)P}{1 + |h|^2 P}\right\} \leq E_h\left\{||h|^2 - h_0^2|P\right\} \\
&= P\sqrt{\text{Var}(|h|^2)}.
\end{aligned}
$$

The proof is complete. ■

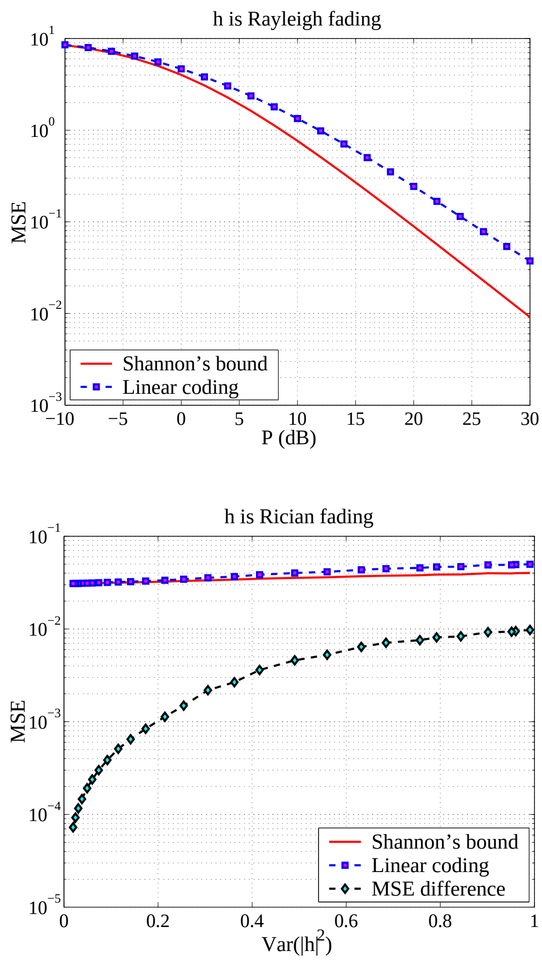

Fig. 4. MSE performance as P or Var$(|h|^2)$ increases. In the top plot, we assume h is Rayleigh fading with Var$(|h|^2) = 1$, but the transmit power P increases. In the bottom plot, we assume P=15dB, and h is Rician fading with increasing variance Var$(|h|^2)$ (by taking different Rician factors).

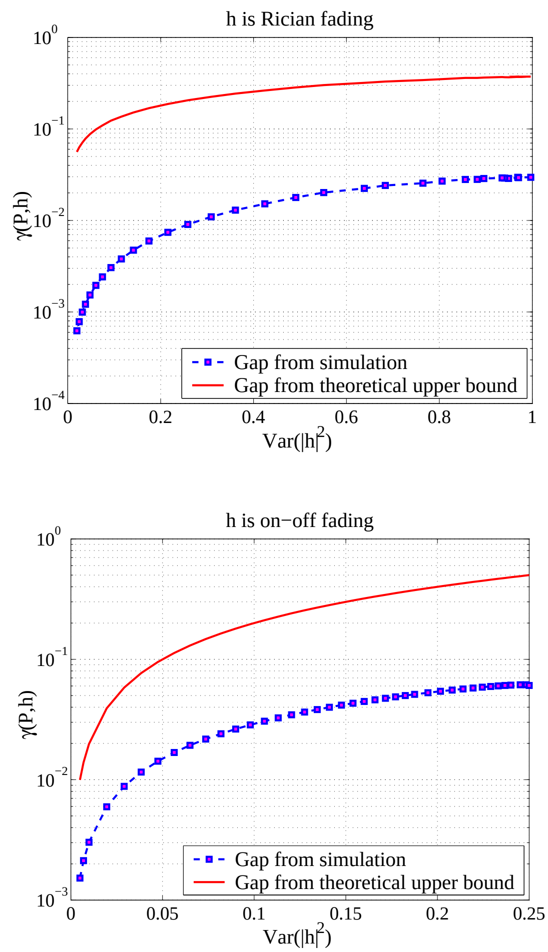

Fig. 5. Performance gap of linear coding compared to the Shannon's bound. The red curves are the universal bound of the gap $\gamma(P, h)$ given in Theorem 7 for all possible distributions. In addition, in the top plot, we assume $P = 0$dB, and h is Rician fading with increasing Var$(|h|^2)$ (by taking different Rician factors); in the bottom plot, we assume $P = 0$dB, and h is on-off fading with increasing Var$(|h|^2)$.

III. Numerical Results

From Theorem 7 we can see that if the performance of linear coding is close to Shannon's bound the product $P\sqrt{\text{Var}(|h|^2)}$ is small. Thus the linear coding is well suited for the source-channel communication applications in an energy-constrained network (which includes the wireless sensor network as an example) [5], where the stringent power constraint usually leads to small transmit power budget P. In what follows, we show some plots to demonstrate the performance gap of linear coding compared to the Shannon's bound for different fading distributions and power budget.

The plot of the MSE performance in the case of Rayleigh fading with increasing power (when the fading distribution h is fixed), or Rician fading with increasing variance (when power P is fixed) are plotted in Fig. 4. We can see that when either the transmit power or the variance of fading gain is relatively small, the MSE performance gap is negligible. In Fig. 5, we numerically examine the behavior of the gap coefficient $\gamma(P, h)$ for different fading distributions. We conclude that the gap coefficients $\gamma(P, h)$ from both theory and experiments are small ($\ll 1$).

IV. Conclusion and Future Work

We have studied the performance linear coding of a discrete memoryless Gaussian source transmitted through a discrete memoryless fading channel. We show that linear coding is optimal among all single-letter codes, but the performance of linear coding can not be improved by increasing the block length. Thus, in general, the linear coding can not achieve the Shannon's bound unless the magnitude of the fading gain is a constant. We bound the performance gap of linear coding from the optimal coding in terms of the variance of the fading gain and transmit power budget. Both theoretical analysis and simulation show that the gap is negligible if either transmit power is or the variance of fading gain is relatively small.

As future directions, we plan to investigate simple joint source-channel codes that can approach or obtain the Shannon's bound for transmission of Gaussian sources through the fading channels. We have shown in this paper that linear coding is suboptimal, thus, other simple but optimal coding schemes need to be proposed. We have also shown that linear coding is optimal among all single-letter codes. In general, it is worthwhile to find the best coding schemes among source-channel codes with fixed block length. We are interested in characterizing necessary/sufficient conditions for such coding schemes for general source-channel pairs.

Appendix: Optimal Power Loading

We first show that uniform power allocation is optimal when there is no transmitter CSI. Suppose P is the average power budget. The mean square distortion averaged over time is

$$\begin{aligned} E(|S-\hat{S}|^2) &\overset{\text{def}}{=} \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}E(|S(i)-\hat{S}(i)|^2 \\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}E_{h(i)}\left\{\frac{\sigma_S^2}{1+|h(i)|^2P(i)}\right\} \\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}E_{h}\left\{\frac{\sigma_S^2}{1+|h|^2P(i)}\right\} \\ &\geq \sigma_S^2 E_h\left\{\frac{1}{1+|h|^2P}\right\}, \end{aligned}$$

where the last step is due to the convexity of the function $E_h\left\{\frac{\sigma_S^2}{1+|h|^2P(i)}\right\}$ in terms of $P(i)$, and the equality holds iff $P(i)=P$ for all i. Thus when there is no transmitter CSI, uniform power allocation is optimal, and the corresponding achieved MSE

$$D_l(P;\text{no Tx CSI}) = \sigma_S^2 E_h\left\{\frac{1}{1+|h|^2P}\right\}. \quad (13)$$

If there is CSI at the transmitter, the optimal power loading (along the time i) is not uniform anymore. Instead, the power can be loaded according to the CSI to achieve better distortion performance. Suppose when the channel state is h, the corresponding power loading is $P(h)$. Then the achieved average distortion is

$$D = E_h\left\{\frac{\sigma_s^2}{1+|h|^2P_2(h)}\right\}.$$

The optimal power loading $P^*(h)$ can be solved from the following problem.

$$\begin{aligned} \min \quad & E_h\left\{\frac{\sigma_s^2}{1+|h|^2P_2(h)}\right\} \\ \text{s.t.} \quad & E_h\{P(h)\}=P, \quad P(h)\geq 0. \end{aligned}$$

Due to the concavity of the function $\frac{\sigma_s^2}{1+|h|^2P_2(h)}$ in terms of $P(h)$, we obtain that for all h that have the same magnitude, the optimal power loading $P^*(h)$ should also be the same, i.e., $P^*(h)$ is only a function of $|h|$. We thus can do the power loading according to $|h|$ without loss of optimality. If $|h|$ has finite states, assuming $P(|h|=g_i)=f_i,\ i=1,2,\ldots,L$, we have

$$\begin{aligned} \min \quad & \sum_{i=1}^{L}\frac{\sigma_s^2}{1+g_i^2P_i}f_i \\ \text{s.t.} \quad & \sum_{i=1}^{L}P_if_i=P, \quad P_i\geq 0. \end{aligned}$$

The Lagrangian is

$$G(P,\mu,\lambda)=\sum_{i=1}^{L}\frac{\sigma_s^2}{1+g_i^2P_i}f_i+\mu\left(\sum_{i=1}^{L}P_if_i-P\right)-\sum_{i=1}^{L}\lambda_iP_i.$$

We obtain the following Karush-Kuhn-Tucker (KKT) conditions [6]:

$$\begin{aligned} \frac{\partial G}{\partial P_i} = -\frac{\sigma_s^2g_i^2}{(1+g_i^2P_i)^2}f_i+\mu f_i-\lambda_i=0, \\ P_i\lambda_i=0, \quad i=1,2,\ldots L, \\ \sum_{i=1}^{L}P_if_i=P. \end{aligned}$$

For all those $P_i\neq 0$, we have $\lambda_i=0$, and $\frac{\sigma_s^2g_i^2}{(1+g_i^2P_i)^2}=\mu$. Therefore,

$$P_i^* = \frac{1}{g_i}\left(\mu-\frac{1}{g_i}\right)^+,$$

where μ is a common threshold for all states, and can be solved from

$$\sum_{i=1}^{L}P_i^*f_i = \sum_{i=1}^{L}\frac{1}{g_i}\left(\mu-\frac{1}{g_i}\right)^+f_i = P.$$

If $|h|$ has a non-discrete pdf, then by discretizing the pdf and taking the limit, we can obtain

$$P^*(|h|) = \frac{1}{|h|}\left(\mu-\frac{1}{|h|}\right)^+, \quad (14)$$

where μ is solved from

$$\begin{aligned} \int_0^\infty P^*(|h|)f(|h|)d|h| &= \int_{\frac{1}{\mu}}^{\infty}\frac{1}{|h|}\left(\mu-\frac{1}{|h|}\right)f(|h|)d|h| \\ &= P. \end{aligned} \quad (15)$$

With this power allocation strategy, the achieved average MSE

$$D_l(P;\text{with Tx CSI}) = \sigma_S^2E_h\left\{\frac{1}{1+|h|^2P^*(|h|)}\right\}. \quad (16)$$

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379-423, 623-656, 1948.

[2] T. J. Goblick, "Theoretical limitations on the transmission of data from analog sources," *IEEE Transactions on Information Theory*, vol. 11, pp. 558-567, Oct. 1965.

[3] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: lossy source channel communication revisited", *IEEE Transactions on Information Theory*, Vol. 49, pp. 1147-1158, May 2003.

[4] K. M. Abadir, *Matrix algebra*, Cambridge University Press, New York, 2005.

[5] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Communication Magazine*, pp. 8-27, Aug. 2002.

[6] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge Univ. Press, Cambridge, U.K., 2003.