Channel Dependent Adaptive Modulation and Coding *without* Channel State Information at the Transmitter

Bradford D. Boyle, John MacLaren Walsh, and Steven Weber

January 7, 2014

Abstract

In this paper, we consider the problem of communicating over an Additive White Gaussian Noise (AWGN) channel with an unknown noise power *without* feedback. For the case where the noise power is drawn from a distribution with K distinct values, we model the channel as a K-user broadcast channel, with a user for each of the K possible noise powers. We propose two metrics for measuring the loss of using a broadcast code compared to an omniscient transmitter. Using these metrics, we solve for the optimal power allocation in a superposition code. For the case of $K = 2$, we prove several interesting properties of the optimal power allocation and the minimum loss.

1 Introduction

Traditional Adaptive Modulation and Coding (AMC) proceeds by first defining a finite collection of codes and modulation schemes associated with different information rates r_k measured in bits per channel use. The index $k \in \{1, \ldots, K\}$ indicating which scheme to use is called the Modulation and Coding Scheme (MCS) index. The receiver measures the channel quality using reference or training signals, or pilots, and determines the information rate among this finite set corresponding to a modulation and coding scheme achieving a given target probability of error. The associated index k, or some quantization of it, is then fed back to the transmitter under the label Channel Quality Index (CQI). The transmitter then takes into consideration factors such as the amount of data waiting to be sent to the various receivers associated with it and their necessary quality of service, then selects the modulation and coding scheme to use when transmitting to them.

The key capability that such a scheme allows is the adaptability of the amount of information flowing across the link to the fluctuations in channel state. Such a scheme relies on explicit feedback from the receiver to the transmitter and this feedback can account for a non-trivial amount of the channel capacity. An omniscient transmitter would know which state the channel is in and could select an appropriate code and modulation scheme without explicit feedback from the receiver.

The problem considered here fits into the larger class of problems dealing with coding for channels with state. The writing on dirty paper result, which is a specialization of the Gelfand-Pinsker theorem to AWGN channels with state, shows that if the state information is available noncausally at the transmitter, then the effect of the state can be canceled [1]. Our model assumes that the state information is not available (causally or noncausally) at the transmitter.

In this paper, we propose a broadcast channel model for a transmitter to effect channel dependent AMC without Channel State Information (CSI), as suggested by Cover's approach to compound channels [2] (§2) along with two related metrics for measuring the loss between our model and an omniscient transmitter (§3). We provide a characterization of a transmitter that minimizes these losses (§3.1). For the special case of $K = 2$, we prove several non-intuitive properties of the optimal solution (§3.2). Finally, we provide plots of the optimal solution for both the general case (§4.1) and the special case $K = 2$ (§4.2).

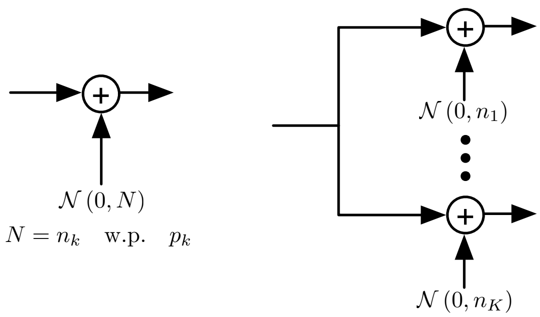


Figure 1: Trying to guarantee messages with information rates that are a function of an unknown channel SNR can be understood as a broadcast channel model. The receivers on the right do not in reality simultaneously exist, however, the code must ensure a certain message is successfully received under each of the K possible channel states, and this can be thought of as reliably transmitting those messages to K simultaneous receivers.

2 Problem Model

Consider a point-to-point AWGN channel between a basestation and mobile handset. The unknown noise power of the channel is modeled as a random variable N with some probability mass function p_k defined on a set of K possible values $n_1 \leq n_2 \leq \cdots \leq n_K$. Under the assumption of a fixed transmit power P at the basestation, we can consider the random variable $\Gamma = {}^{P}\!/_{N}$ with probability mass function defined on $\gamma_1 = {}^{P}\!/_{n_1} \geq \cdots \gamma_K = {}^{P}\!/_{n_K}$. This is depicted in the left-hand side of Fig. 1. We are interested in a variable-to-fixed coding strategy [3], wherein a fixed number of symbols are transmitted across the channel and the decoder recovers a variable number of information bits, depending on the state of the channel. By utilizing a fixed blocklength, we ensure the the receiver decodes some information in a fixed amount of time. This is in contrast to a fixed-to-variable coding strategy [3] (e.g., rateless fountain codes) where a variable number of symbols (and hence variable delay) are transmitted in order for the receiver to decode a fixed number of information bits. We note that in an AMC scheme with feedback, the collection of codes do not need to be all of the same length.

For channels with state, a lower bound for the variable-to-fixed channel capacity is known in terms of the capacity region for a broadcast channel with degraded message sets [3]. For the AWGN channel model considered here, this bound is tight. The capacity region of the K-user broadcast channel with independent messages is [4, 5]

$$R_k \leq \mathsf{C}\left(\frac{\alpha_k}{\sum_{i<k} \alpha_i + \gamma_k^{-1}}\right) \quad k = 1, \ldots, K \tag{1}$$

where $\gamma_1 \geq \cdots \geq \gamma_K$ are the SNRs ($\gamma_k = {}^{P}\!/_{n_k}$) associated with the selected adaptive modulation and coding rates, $\mathsf{C}(\gamma) = {}^{1}\!/_{2} \log_2(1+\gamma)$, and $\boldsymbol{\alpha} \geq 0$ such that $\sum_k \alpha_k = 1$. By consider the superposition coding achievability proof of (1), we can understand the α_k's as being the fraction of the transmitter's total power P that is used for encoding the message for the k-th receiver. The message containing R_k bits is reliably decoded by all receivers $1, \ldots, k$; that is, the k-th receiver reliably receives $\sum_{i=k}^{K} R_k$ bits and in particular, all receivers reliably receive R_K bits and this can be interpreted as representing common information for all receivers.

3 Gap To Omniscience

We must select a point on the Pareto frontier of (1) by selecting $\alpha_k \geq 0$, $k = 1, \ldots, K$ such that $\sum_k \alpha_k = 1$, and we must additionally measure the capacity loss from not knowing the SNR and attempting to transmit a single message (with a single, fixed block length) that is capable of being decoded at several different SNRs as compared to an omniscient (i.e., perfect CSI estimation and instantaneous feedback) model. It is natural, then, to combine these two issues together by selecting the point in the rate region which minimizes an appropriately chosen loss metric.

If we have a prior distribution indicating the probability that the channel is in each of the K states, we could seek to minimize expected capacity loss

$$J_{\mathbb{E},|\cdot|}(\boldsymbol{\alpha}) = \sum_{k=1}^{K} p_k \left(\mathsf{C}(\gamma_k) - \sum_{i=k}^{K} \mathsf{C}\left(\frac{\alpha_i}{\sum_{j<i} \alpha_j + \gamma_i^{-1}} \right) \right) \tag{2}$$

where p_k is the probability of the channel being in state γ_k. The right-hand side of (2) is the difference between the variable-to-fixed channel capacity if the AWGN channel state were known to the transmitter and when the state is unknown [3].

3.1 General Solution

Once we have selected an appropriate loss metric, we find $\alpha_1, \ldots, \alpha_K$ by solving the following optimization

$$\begin{aligned} \underset{\boldsymbol{\alpha}}{\text{minimize}} \quad & J(\boldsymbol{\alpha}) \\ \text{subject to} \quad & \alpha_k \geq 0 \quad k = 1, \ldots, K \\ & \sum_{k=1}^{K} \alpha_k = 1. \end{aligned} \tag{3}$$

We can alternatively express the expected capacity loss as

$$J_{\mathbb{E},|\cdot|}(\boldsymbol{c}) = \mathbb{E}\left[\mathsf{C}(\Gamma)\right] - \sum_{k=1}^{K} f_k \mathsf{C}\left(\frac{c_k - c_{k-1}}{c_{k-1} + \gamma_k^{-1}} \right) \tag{4}$$

where

$$f_k \triangleq \sum_{i=1}^{k} p_i = \mathbb{P}\left[\Gamma \geq \gamma_k\right] \tag{5}$$

and

$$c_k = \sum_{i=1}^{k} \alpha_i, \quad \alpha_k = c_k - c_{k-1} \tag{6}$$

We observe that only the negative sum in (4) depends on $\boldsymbol{c}$, and hence we can minimize (4) by selecting the $\boldsymbol{c}$ that solves

$$\begin{aligned} \underset{\boldsymbol{c}}{\text{maximize}} \quad & \sum_{k=1}^{K} f_k \mathsf{C}\left(\frac{c_k - c_{k-1}}{c_{k-1} + \gamma_k^{-1}} \right) \\ \text{subject to} \quad & c_k \geq c_{k-1} \quad k = 1, \ldots, K. \end{aligned} \tag{7}$$

We note that the objective function in both (3) and (7) is not convex; as a simple counter example, let $K = 2$, $\gamma_1 = 2$, $\gamma_2 = 1$, and $p_1 = 1/2$. The second derivative w.r.t. α using these parameters is not positive semi-definite (see Appendix B).

For notational brevity in the sequel, we will define the following quantity

$$\hat{c}_i \triangleq \frac{f_i \gamma_i - f_{i+1} \gamma_{i+1}}{\gamma_i \gamma_{i+1} (f_{i+1} - f_i)}. \tag{8}$$

Implicit in the above definition is the assumption that $p_i > 0$ for all i.

Theorem 1. *If the f_i's and γ_i's are such that (8) is an* increasing *sequence in i, the solution to (7) is is given as*

$$c_i^* = [\hat{c}_i]_0^1 \tag{9}$$

where $[\cdot]_0^1 = \max\{\min\{\cdot, 1\}, 0\}$.

Proof. For all feasible $\boldsymbol{c}$, there are at most $K-1$ active constraints; otherwise, we would have the contradiction $0 = 1$. The structure of the inequality constraints are such that we can pick any $K-1$ of the gradient vectors at a feasible point and have a set of $K-1$ linearly independent vectors. Therefore every feasible $\boldsymbol{c}$ is *regular*.

Forming the Lagrangian for this optimization we have

$$L(\boldsymbol{c}, \boldsymbol{\mu}) = -\sum_{i=1}^{K} f_i \mathsf{C}\left(\frac{c_i - c_{i-1}}{c_{i-1} + \gamma_i^{-1}}\right) + \sum_{i=1}^{K} \mu_i (c_{i-1} - c_i). \tag{10}$$

The first order derivative of the Lagrangian is

$$\begin{aligned}\frac{\partial}{\partial c_i} L(\boldsymbol{c}, \boldsymbol{\mu}) &= -f_i \frac{\partial}{\partial c_i} \mathsf{C}\left(\frac{c_i - c_{i-1}}{c_{i-1} + \gamma_i^{-1}}\right) - f_{i+1} \frac{\partial}{\partial c_i} \mathsf{C}\left(\frac{c_{i+1} - c_i}{c_i + \gamma_{i+1}^{-1}}\right) - \mu_i + \mu_{i+1} \\ &= \frac{-f_i}{2(c_i + \gamma_i^{-1}) \ln 2} + \frac{-f_{i+1}}{2(c_i + \gamma_{i+1}^{-1}) \ln 2} - \mu_i + \mu_{i+1}\end{aligned}$$

For a local minimum $\boldsymbol{c}^*$, first order necessary condition yields

$$c_i^* - \hat{c}_i = 2 \ln 2 (\mu_i - \mu_{i+1}) \frac{(c_i^* + \gamma_i^{-1})(c_i^* + \gamma_{i+1}^{-1})}{f_{i+1} - f_i}. \tag{11}$$

Two key observations are:

1. If $p_i > 0$, then
$$\frac{(c_i^* + \gamma_i^{-1})(c_i^* + \gamma_{i+1}^{-1})}{f_{i+1} - f_i} > 0$$
2. If $\mu_i^* > 0$, then by complimentary slackness, we must have $c_i^* = c_{i-1}^*$.

We first consider the case $c_i^* > \hat{c}_i$. The left-hand side of (11) is positive and we must have that $\mu_i^* - \mu_{i+1}^* > 0$ which implies $\mu_i^* > 0$ and that $c_{i-1}^* = c_i^*$. We have assumed that $\hat{c}_{i-1} < \hat{c}_i$ for all $i = 1, \ldots, K$. We have shown that if $c_i^* > \hat{c}_i$ then $c_{i-1}^* = c_i^*$. Combining these three facts would give us that $c_{i-1}^* = c_i^* > \hat{c}_i > \hat{c}_{i-1}$ and in particular $c_{i-1}^* > \hat{c}_{i-1}$. Repeating, we come to the conclusion if $c_i^* > \hat{c}_i$ then $c_i^* = c_{i-1}^* = \cdots = c_0^* = 0$.

Next, we consider the case $c_i^* < \hat{c}_i$. The left-hand side of (11) is negative and we must have that $\mu_i^* - \mu_{i+1}^* < 0$ which implies that $\mu_{i+1}^* > 0$ and that $c_{i+1}^* = c_i^*$. We have $c_{i+1}^* = c_i^* < \hat{c}_i < \hat{c}_{i+1}$ and in particular $c_{i+1}^* < \hat{c}_{i+1}$. Repeating, we come to the conclusion if $c_i^* < \hat{c}_i$ then $c_i^* = c_{i+1}^* = \cdots = c_K^* = 1$. ☐

Remark. *The requirement in Theorem 1 that the f_k's and γ_k's are such that (8) is an increasing sequence in k holds true in some natural cases of interest. If we consider uniformly distributed ($p_k = 1/K$), exponentially spaced SNRs ($\gamma_k = \delta^{K-k}\gamma_K$ with $\delta > 1$), then (8) becomes*

$$c_k^* = \frac{k(\delta - 1) - 1}{\delta^{K-k}\gamma_K}$$

which is increasing in k.

A closely related metric is the fractional expected capacity loss

$$J_{\%,\mathbb{E}}(\boldsymbol{\alpha}) = \frac{J_{\mathbb{E},|\cdot|}(\boldsymbol{\alpha})}{\sum_{k=1}^{K} p_k \mathsf{C}(\gamma_k)} \tag{12}$$

which is the expected capacity loss multiplied by a scalar that does not depend on $\boldsymbol{\alpha}$ and so has the same optimal $\boldsymbol{\alpha}$.

3.2 Channels with Two States

We now consider in detail the special case where the channel is in one of two possible states ($K = 2$). We will write $\gamma_1 = \delta\gamma_2$ for some $\delta \geq 1$ and consider how the selected point on the Pareto frontier of (1) and the resulting minimum loss behave as a function of δ for the fixed parameters γ_2 and p_1 (i.e., the probability of being in the better state). For the case of two users, a point on the Pareto frontier is parameterized by $\boldsymbol{\alpha} = (\alpha_1, \alpha_2) = (\alpha, 1 - \alpha)$.

Corollary 1. *For $K = 2$ and $\gamma_1 = \delta\gamma_2$, $\delta \geq 1$, the α that minimizes (2) and (12) is given as*

$$\alpha^* = \left[\frac{p_1\delta - 1}{\delta\gamma_2(1 - p_1)}\right]_0^1 \tag{13}$$

Proposition 1. *For $K = 2$ and $\gamma_1 = \delta\gamma_2$, $\delta \geq 1$, (i) $\alpha^* = 0$ for $1 \leq \delta \leq {}^1/_{p_1}$; (ii) If $p_1 < 1$ then α^* is monotone increasing in δ, and; (iii) If $p_1 \leq {}^{\gamma_2}/_{\gamma_2+1}$, then*

$$\lim_{\delta\to\infty} \alpha^* = \frac{p_1}{\gamma_2(1 - p_1)}. \tag{14}$$

If $p_1 > {}^{\gamma_2}/_{\gamma_2+1}$, then

$$\alpha^*|_{\delta = \frac{1}{p_1 - \gamma_2(1-p_1)}} = 1. \tag{15}$$

Proof. (i) Follows immediately from (13).

(ii) Taking the first derivate we have

$$\frac{\partial}{\partial\delta} \frac{p_1\delta - 1}{\delta\gamma_2(1 - p_1)} = \frac{\gamma_2(1 - p_1)}{(\delta\gamma_2(1 - p_1))^2} > 0$$

(iii) If $p_1 \leq {}^{\gamma_2}/_{\gamma_2+1}$, then there exists no positive δ for which $\alpha^* = 1$ and we have

$$\lim_{\delta\to\infty} \alpha^* = \lim_{\delta\to\infty} \frac{p_1\delta - 1}{\delta\gamma_2(1 - p_1)}.$$

If $p_1 > {}^{\gamma_2}/_{\gamma_2+1}$ then simple algebra shows that when $\delta = {}^1/_{p_1 - \gamma_2(1-p_1)}$, $\alpha^* = 1$.

□

The first part of Proposition 1 says the the better SNR (γ_1) needs to be larger than worse SNR (γ_2) by a minimum amount before the transmitter should allocate any of its power to encoding a message for this better channel. This threshold is inversely proportional to the probability of the high SNR — the less likely $\Gamma = \gamma_1$, the larger γ_1 needs to be before the transmitter will allocate power for this channel state. The last part of Proposition 1 says that unless the probability of being in the better state is sufficiently high, the transmitter should not allocate all of its power for this channel despite how much better the channel may be. Conversely, if the probability of $\Gamma = \gamma_1$ is high enough then there exists a threshold on γ_1 above which the transmitter should allocate all of its power to that state.

Proposition 2. *For $K = 2$ and $\gamma_1 = \delta\gamma_2$ with $\delta \geq 1$, the expected capacity loss $J_{\mathbb{E},|\cdot|}(\alpha^*)$ is: (i)*

$$J_{\mathbb{E},|\cdot|}(\alpha^*)\big|_{\delta=1} = 0; \tag{16}$$

(ii) increasing in δ, and; (iii) bounded as $\delta \to \infty$. In particular, if $p_1 \leq {}^{\gamma_2}/_{\gamma_2+1}$,

$$\lim_{\delta\to\infty} J_{\mathbb{E},|\cdot|}(\alpha^*) = \frac{p_1}{2}\log_2\left(\frac{\gamma_2(1 - p_1)}{(1 + \gamma_2)p_1}\right) - \frac{1}{2}\log_2(1 - p_1) \tag{17}$$

and if $p_1 > {}^{\gamma_2}/_{\gamma_2+1}$

$$\lim_{\delta\to\infty} J_{\mathbb{E},|\cdot|}(\alpha^*) = \frac{1 - p_1}{2}\log_2(1 + \gamma_2) \tag{18}$$

Proof. (i) Follows from Corollary 1 by setting $\delta = 1$.

(ii) For $1 \le \delta \le 1/p_1$,

$$J_{\mathbb{E},|\cdot|}(\alpha^*) = \frac{p_1}{2}\log_2\left(\frac{1+\delta\gamma_2}{1+\gamma_2}\right)$$

which is increasing in δ.
If $p_1 \le \gamma_2/\gamma_2+1$ and $\delta \ge 1/p_1$,

$$\frac{\partial}{\partial\delta}J_{\mathbb{E},|\cdot|}(\alpha^*) = \frac{1-\delta(p_1-\gamma_2(1-p_1))}{2(\delta-1)\delta(1+\delta\gamma_2)} \ge 0$$

for all $\delta \ge 1/p_1$.
If $p_1 > \gamma_2/\gamma_2+1$ and

$$\frac{1}{p_1} \le \delta \le \frac{1}{p_1-\gamma_2(1-p_1)}$$

then

$$\frac{\partial}{\partial\delta}J_{\mathbb{E},|\cdot|}(\alpha^*) = \frac{1-\delta(p_1-\gamma_2(1-p_1))}{2(\delta-1)\delta(1+\delta\gamma_2)}$$

which is non-negative over the assumed range for δ.
If $p_1 > \gamma_2/\gamma_2+1$, then for

$$\delta \ge \frac{1}{p_1-\gamma_2(1-p_1)}$$

we have

$$J_{\mathbb{E},|\cdot|}(\alpha^*) = \frac{1-p_1}{2}\log_2(1+\gamma_2)$$

(iii) If $p_1 \le \gamma_2/\gamma_2+1$ then for $\delta \ge 1/p_1$

$$J_{\mathbb{E},|\cdot|}(\alpha^*) = \frac{p_1}{2}\log_2\left(\frac{(1+\delta\gamma_2)(1-p_1)}{p_1(\delta-1)(1+\gamma_2)}\right) + \frac{1}{2}\log_2\left(\frac{(1+\gamma_2)(\delta-1)}{\delta(1-p_1)(1+\gamma_2)}\right)$$

which follows from (13), Prop. 1, and some algebraic simplification. If $p_1 > \gamma_2/\gamma_2+1$ and $\delta > 1/p_1-\gamma_2(1-p_1)$ then $\alpha^* = 1$ and

$$J_{\mathbb{E},|\cdot|}(\alpha^*) = \frac{1-p_1}{2}\log_2(1+\gamma_2)$$

which follows from Prop. 1. Taking the limits of both expressions gives the results of the proposition. □

Proposition 3. *For $K = 2$ and $\gamma_1 = \delta\gamma_2$ with $\delta \ge 1$, the expected capacity $\mathbb{E}[\mathsf{C}(\Gamma)]$ is positive, monotonically increasing in δ, and approaches ∞ as $\delta \to \infty$.*

Proof. Immediate from the properties of $\log_2(x)$. □

Proposition 4. *For $K = 2$ and $\gamma_1 = \delta\gamma_2$ with $\delta \ge 1$, the fractional expected capacity loss $J_{\%,\mathbb{E}}(\boldsymbol{\alpha})$ is: (i) increasing in δ for $1 \le \delta \le 1/p_1$; (ii) non-monotonic in δ, and; (iii)*

$$\lim_{\delta\to\infty} J_{\%,\mathbb{E}}(\boldsymbol{\alpha}) = 0 \tag{19}$$

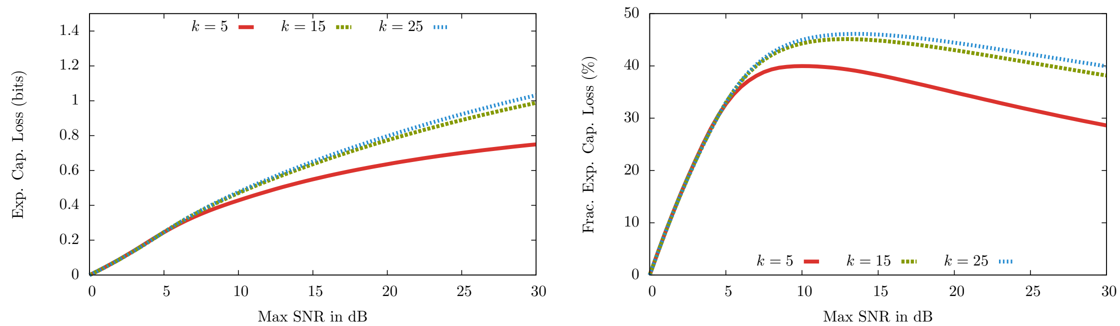
Proof. (i) For $1 \le \delta \le 1/p_1$ we have

$$\frac{\partial}{\partial\delta}J_{\%,\mathbb{E}}(\boldsymbol{\alpha}) = \frac{p_1\gamma_2\mathsf{C}(\gamma_2)}{\log(2)(1+\delta\gamma_2)\left(\mathbb{E}[\mathsf{C}(\Gamma)]\right)^2} > 0. \tag{20}$$

(ii) Follows from Propositions 2(ii), 2(iii), 3, & 4(i).

(iii) Follows from Propositions 2(iii) & 3.

□



(a) $J_{\mathbb{E},|\cdot|}\left(\boldsymbol{\alpha}^*\right)$ (b) $J_{\%,\mathbb{E}}\left(\boldsymbol{\alpha}^*\right)$

Figure 2: Minimum expected capacity loss (a) and minimum fractional expected capacity loss (b) as a function of maximum SNR (γ_1) and the number (K) of AMC levels. The minimum SNR was $\gamma_K = 0$ dB.

4 Results

In this section, we present plots of the minimum expected capacity loss and minimum fractional expected capacity loss for general K and specifically for $K = 2$.

4.1 K-State Channel Results

Fig. 2a shows the minimum expected capacity loss and Fig. 2b shows the minimum fractional expected capacity loss plotted as functions of γ_1 in dB for different values of K. In both plots, $\gamma_K = 0$ dB and $\gamma_2, \ldots, \gamma_{K-1}$ are evenly spaced (in dB) between γ_1 and γ_K. The assumed prior distribution was the discrete uniform distribution (i.e., $p_k = {}^1/_K$). From Fig. 2a, we observe that the minimum expected capacity loss is increasing in γ_1 and for small values of γ_1 (i.e., $\gamma_1 \leq 5$ dB), the minimum expected capacity loss looks linear in γ_1 and independent of K. From Fig. 2b, we observe that, like Fig. 2a, for $\gamma_1 \leq 5$ dB, the minimum fractional expected capacity loss looks linear in γ_1 and independent of K. Unlike Fig. 2a, the minimum fractional expected capacity loss is *not* monotonic in γ_1.

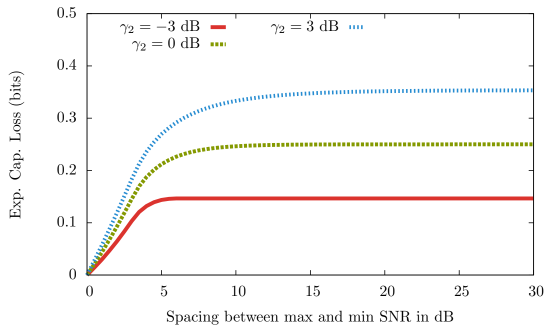
4.2 Two-State Channel Results

For all plots in this section, $p_1 = {}^1/_2$ and show results for: (i) $\gamma_2 = -3$ dB, shown in solid red line; (ii) $\gamma_2 = 0$ dB, shown in dashed blue line, and; (iii) $\gamma_2 = 3$ dB, shown in dotted green line. . With this value for p_1 and the given values of γ_2, we have the following three conditions: (i) $p_1 > {}^{\gamma_2}/_{\gamma_2+1}$; (ii) $p_1 = {}^{\gamma_2}/_{\gamma_2+1}$, and; (iii) $p_1 < {}^{\gamma_2}/_{\gamma_2+1}$.

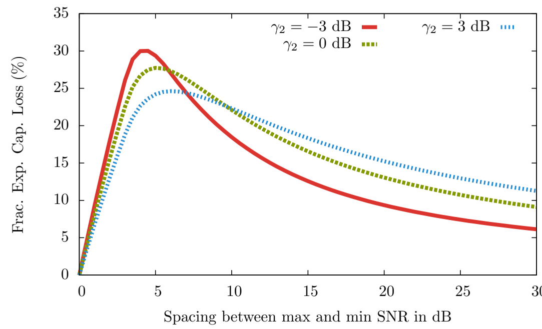
Fig. 3c shows the value of α that minimizes both the expected capacity loss and fractional expected capacity loss as a function of δ for the previously mentioned values of γ_2. We observe from Fig. 3c, that α^* is identically 0 for $\delta \leq 3$ dB (Prop. 1(i)) and is monotonically increasing in δ (Prop. 1(ii)). Additionally, we see that α^*approaches the limits (i) 1 at $\delta = 6$ dB for $\gamma_2 = -3$ dB, (ii) 1 for $\gamma_2 = 0$ dB, and (iii) 0.5 for $\gamma_2 = 3$ dB (Prop. 1(iii)).

Fig. 3a shows the expected capacity loss $J_{\mathbb{E},|\cdot|}\left(\alpha^*\right)$ as a function of δ for the previously mentioned values of γ_2. We observe from Fig. 3a that $J_{\mathbb{E},|\cdot|}\left(\alpha^*\right) = 0$ for $\delta = 1$ (0 dB) (Prop. 2(i)) and is monotonically increasing in δ (Prop. 2(ii)). The expected capacity loss is bounded by (i) 0.1465 bits for $\gamma_2 = -3$ dB, (ii) 0.2500 bits for $\gamma_2 = 0$ dB, and (iii) 0.3535 bits for $\gamma_2 = 3$ dB (Prop. 2(iii)).

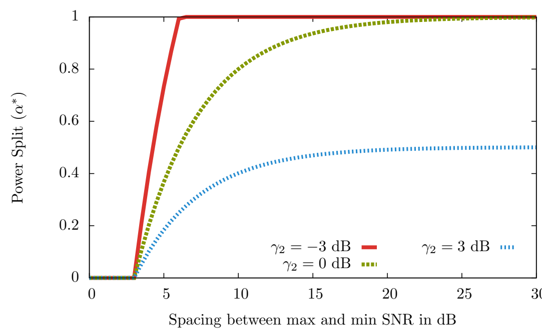
Fig. 4 shows the expected capacity loss $\mathbb{E}\left[\mathsf{C}(\Gamma)\right]$ as a function of δ for the previously mentioned values of γ_2. We observe from Fig. 4 that $\mathbb{E}\left[\mathsf{C}(\Gamma)\right]$ is positive, monotonically increasing in δ, and approaches ∞ as $\delta \to \infty$.



(a) $J_{\mathbb{E},|\cdot|}(\alpha^*)$



(b) $J_{\%,\mathbb{E}}(\alpha^*)$



(c) α^*

Figure 3: Minimum expected capacity loss (a), minimum fractional expected capacity loss (b), and the optimal power split (c) that solves (3) as a function of δ in dB for different values of γ_2.

Fig. 3b shows the fractional expected capacity loss $J_{\%,\mathbb{E}}(\boldsymbol{\alpha})$ as a function of δ for the previously mentioned values of γ_2. We observe from Fig. 3b that $J_{\%,\mathbb{E}}(\boldsymbol{\alpha})$ is increasing in δ for $0 \text{ dB} \leq \delta \leq 3 \text{ dB}$ (Prop. 4(i)) and is non-monotonic over the full range of δ (Prop. 4(ii)). We conclude by noting that while the fractional expected capacity loss approaches the limit 0 as $\delta \to \infty$ (Prop. 4(iii)), this convergence is rather slow.

5 Conclusions & Future Work

Using the K-user broadcast channel as a model for channel with K unknown states, we have characterized the optimal power allocation across a family of fixed block length codes. In the case of $K = 2$, we have shown that the spacing between the better SNR and the worst SNR must be above some threshold before the optimal transmitter will allocate power to the codeword associated with the better SNR. Additionally, we have shown that if the probability of being in the state with the better SNR is not high enough, the optimal transmitter will never allocate its entire power budget for the codeword associated with the better SNR. Looking at Fig. 2a and Fig. 2b, we observe that as the number of possible channel states gets large, the minimum expected capacity loss and fractional expected capacity loss appear to be approaching some limiting function. In subsequent work, we will use calculus of variations to characterize this limiting function as $K \to \infty$. The case of an uncountably infinite number of AMC levels was handled for Rayleigh distributed channel fades in [6]. Here we focused on a finite number of AMC values, and allow for an arbitrary distribution on the associated SNRs.

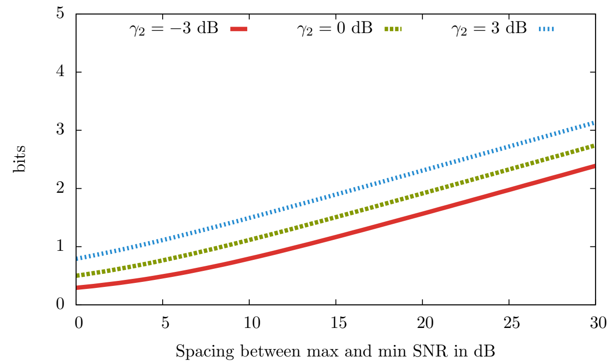


Figure 4: Expected capacity $\mathbb{E}[\mathsf{C}(\Gamma)]$ as a function of δ in dB for different values of γ_2.

References

[1] Abbas El Gamal and Young-Han Kim, *Network Information Theory*, Cambridge University Press, 2011.

[2] Thomas M. Cover, “Broadcast channels,” *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[3] Sergio Verdú and Shlomo Shamai (Shitz), “Variable-rate channel capacity,” *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2651–2667, June 2010.

[4] Patrick P. Bergmans, “Random coding theorem for broadcast channels with degraded components,” *IEEE Trans. Inf. Theory*, vol. 19, no. 2, pp. 197–207, Jan. 1973.

[5] Patrick P. Bergmans, “A simple converse for broadcast channels with additive white gaussian noise (corresp.),” *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 279–280, March 1974.

[6] Shlomo Shamai (Shitz) and Avi Steiner, “A broadcast approach for a single-user slowly fading mimo channel,” *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2617–2635, 2003.

A Worst-Case δ

Looking at Fig. 3b, we would like to characterize the δ that gives the largest fractional expected capacity loss for the case of two channel states.

Conjecture 1. *For $K = 2$ and $\gamma_1 = \delta\gamma_2$ with $\delta \geq 1$, the δ that gives the maximum fractional expected capacity loss is the solution to the following non-linear equation*

$$\begin{aligned} 0 &= (1-p_1)(\delta\gamma_2+1)(1-\delta p_1)\ln(\gamma_2+1) \\ &+ p_1\left((\delta-1)\delta\gamma_2\left(\ln\left(\frac{\delta(\gamma_2+1)(1-p_1)}{\delta-1}\right)\right.\right. \\ &+ \left.\left. p_1\ln\left(\frac{(\delta-1)p_1}{1-p_1}\right)\right) + (\delta\gamma_2+1)(1-\delta p_1)\ln(\delta\gamma_2+1)\right) \end{aligned} \tag{21}$$

Proof Sketch. It is straight-forward to show that the right-hand side of (21) is the first derivative of $J_{\%,\mathbb{E}}(\boldsymbol{\alpha})$ with respect to δ when $0 < \alpha^* < 1$. The numerical solution of (21) matches with the observed maximum values from Fig. 3b. It remains to prove uniqueness of this stationary point. □

B Counter-Proof for Convexity

Proposition 5. *The expected capacity loss $J_{\mathbb{E},|\cdot|}(\alpha^*)$ is, in general, not convex.*

Proof. We construct the following counter example. Let $K = 2$ and $\gamma_1 = \delta\gamma_2$ for $\delta \geq 1$. We then have

$$\frac{\partial^2}{\partial\delta^2}J_{\mathbb{E},|\cdot|}(\alpha^*) = \frac{\gamma_2^2(\delta(\alpha\gamma_2(\delta((p_1-1)\alpha\gamma_2+2p_1)-2)+p_1\delta)-1)}{\ln(4)(\alpha\gamma_2+1)^2(\alpha\gamma_2\delta+1)^2} \tag{22}$$

Letting $p_1 = 1/2$, $\delta = 2$, and $\gamma_2 = 1$, the (22) becomes

$$\frac{\partial^2}{\partial\delta^2}J_{\mathbb{E},|\cdot|}(\alpha^*) = \frac{1-2\alpha^2}{(2\alpha^2+3\alpha+1)^2\ln(4)} \tag{23}$$

which is positive for $\alpha < 1/\sqrt{2}$ and is negative for $\alpha > 1/\sqrt{2}$. □

What is interesting, beyond the non-convexity of the expected capacity loss, is that this non-convexity occurs for a non-pathological choice of the parameters p_1, δ, and γ_2.