Coded Modulation with Mismatched Power Control over Block-Fading Channels

Tùng T. Kim
Department of Electrical Engineering
Princeton University
Princeton, NJ 08544, USA
Email: thanhkim@princeton.edu

Albert Guillén i Fàbregas
Department of Engineering
University of Cambridge
Cambridge CB2 1PZ, UK
Email: albert.guillen@eng.cam.ac.uk

***Abstract*—Communication over delay-constrained block-fading channels with discrete inputs and imperfect channel state information at the transmitter (CSIT) is studied. The CSIT mismatch is modeled as a Gaussian random variable, whose variance decays as a power of the signal-to-noise ratio (SNR). We focus on the large-SNR behavior of the outage probability when transmit power control is used. We derive the outage exponent as a function of the system parameters, including the CSIT noise variance exponent and the exponent of the peak power constraint. It is shown that CSIT, even if noisy, is always beneficial and leads to significant gains in terms of exponents. It is also shown that when precoders are used at the transmitter, further exponent gains can be attained at the expense of higher decoding complexity.**

I. Introduction

Temporal power control across fading states can lead to dramatic improvement in the outage performance of block fading channels [1]. The intuition behind this phenomenon is that power saved in particularly bad channel conditions can be used in good channel realizations. Power control over block fading channels was originally studied under the idealistic assumptions of perfect channel state information (CSI) at the transmitter (CSIT) and Gaussian signal constellations [1]. Acquiring perfect CSIT is however a challenging task due to the temporal variation of wireless media, as well as due to the processing and transmission.

This work considers a block fading channel with *discrete input*, where the transmitter has access to a noisy version of the CSI. Similarly to [2], we model the CSIT noise as a Gaussian random variable whose variance decays as a negative power of the SNR. The rate of decaying of the CSIT noise can also be related to practical parameters in wireless systems [3]. In sharp contrast to the assumption of using Gaussian codebooks [2]–[4], the current work assumes that the input symbols are taken from a *discrete* distribution such as M-QAM or PSK. Focusing on the high signal-to-noise ratio (SNR) regime, we establish the diversity gain of block fading channels under the noisy CSIT model of interest. Note that unlike in the diversity–multiplexing tradeoff analysis [5] where the code rate grows with the SNR, herein we keep the constellation size to be 2^M at all values of the SNR and we do not let the code rate scale with the SNR. The results will shed some light into the interplay in the high-SNR regime between the number of receive antennas, the number of fading blocks, the constellation size, the code rate, as well as the SNR exponent of the CSIT noise variance and the peak exponent constraint.

II. System Model

Consider transmission over a block-fading channel with B sub-channels, where each sub-channel has a single transmit and m receive antennas (cf. Fig. 1). The mutually independent channel vectors $\boldsymbol{h}_1, \dots, \boldsymbol{h}_B$ have independent and identically distributed (i.i.d.) complex Gaussian components with zero mean and unit variance. The channel gains are constant during one fading block but change independently from one block to the next. This models a typical delay-limited scenario in wireless communications, where the delay constraint dictated by higher-layer applications prevents the system from fully exploiting time diversity [1].

The corresponding discrete-time complex baseband input-output relation for the ith sub-channel can be written as

$$\boldsymbol{Y}_i = \boldsymbol{h}_i \sqrt{P_i}\, \boldsymbol{x}_i^{\mathrm{T}} + \boldsymbol{W}_i \tag{1}$$

where $\boldsymbol{Y}_i \in \mathbb{C}^{m\times L}$ is the received signal matrix corresponding to block i, $\boldsymbol{x}_i \in \mathbb{C}^L$ is the transmitted vector in block i, $\boldsymbol{x}^{\mathrm{T}}$ denotes the transpose of $\boldsymbol{x}$, and $\boldsymbol{W}_i \in \mathbb{C}^{m\times L}$ denotes the complex additive white Gaussian noise whose entries are i.i.d. with zero mean and unit variance. We denote the block length by L and the power in block i by P_i. Hence, a codeword corresponds to BL channel uses.

We assume perfect CSIR, i.e., the receiver has perfect knowledge about all the channel gains and the powers P_i. Furthermore, we assume that the transmitter has access to a noisy version $\widehat{\boldsymbol{h}}_i$ of the true channel realization $\boldsymbol{h}_i$, so that

$$\boldsymbol{h}_i = \widehat{\boldsymbol{h}}_i + \boldsymbol{e}_i, \qquad i = 1, \dots, B \tag{2}$$

where $\boldsymbol{e}_i \in \mathbb{C}^m$ is the CSIT noise vector, independent of $\widehat{\boldsymbol{h}}_i$, with i.i.d. Gaussian components with zero mean and variance σ_{e}^2. This model of the CSIT has been well motivated in many different contexts, such as in scenarios with delayed feedback, noisy feedback, or in systems exploiting channel reciprocity [6], [7]. We further assume, as in [2], that the CSIT noise variance decays as a power of the SNR, $\sigma_{\mathrm{e}}^2 = \mathrm{SNR}^{-d_{\mathrm{e}}}$, for some $d_{\mathrm{e}} > 0$. Thus we consider a family of channels where the second-order statistic of the CSIT noise varies with

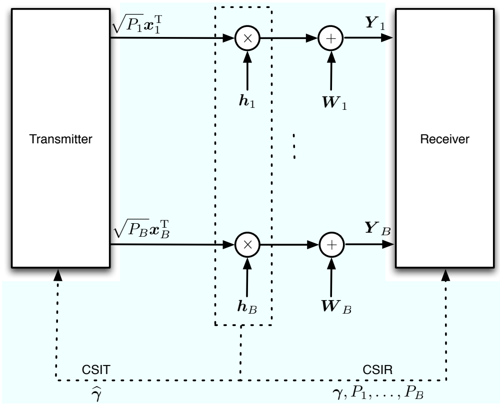

Fig. 1. System model and CSI assumptions.

SNR. If the CSIT for example is estimated from the reverse link (exploiting reciprocity in a TDD system), its quality will depend on the SNR of reverse link and not the forward link. However, while the SNRs of the forward and reverse links are different, this difference will be fully captured by changing the values of d_{e}. For convenience, we introduce the normalized channel gains $\bar{\boldsymbol{h}}_i = \frac{\sqrt{2}}{\sigma_{\mathrm{e}}}\boldsymbol{h}_i$. Given $\widehat{\boldsymbol{h}}_i$ then $\bar{\boldsymbol{h}}_i$ is complex Gaussian with mean $\frac{\sqrt{2}}{\sigma_{\mathrm{e}}}\widehat{\boldsymbol{h}}_i$ and a scaled identity covariance matrix.

Let $\gamma_i \triangleq \|\boldsymbol{h}_i\|^2$ be the fading magnitude of block i and $\boldsymbol{\gamma} = [\gamma_1 \cdots \gamma_B]$. Further denote $\bar{\gamma}_i \triangleq \|\bar{\boldsymbol{h}}_i\|^2$, $\hat{\gamma}_i \triangleq \|\widehat{\boldsymbol{h}}_i\|^2$, $\bar{\boldsymbol{\gamma}} \triangleq [\bar{\gamma}_1 \cdots \bar{\gamma}_B]$ and $\widehat{\boldsymbol{\gamma}} \triangleq [\hat{\gamma}_1 \cdots \hat{\gamma}_B]$.

III. Preliminaries

We assume transmission at a fixed-rate R using a coded modulation scheme $\mathcal{M} \subset \mathbb{C}^{BL}$ of length BL constructed over a signal constellation $\mathcal{X} \subset \mathbb{C}$ of size 2^M such as 2^M-PSK or QAM. We denote the codewords of $\mathcal{M}$ by $\boldsymbol{x} = (\boldsymbol{x}_1^{\mathrm{T}}, \ldots, \boldsymbol{x}_B^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{C}^{BL}$. We assume that the signal constellation $\mathcal{X}$ is zero mean and normalized in energy, i.e., $\mathbb{E}[X] = 0$ and $\mathbb{E}[|X|^2] = 1$, where X denotes the corresponding random variable. We denote the input distribution as $Q(x)$. The instantaneous input-output mutual information of the channel is given by $I(\boldsymbol{\gamma}) = \frac{1}{B}\sum_{i=1}^{B} I_{\mathcal{X}}(P_i\gamma_i)$ where

$$I_{\mathcal{X}}(s) = \mathbb{E}\left[\log_2 \frac{e^{-|Y-\sqrt{s}X|^2}}{\sum_{x'\in\mathcal{X}} Q(x')e^{-|Y-\sqrt{s}x'|^2}}\right] \tag{3}$$

is the input-output mutual information of an AWGN channel with SNR s using the signal constellation $\mathcal{X}$.

The outage probability is commonly defined as [8], [9] $P_{\mathrm{out}}(R) \triangleq \Pr\{I(\boldsymbol{\gamma}) < R\}$. In this work, we are interested in the SNR exponents of the outage probability [5], [10], i.e., $d_{\mathrm{out}} \triangleq \lim_{\mathrm{SNR}\to\infty} -\frac{\log P_{\mathrm{out}}(R)}{\log \mathrm{SNR}}$. We adopt the notation $g(\mathrm{SNR}) \doteq \mathrm{SNR}^a \Leftrightarrow \lim_{\mathrm{SNR}\to\infty} \frac{\log g(\mathrm{SNR})}{\log \mathrm{SNR}} = a$.

It has been shown in [10], [11] that the outage exponent without CSIT is given by $d_{\mathrm{out}} = md_{\mathrm{sb}}(R)$ where

$$d_{\mathrm{sb}}(R) \triangleq 1 + \left\lfloor B\left(1 - \frac{R}{M}\right)\right\rfloor = B - \left\lceil \frac{BR}{M}\right\rceil + 1, \tag{4}$$

with $\lfloor x \rfloor$ being the largest integer that is not larger than x and $\lceil x \rceil$ being the smallest integer that is not smaller than x, is the Singleton bound on the block-diversity of the coded modulation scheme $\mathcal{M}$ [10], [12], [13].

The transmitter can adapt the transmitted powers P_i to the (noisy) channel conditions $\widehat{\boldsymbol{\gamma}}$. We consider power allocation algorithms that treat the noisy CSIT $\widehat{\boldsymbol{\gamma}}$ as if it were perfect. We consider an average power constraint, such that

$$\mathbb{E}\left[\frac{1}{B}\sum_{i=1}^{B} P_i(\widehat{\boldsymbol{\gamma}})\right] = \mathbb{E}\left[P(\widehat{\boldsymbol{\gamma}})\right] \le \mathrm{SNR} \tag{5}$$

where $P(\widehat{\boldsymbol{\gamma}}) = \frac{1}{B}\sum_{i=1}^{B} P_i(\widehat{\boldsymbol{\gamma}})$ is the instantaneous average power allocated given $\widehat{\boldsymbol{\gamma}}$. The SNR herein has the meaning of the average transmit power over infinitely many fading blocks. We introduce a peak-to-average power constraint

$$P(\widehat{\boldsymbol{\gamma}}) \le \mathrm{SNR}^{d_{\mathrm{peak}}} \tag{6}$$

where d_{peak} is interpreted as the peak-to-average power SNR exponent. The case $d_{\mathrm{peak}} = 1$ represents a system whose allocated power is dominated by the peak-power constraint. By allowing d_{peak} to take an arbitrary value, we can model a family of systems with different behavior in the peak power constraint. In the high-SNR regime of interest, we can for example scale the right hand side of (6) by a constant without changing any conclusion. That is, any constant, finite ratios between the peak and the average power provides the same asymptotic behavior as $d_{\mathrm{peak}} = 1$.

The corresponding minimum-outage power allocation rule is the solution to the following problem

$$\begin{cases} \text{Minimize} & P_{\mathrm{out}}(R) \\ \text{subject to} & \mathbb{E}\left[\frac{1}{B}\sum_{i=1}^{B} P_i(\widehat{\boldsymbol{\gamma}})\right] \le \mathrm{SNR} \\ & \frac{1}{B}\sum_{i=1}^{B} P_i(\widehat{\boldsymbol{\gamma}}) \le \mathrm{SNR}^{d_{\mathrm{peak}}} \\ & P_i(\widehat{\boldsymbol{\gamma}}) \ge 0, \quad i = 1, \ldots, B. \end{cases} \tag{7}$$

Solving the minimum-outage power allocation problem even numerically is difficult in general, given our noisy CSIT model and the discreteness of $\mathcal{X}$. Nevertheless, we can characterize the asymptotic behavior of the optimal solution in the high SNR regime. Following [5], note that the outage exponent of the optimal algorithm is the same as that of a power control system that allocates power uniformly across the blocks, i.e, $P_i(\widehat{\boldsymbol{\gamma}}) = P(\widehat{\boldsymbol{\gamma}})$, $\forall i = 1, \ldots, B$.

IV. Asymptotic Outage Behavior

A. Main Result

In this section, we study the asymptotic behavior of the outage probability. In particular, our main results in terms of outage SNR exponents are stated as follows.

Theorem 1: Consider transmission at rate R over a block-fading channel described by (1) with Rayleigh fading with mismatched CSIT modeled by (2) with inputs drawn from $\mathcal{X}$. The transmitter uses power control with an average power

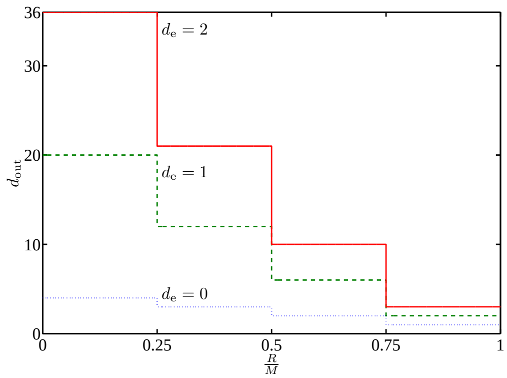

Fig. 2. Outage exponents for $B = 4$, $m = 1$ and $d_{\text{peak}} > 1 + md_e d_{\text{sb}}(R)$.

constraint (5) and a peak-to-average power constraint (6). Then, the outage exponents are given by

$$d(R, d_e, d_{\text{peak}}) = \begin{cases} md_{\text{sb}}(R)d_{\text{peak}} & d_{\text{peak}} \leq 1 + md_{\text{sb}}(R)d_e, \\ md_{\text{sb}}(R)\left(1 + md_{\text{sb}}(R)d_e\right) & \text{otherwise.} \end{cases} \tag{8}$$

Proof: See Appendix A. ∎

In Fig. 2 we plot the outage exponents for $B = 4$, $m = 1$ with no CSIT (or $d_e = 0$) and with noisy CSIT with $d_e = 1, 2$ when $d_{\text{peak}} > 1 + md_e d_{\text{sb}}(R)$. As we observe from the figure, increasing d_e yields a better exponent. Note that in this case, when the CSIT is perfect the exponent is infinitely large. Even with imperfect CSIT, large gains are possible by power control.

In the extreme case $d_{\text{peak}} = 1$, we obtain $d(R, d_e, 1) = md_{\text{sb}}(R)$, which is the outage exponent for a system without power control [10], [11]. Increasing d_{peak} subsequently leads to an improvement in the outage performance. However, when d_{peak} exceeds a certain threshold, the diversity gain is "saturated" due to the limitation on the accuracy of the CSIT. In other words, a stringent constraint on the peak power exponent leads to a lot more pronounced detrimental effect in the case of accurate CSIT than in the case of very noisy CSIT.

In the case $d_e \downarrow 0$ (very noisy CSIT) we have $d(R, d_e, d_{\text{peak}}) \to md_{\text{sb}}(R)$, equal the outage exponent when there is no CSIT [10], [11]. Thus if the CSIT noise variance has the same order of magnitude as the channel gains, then no extra diversity gain can be obtained from power control.

In case $d_e \to \infty$ (CSIT noise variance decays exponentially or faster with the SNR), $d(R, d_e) \to \infty$, $\forall R < M$, as long as the peak exponent constraint is also relaxed to satisfy $d_{\text{peak}} > 1 + md_{\text{sb}}(R)d_e$. For strictly positive and finite d_e, using power control, even with noisy CSIT, provides an extra diversity gain of $\left(md_{\text{sb}}(R)\right)^2 d_e$ compared to the no-CSIT case, as long as the peak constraint is sufficiently relaxed.

We also learn from the analysis in Appendix A that at high SNR, when $d_{\text{peak}} \geq 1 + md_e\left(B - \left\lceil\frac{BR}{M}\right\rceil + 1\right)$, the dominant outage event occurs when exactly $\left\lceil\frac{BR}{M}\right\rceil - 1$ of the channel gain estimates $\hat{\gamma}_i$'s are much larger than the noise variance σ_e^2, and the remaining $B - \left\lceil\frac{BR}{M}\right\rceil + 1$ channel estimates have the same order of magnitude as σ_e^2. When d_{peak} is smaller, however, the system cannot "invert" the worst channel realizations and the peak exponent becomes the limiting factor.

B. Improving the Outage Exponent with Rotations

In [14], it is shown that a precoding technique can be used to improve the outage exponent over fading channels with discrete inputs. We demonstrate how that idea can be applied in the current noisy CSIT setting of interest to further improve the outage exponents. To simplify the presentation, we remove the peak exponent constraint (setting $d_{\text{peak}} = \infty$), focusing only on the effects of the CSIT noise.

We briefly recall the precoding technique of [14]. First consider reformatting the codewords $\boldsymbol{x} \in \mathcal{M}$ as matrices $\boldsymbol{X} \in \mathbb{C}^{B \times L}$. We now obtain $\boldsymbol{X}$ as

$$\boldsymbol{X} = \boldsymbol{M}\boldsymbol{S} \tag{9}$$

where

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{M}_1 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{M}_K \end{pmatrix} \in \mathbb{C}^{B \times B} \tag{10}$$

is a unitary block-diagonal matrix, and the entries of $\boldsymbol{S} \in \mathbb{C}^{B \times L}$ belong to the signal constellation $\mathcal{X}$ with size 2^M symbols. The matrices $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_K \in \mathbb{C}^{N \times N}$ are the K unitary rotation matrices of dimension N each. $\boldsymbol{M}_k(\boldsymbol{s} - \boldsymbol{s}') \neq \boldsymbol{0}$ componentwise, for all $\boldsymbol{x} \neq \boldsymbol{x}' \in \mathcal{X}^N$. This implies that if the vector $(\boldsymbol{s} - \boldsymbol{s}')$ has a positive number of nonzero entries, then, its rotated version will have all N entries different from zero.

According to [14], with no CSIT we obtain the exponent $d_{\text{out}} = md_{\text{sb}}^{\text{rot}}(R)$ where

$$d_{\text{sb}}^{\text{rot}}(R) \triangleq N\left(1 + \left\lfloor\frac{B}{N}\left(1 - \frac{R}{M}\right)\right\rfloor\right) = B + N - N\left\lceil\frac{BR}{MN}\right\rceil.$$

With noisy CSIT, completely similarly to the previous section we have the following result.

Theorem 2: Consider transmission over a block-fading channel described by (1) with mismatched CSIT modeled by (2) with inputs obtained as the rotation of a coded modulation scheme over $\mathcal{X}$ as described by (9), using full diversity rotations. The transmitter uses power control with an average power constraint (5). Then, the outage exponents are given by

$$d(R, d_e) = md_{\text{sb}}^{\text{rot}}(R)\left(1 + md_{\text{sb}}^{\text{rot}}(R)d_e\right). \tag{11}$$

We illustrate in Fig. 3 the effect of full-diversity rotation matrices on the outage exponent of the coded modulation system with mismatched CSIT. This precoding method clearly leads to a higher diversity gain even at high code rates, at the expense of increasing receiver complexity.

In the case $N = B$, i.e. when a single matrix that rotates all B output symbols is used, then $d(R, d_e) = mB(1 + mBd_e)$. This is the maximum diversity gain we can achieve in this scenario, even with codes drawn from a Gaussian ensemble [3]. However, the receiver complexity will increase exponentially in N. Note also in terms of exponents, there is nothing to gain

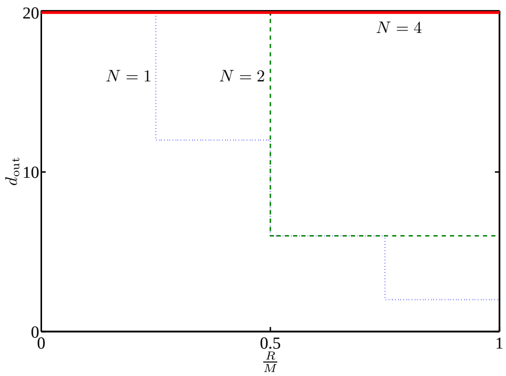

Fig. 3. Outage exponents for $B=4$, $m=1$, $d_{\text{e}}=1$ and full-diversity rotations of size $N=1$ (dotted line), $N=2$ (dashed line) and $N=4$ (solid line).

in optimizing the full precoding matrix. Finally, we notice that determining the outage exponent in the more general multiple-input multiple-transmit settings remains an open problem.

Appendix A
Proof of Theorem 1 (Sketch)

Let us invoke the standard change of variables as in [5], $\bar{\alpha}_i = -\frac{\log \bar{\gamma}_i}{\log \text{SNR}}$ and $\hat{\alpha}_i = -\frac{\log \hat{\gamma}_i}{\log \text{SNR}}$. We also perform the change of variable $\pi(\widehat{\boldsymbol{\gamma}}) \equiv \pi(\widehat{\boldsymbol{\alpha}}) \triangleq \frac{\log P(\widehat{\boldsymbol{\gamma}})}{\log \text{SNR}}$.

The power constraint (5) asymptotically becomes [3]

$$\int \text{SNR}^{\pi(\widehat{\boldsymbol{\gamma}})} f(\widehat{\boldsymbol{\gamma}}) d\widehat{\boldsymbol{\gamma}} \dot{\leq} \text{SNR}^1. \tag{12}$$

The $\hat{\gamma}_i$'s are mutually independent and follow Chi-square distribution with $2m$ degrees of freedom. Also, $\text{E}[\gamma_i] = \text{E}[\|\boldsymbol{h}_i\|^2] - \text{E}[\|\boldsymbol{e}_i\|^2] \doteq \text{SNR}^0$. Changing variables from $\widehat{\boldsymbol{\gamma}}$ to $\widehat{\boldsymbol{\alpha}}$, we readily obtain

$$\int_{\widehat{\boldsymbol{\alpha}} \in \mathbb{R}_+^B} \text{SNR}^{\pi(\widehat{\boldsymbol{\alpha}})} \text{SNR}^{-m\sum_{i=1}^B \hat{\alpha}_i} d\widehat{\boldsymbol{\alpha}} \dot{\leq} \text{SNR}^1. \tag{13}$$

Herein we have neglected the terms irrelevant to the SNR exponent, noticing that for any set containing $\alpha_i < 0$, its probability measure decays exponentially in SNR [5]. Applying Varadhan's integral lemma we then have

$$\sup_{\widehat{\boldsymbol{\alpha}} \in \mathbb{R}_+^B} \left\{ \pi(\widehat{\boldsymbol{\alpha}}) - m \sum_{i=1}^B \hat{\alpha}_i \right\} \leq 1. \tag{14}$$

Since outage probability is a non-increasing function of transmit power, we conclude that with the optimal power allocation,

$$\pi(\widehat{\boldsymbol{\alpha}}) = \min \left(d_{\text{peak}}, 1 + m \sum_{i=1}^B \hat{\alpha}_i, \right) \tag{15}$$

where we need to introduce d_{peak} to take into account (6).

From [10] it is known that as SNR $\to \infty$ the mutual information in sub-channel i, $I_{\mathcal{X}}\left(P(\widehat{\boldsymbol{\gamma}})\gamma_i\right)$, tends to either M or 0 depending only on the behavior of the term

$$P(\widehat{\boldsymbol{\gamma}})\gamma_i = \text{SNR}^{\min\left(d_{\text{peak}}, 1+m\sum_{j=1}^B \hat{\alpha}_j\right) - d_{\text{e}} - \bar{\alpha}_i}. \tag{16}$$

In particular, if $\bar{\alpha}_i \leq \pi(\widehat{\boldsymbol{\alpha}}) - d_{\text{e}}$ then $I_{\mathcal{X}}\left(P(\widehat{\boldsymbol{\gamma}})\gamma_i\right) \to M$ bits per channel use. Otherwise $I_{\mathcal{X}}\left(P(\widehat{\boldsymbol{\gamma}})\gamma_i\right) \to 0$.

Thus the asymptotic outage set is given by

$$\mathcal{O} = \{\bar{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\alpha}} : \\ \sum_{i=1}^B \mathbf{1}\left(\bar{\alpha}_i \leq \min\left(d_{\text{peak}}, 1 + m \sum_{j=1}^B \hat{\alpha}_j \right) - d_{\text{e}} \right) < \frac{BR}{M} \right\}$$

where $\mathbf{1}(\cdot)$ is the indicator function. We then have

$$P_{\text{out}}(R) \doteq \int_{\mathcal{O}} f(\bar{\boldsymbol{\alpha}}|\widehat{\boldsymbol{\alpha}}) f(\widehat{\boldsymbol{\alpha}}) d\bar{\boldsymbol{\alpha}} d\widehat{\boldsymbol{\alpha}}. \tag{17}$$

Notice that $f(\bar{\boldsymbol{\gamma}}|\widehat{\boldsymbol{\gamma}}) = \prod_{i=1}^B f(\bar{\gamma}_i|\hat{\gamma}_i)$, where the conditional p.d.f $f(\bar{\gamma}_i|\hat{\gamma}_i)$ is a non-central chi-square one with $2m$ degrees of freedom. Using a result in [3], we can asymptotically expand the integral (17), showing that $d(R; d_{\text{e}}, d_{\text{peak}}) = \min(d_0, \ldots, d_B)$ with d_n being defined such that

$$\int_{\mathcal{O} \cap \mathcal{B}_n} \prod_{i=1}^{B-n} \text{SNR}^{-m\hat{\alpha}_i - m\bar{\alpha}_i} \prod_{j=B-n+1}^{B} \text{SNR}^{-m\hat{\alpha}_j} d\bar{\boldsymbol{\alpha}} d\widehat{\boldsymbol{\alpha}} \doteq \text{SNR}^{-d_n}$$

where

$$\begin{aligned} \mathcal{B}_n \triangleq \{ & \bar{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\alpha}} : \{\bar{\alpha}_1 > 0, \hat{\alpha}_1 \geq d_{\text{e}}\} \cap \cdots \\ & \cap \{\bar{\alpha}_{B-n} > 0, \hat{\alpha}_{B-n} \geq d_{\text{e}}\} \\ & \cap \{0 \leq \hat{\alpha}_{B-n+1} < d_{\text{e}}, \bar{\alpha}_{B-n+1} = \hat{\alpha}_{B-n+1} - d_{\text{e}}\} \cap \cdots \\ & \cap \{0 \leq \hat{\alpha}_B < d_{\text{e}}, \bar{\alpha}_B = \hat{\alpha}_B - d_{\text{e}}\}\}. \end{aligned}$$

Thus applying Varadhan's integral lemma gives

$$d_n = \inf_{\bar{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\alpha}} \in \mathcal{O} \cap \mathcal{B}_n} \left\{ m \sum_{i=1}^B \hat{\alpha}_i + m \sum_{j=1}^{B-n} \bar{\alpha}_j \right\}. \tag{18}$$

Over $\mathcal{B}_n$, we have $\bar{\alpha}_i = \hat{\alpha}_i - d_{\text{e}}$ for all $i \geq B-n+1$, thus

$$\begin{aligned} \mathcal{O} = \Bigg\{ \bar{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\alpha}} : & \sum_{i=1}^{B-n} \mathbf{1}\left(\bar{\alpha}_i \leq \min\left(d_{\text{peak}}, 1 + m \sum_{j=1}^B \hat{\alpha}_j \right) - d_{\text{e}} \right) \\ & + \sum_{i=B-n+1}^{B} \mathbf{1}\left(\hat{\alpha}_i \leq \min\left(d_{\text{peak}}, 1 + m \sum_{j=1}^B \hat{\alpha}_j \right) \right) < \frac{BR}{M} \Bigg\}. \end{aligned}$$

To compute d_n, we consider two mutual exclusively cases.

Case 1: $d_{\text{peak}} < 1 + m\sum_{j=1}^B \hat{\alpha}_j$. Denote the SNR exponent over the intersection of this region and $\mathcal{B}_n$ as $d_n^{(1)}$.

Case 1.1: If $d_{\text{peak}} < d_{\text{e}}$ then $\mathbf{1}\left(\bar{\alpha}_i \leq d_{\text{peak}} - d_{\text{e}}\right) = 0$, $\forall i \in \{1, \ldots, B-n\}$. The outage set reduces to

$$\mathcal{O} = \left\{ \widehat{\boldsymbol{\alpha}} : \sum_{i=B-n+1}^{B} \mathbf{1}\left(\hat{\alpha}_i \leq d_{\text{peak}}\right) < \frac{BR}{M} \right\}. \tag{19}$$

Because for $i = 1, \ldots, B-n$, the terms $\hat{\alpha}_i$ and $\bar{\alpha}_i$ are not present in the outage set, we have the optimal solution to (18) $\bar{\alpha}_1^* = \cdots = \bar{\alpha}_{B-n}^* = 0$ and $\sum_{i=1}^{B-n} \hat{\alpha}_i^* = \max(d_{\text{peak}} - 1, m(B-n)d_{\text{e}})$, due to the constraint $d_{\text{peak}} < 1 + m\sum_{j=1}^B \hat{\alpha}_j$.

After some manipulation, we have that if $d_{\mathrm{peak}} < d_{\mathrm{e}}$ then

$$d_n^{(1)} = \begin{cases} m(B-n)d_{\mathrm{e}} & \text{if } \frac{BR}{M} > n, \\ m(B-n)d_{\mathrm{e}} + md_{\mathrm{peak}}\left(n - \left\lceil \frac{BR}{M} \right\rceil + 1\right) & \text{if } \frac{BR}{M} \le n. \end{cases} \tag{20}$$

Case 1.2: On the other hand, if $d_{\mathrm{peak}} \ge d_{\mathrm{e}}$, then for $i = B-n+1, \ldots, B$ we have $\mathbf{1}(\hat{\alpha}_i \le d_{\mathrm{peak}}) = 1$ because in $\mathcal{B}_n$, $\hat{\alpha}_i < d_{\mathrm{e}}$ for these values of i. The outage set reduces to

$$\mathcal{O} = \left\{ \bar{\boldsymbol{\alpha}}, \hat{\boldsymbol{\alpha}} : \sum_{i=1}^{B-n} \mathbf{1}\left(\bar{\alpha}_i \le d_{\mathrm{peak}} - d_{\mathrm{e}}\right) < \frac{BR}{M} - n \right\}. \tag{21}$$

Note that if $\frac{BR}{M} \le n$ then $d_n^{(1)} = \infty$ because the set of "bad" channel realizations is empty [4].

After some manipulation, we have that if $d_{\mathrm{peak}} \ge d_{\mathrm{e}}$ then

$$d_n^{(1)} = \begin{cases} m(d_{\mathrm{peak}} - d_{\mathrm{e}})\left(B - n + 1 - \left\lceil \frac{BR}{M} - n \right\rceil\right) & \\ + \max\left(d_{\mathrm{peak}} - 1, m(B-n)d_{\mathrm{e}}\right) & \text{if } \frac{BR}{M} > n \\ \infty & \text{if } \frac{BR}{M} \le n. \end{cases} \tag{22}$$

Case 2: $d_{\mathrm{peak}} \ge 1 + m\sum_{j=1}^{B} \hat{\alpha}_j$. Note that over $\mathcal{B}_n$ we have $\sum_{j=1}^{B} \hat{\alpha}_j \ge (B-n)d_{\mathrm{e}}$ thus Case 2 can only happen if $d_{\mathrm{peak}} \ge 1 + m(B-n)d_{\mathrm{e}}$. For n such that $d_{\mathrm{peak}} < 1 + m(B-n)d_{\mathrm{e}}$, we use the convention $d_n^{(2)} = \infty$. Then, over $\mathcal{B}_n$

$$\mathcal{O} = \left\{ \bar{\boldsymbol{\alpha}}, \hat{\boldsymbol{\alpha}} : \sum_{i=1}^{B-n} \mathbf{1}\left(\bar{\alpha}_i \le 1 + m\sum_{j=1}^{B} \hat{\alpha}_j - d_{\mathrm{e}}\right) < \frac{BR}{M} - n \right\}.$$

If $\frac{BR}{M} \le n$ then the outage probability decays exponentially in SNR. We obtain $\hat{\alpha}_1^* = \cdots = \hat{\alpha}_{B-n}^* = d_{\mathrm{e}}$ and $\hat{\alpha}_{B-n+1}^* = \cdots = \hat{\alpha}_B^* = 0$. We also have $\bar{\alpha}_i^* = 1 + m(B-n)d_{\mathrm{e}} - d_{\mathrm{e}}$, for exactly $B - n - \left\lceil \frac{BR}{M} - n \right\rceil + 1$ of the $\bar{\alpha}_i$'s, and the other $\bar{\alpha}_i$'s are zero. Thus

$$d_n^{(2)} = \begin{cases} m(B-n)d_{\mathrm{e}} + m\left(B - n - \left\lceil \frac{BR}{M} - n \right\rceil + 1\right) \times & \\ \times(1 + m(B-n)d_{\mathrm{e}} - d_{\mathrm{e}}) & \text{if } \frac{BR}{M} > n, \\ \infty & \text{if } \frac{BR}{M} \le n. \end{cases} \tag{23}$$

We combine the results in Case 1 and 2 to find $d(R, d_{\mathrm{e}}, d_{\mathrm{peak}})$. If $d_{\mathrm{peak}} < d_{\mathrm{e}}$ then we have

$$\begin{aligned} d(R, d_{\mathrm{e}}, d_{\mathrm{peak}}) &= \min(d_0^{(1)}, d_1^{(1)}, \ldots, d_B^{(1)}) \\ &= md_{\mathrm{peak}}\left(B - \left\lceil \frac{BR}{M} \right\rceil + 1.\right) \end{aligned} \tag{24}$$

We now consider the case $d_{\mathrm{peak}} \ge d_{\mathrm{e}}$, where the $d_n^{(1)}$'s are given by (22). There are three possibilities.

Case A: If $d_{\mathrm{peak}} \ge 1 + mBd_{\mathrm{e}}$ then $d_{\mathrm{peak}} \ge 1 + m(B-n)d_{\mathrm{e}}$, $\forall n = 0, \ldots, B$. Thus

$$d_n^{(1)} = \begin{cases} m(d_{\mathrm{peak}} - d_{\mathrm{e}})\left(B - n + 1 - \left\lceil \frac{BR}{M} - n \right\rceil\right) & \\ + d_{\mathrm{peak}} - 1 & \text{if } \frac{BR}{M} > n \\ \infty & \text{if } \frac{BR}{M} \le n. \end{cases} \tag{25}$$

It can then be shown that

$$\begin{aligned} &d(R, d_{\mathrm{e}}, d_{\mathrm{peak}}) = \min(d_0^{(2)}, d_1^{(2)}, \ldots, d_B^{(2)}) \\ &= m\left(B - \left\lceil \frac{BR}{M} \right\rceil + 1\right)\left(1 + m\left(B - \left\lceil \frac{BR}{M} \right\rceil + 1\right)d_{\mathrm{e}}\right). \end{aligned}$$

Case B: $1 + md_{\mathrm{e}}\left(B - \left\lceil \frac{BR}{M} \right\rceil + 1\right) < d_{\mathrm{peak}} < 1 + mBd_{\mathrm{e}}$. This implies $\frac{BR}{M} \ge \left\lceil \frac{BR}{M} \right\rceil - 1 > B - \frac{d_{\mathrm{peak}} - 1}{md_{\mathrm{e}}}$. It can be shown that in this case

$$\begin{aligned} &d(R, d_{\mathrm{e}}, d_{\mathrm{peak}}) \\ &= d^{(2)}_{\left\lceil \frac{BR}{M} \right\rceil - 1} \\ &= m\left(B - \left\lceil \frac{BR}{M} \right\rceil + 1\right)\left(1 + m\left(B - \left\lceil \frac{BR}{M} \right\rceil + 1\right)d_{\mathrm{e}}\right). \end{aligned}$$

Case C: $d_{\mathrm{peak}} \le 1 + md_{\mathrm{e}}\left(B - \left\lceil \frac{BR}{M} \right\rceil + 1\right)$. This implies $\left\lceil \frac{BR}{M} \right\rceil - 1 \le B - \frac{d_{\mathrm{peak}} - 1}{md_{\mathrm{e}}}$. Thus for any integer n such that $n < \frac{BR}{M}$ then $n < B - \frac{d_{\mathrm{peak}} - 1}{md_{\mathrm{e}}}$ leading to $d_{\mathrm{peak}} < 1 + md_{\mathrm{e}}(B-n)$. Hence from (22) we have

$$d_n^{(1)} = \begin{cases} m(d_{\mathrm{peak}} - d_{\mathrm{e}})\left(B - n + 1 - \left\lceil \frac{BR}{M} - n \right\rceil\right) & \\ + md_{\mathrm{e}}(B-n) & \text{if } \frac{BR}{M} > n \\ \infty & \text{if } \frac{BR}{M} \le n. \end{cases}$$

Since $n < \frac{BR}{M}$ leads to $d_{\mathrm{peak}} < 1 + md_{\mathrm{e}}(B-n)$, we also have $d_n^{(2)} = \infty$, $\forall n$. Thus

$$\begin{aligned} d(R, d_{\mathrm{e}}, d_{\mathrm{peak}}) &= \min(d_0^{(1)}, \ldots, d_B^{(1)}) \\ &= md_{\mathrm{peak}}\left(B - \left\lceil \frac{BR}{M} \right\rceil + 1\right). \end{aligned} \tag{26}$$

References

[1] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1468–1489, Jul. 1999.

[2] A. Lim and V. K. N. Lau, "On the fundamental tradeoff of spatial diversity and spatial multiplexing of MISO/SIMO links with imperfect CSIT," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 110–117, Jan. 2008.

[3] T. T. Kim and G. Caire, "Diversity gains of power control in MIMO channels with noisy CSIT," *IEEE Trans. Inf. Theory*, pp. 1618–1626, Apr. 2009.

[4] T. T. Kim and M. Skoglund, "Diversity-multiplexing tradeoff in MIMO channels with partial CSIT," *IEEE Trans. Inf. Theory*, vol. 53, pp. 2743–2759, Aug. 2007.

[5] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1073–1096, May 2003.

[6] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, pp. 2632–2639, Sep. 2001.

[7] G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inf. Theory*, vol. 48, pp. 611–627, Mar. 2002.

[8] L. H. Ozarow, S. Shamai (Shitz), and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 359–378, May 1994.

[9] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.

[10] A. Guillén i Fàbregas and G. Caire, "Coded modulation in the block-fading channel: Coding theorems and code construction," *IEEE Trans. Inf. Theory*, vol. 52, pp. 91–114, Jan. 2006.

[11] K. D. Nguyen, A. Guillén i Fàbregas, and L. K. Rasmussen, "A tight lower bound to the outage probability of discrete-input block-fading channels," *IEEE Trans. Inf. Theory*, vol. 53, pp. 4314–4322, Nov. 2007.

[12] R. Knopp and P. A. Humblet, "On coding for block fading channels," *IEEE Trans. Inf. Theory*, vol. 46, pp. 189–205, Jan. 2000.

[13] E. Malkamäki and H. Leib, "Coded diversity on block-fading channels," *IEEE Trans. Inf. Theory*, vol. 45, pp. 771–781, Mar. 1999.

[14] A. Guillén i Fàbregas and G. Caire, "Multidimensional coded modulation in block-fading channels," *IEEE Trans. Inf. Theory*, vol. 54, pp. 2367–2372, May 2008.