Mismatched CSI Outage Exponents of Block-Fading Channels

A. Taufiq Asyhari
Department of Engineering
University of Cambridge
Cambridge, CB2 1PZ, UK
taufiq-a@ieee.org

Albert Guillén i Fàbregas
Department of Engineering
University of Cambridge
Cambridge, CB2 1PZ, UK
guillen@ieee.org

***Abstract*—We study block-fading channels where both transmitter and receiver do not know the actual channel state information (CSI) but they have access to a noisy version. We study the interplay between estimation error variances at the transmitter and at the receiver to give the optimal outage exponents. We also demonstrate that achieving a reliable channel estimate at the receiver is more important than obtaining a reliable channel state information at the transmitter in terms of outage exponent.**

I. Introduction

The block-fading channel is a widely-accepted model for transmission of delay-limited applications over slowly-varying fading wireless channels [1]. In such a channel, each codeword spans only a finite number of fading blocks. The channel gain remains constant in a block but varies from block to block.

Common assumptions to study the performance of block-fading channels are perfect channel state information (CSI) at the receiver [2], [3] or at both the transmitter and receiver [4]. Some recent works focused the study on imperfect CSI at either side of communication ends. References [5], [6] studied mismatched CSI at the transmitter (CSIT) whereas reference [7] studied mismatched CSI at the receiver (CSIR). This paper proposes a unified framework for studying mismatched CSI at both communication ends. This is a realistic assumption since obtaining channel estimates for both transmitter and receiver can be challenging due to the time-varying nature of the channel, additive noise and the hardware complexity.

The key finding of this paper is an exact characterisation of the outage diversity (signal-to-noise ratio (SNR) exponents) $d_{\rm icsi}$ (see Fig. 1), which in turn describes the interplay between CSIR and CSIT estimation error parameters. Depending on the CSIT and CSIR noises, we identify the region for which mismatched CSIT and CSIR dominate the outage exponent. It is shown that if CSIR is not reliable (i.e. $d_{\hat{e}} \leq 1$), power control schemes cannot improve the outage diversity. Thus, achieving a reliable channel estimate at the receiver is more important than at the transmitter.

Notation: W and $\boldsymbol{W}$ are random scalar and vector. w and $\boldsymbol{w}$ denote scalar and vector. $\lfloor w \rfloor$ ($\lceil w \rceil$) is the largest (smallest) integer smaller (larger) than or equal to w. $[w]_+ \triangleq \max\{0, w\}$.

[1]This work has been partly supported by the Yousef Jameel Scholarship.

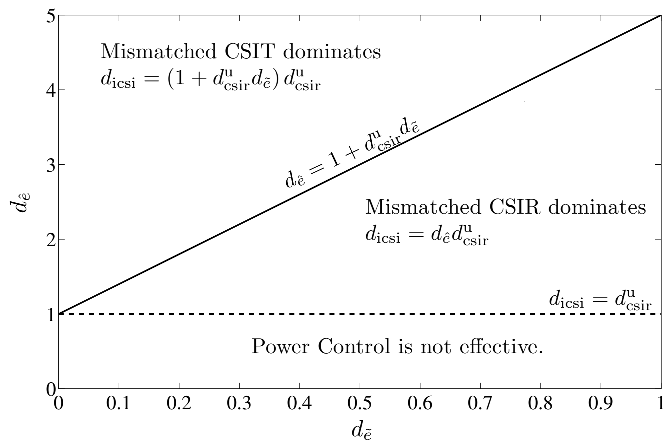

Fig. 1. Interplay between CSIT and CSIR estimation error variance exponents and outage exponents with *full-CSIT* power allocation. The variances of CSIT and CSIR noises are $P^{-d_{\tilde{e}}}$ and $P^{-d_{\hat{e}}}$, where P is the average SNR. $d^{\rm u}_{\rm csir}$ is the outage diversity with uniform power allocation and perfect CSIR [3].

II. System Model

We consider a single-input single-output (SISO) block-fading channel with B fading blocks per codeword. The output of the channel at block b is a $T-$dimensional random vector

$$\boldsymbol{Y}_b = \sqrt{p_b} H_b \boldsymbol{x}_b + \boldsymbol{Z}_b, \quad b = 1, \ldots, B \qquad (1)$$

where $\boldsymbol{Z}_b \in \mathbb{C}^T$ is the additive white Gaussian noise (AWGN) vector; $\boldsymbol{x}_b \in \mathcal{X}^T$ is the transmitted vector; T, $\mathcal{X} \subseteq \mathbb{C}$ and p_b denote the channel block length, the signal constellation and the power allocated for block b, respectively. We assume that the entries of $\boldsymbol{Z}_b$ are independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian random variables with zero mean and unit variance. The fading coefficients, $H_b \in \mathbb{C}$, $\forall\, b = 1, \ldots, B$, are assumed to be blockwise i.i.d. according to a circularly symmetric complex Gaussian distribution with zero mean and unit variance. We denote the fading random vector as $\boldsymbol{H} \triangleq [H_1, \ldots, H_B]$.

We study a case where imperfect CSI is the actual CSI plus AWGN. This model of noisy CSI comes from exploiting channel reciprocity [8], [9] for which the channel realisation is identical at both ends but the channel estimation noises are independent, i.e.

$$\text{CSIT} \quad \tilde{H}_b = H_b + \tilde{E}_b \tag{2}$$

$$\text{CSIR} \quad \hat{H}_b = H_b + \hat{E}_b \tag{3}$$

where $\tilde{E}_b$ and $\hat{E}_b$ are the CSIT and CSIR noises, respectively; they have zero mean and their variances are given by $\sigma^2_{\tilde{e}} = P^{-d_{\tilde{e}}}$ and $\sigma^2_{\hat{e}} = P^{-d_{\hat{e}}}$, respectively, where P is the average SNR. For a fixed fading realisation $H_b = h_b$, the estimates at both ends are independent. The imperfect CSIR model is widely used in a pilot-based channel estimation at the receiver for which the error variance is proportional to the reciprocal of the pilot SNR [10]. The same estimation technique can also be performed at the transmitter, i.e. by transmitting pilot symbols at the reverse link of a time-division duplex (TDD) system [8], [9]. We further incorporate the parameters $d_{\tilde{e}} > 0$ and $d_{\hat{e}} > 0$, denoting the channel estimation error diversities.

A message $m \in \{1, \ldots, 2^{BTR}\}$ is mapped into a codeword $\boldsymbol{x}(m)$ of the code $\mathcal{C}$ with rate $R \triangleq \frac{1}{BT} \log_2 |\mathcal{C}|$; we assume that R is a fixed positive constant. The codebook is drawn i.i.d. from a constellation $\mathcal{X}$. Herein we focus on Gaussian and discrete constellations. The constellation $\mathcal{X}$ is assumed to have unit energy, i.e. $\mathsf{E}\left[|X|^2\right] = 1$. We denote $n(b)$ as the number of fading blocks used for the power adaptation at block b. The power at block b, p_b, is adapted based on the imperfect CSIT vector

$$\tilde{\boldsymbol{H}}_{n(b)} = \left[\tilde{H}_1, \ldots, \tilde{H}_{n(b)}\right] \tag{4}$$

such that $p_b = p_b(\tilde{\boldsymbol{H}}_{n(b)})$. The power allocation vector is then defined as $\boldsymbol{p} \triangleq \left[p_1(\tilde{\boldsymbol{H}}_{n(1)}), \ldots, p_B(\tilde{\boldsymbol{H}}_{n(B)})\right]$. In general, $n(b)$ can be any integer-function of b. In this paper, we focus on the following cases.

1) *Full-CSIT* power allocation if $n(b) = B$ for all $b = 1, \ldots, B$. Imperfect fading estimates for the whole B blocks in a codeword are available at the transmitter *prior* to transmission.
2) *Causal-CSIT* power allocation if $n(b) = b - \tau_{\rm d}$ with a fixed delay $\tau_{\rm d} \geq 0$ for any $b = 1, \ldots, B$. This corresponds to CSIT being limited only to the past imperfect fading estimates due to the delay $\tau_{\rm d}$.
3) *Predictive-CSIT* power allocation if $n(b) = \min(B, b + \tau_{\rm f})$ with a fixed $\tau_{\rm f} \geq 0$ (indicating the number of predicted fading blocks) for any $b = 1, \ldots, B$. This corresponds to CSIT including past, current and a number of predicted future fading estimates.

For the above power allocation schemes, the corresponding long-term average power constraint is given by

$$\mathbb{E}\left[\frac{1}{B}\sum_{b=1}^{B} p_b(\tilde{\boldsymbol{H}}_{n(b)})\right] \leq P. \tag{5}$$

Nearest neighbour decoding is used to infer the transmitted message. Due to its optimality under perfect CSIR and its simplicity, this decoder is widely used in practice even when perfect CSIR is not available. With imperfect CSIR, the decoder treats the imperfect channel estimate as if it were perfect. It first computes the following metric for a given $\boldsymbol{y}$, imperfect CSIR, $\hat{\boldsymbol{H}} = \hat{\boldsymbol{h}} = \left[\hat{h}_1, \ldots, \hat{h}_B\right]$, and power level $\boldsymbol{p} = [p_1(\tilde{\boldsymbol{h}}_{n(1)}), \ldots, p_B(\tilde{\boldsymbol{h}}_{n(B)})]$

$$Q\left(\boldsymbol{y}, \hat{\boldsymbol{h}}, \boldsymbol{p}, \boldsymbol{x}(m)\right) \propto \exp\left(-\sum_{b=1}^{B} \|\boldsymbol{y}_b - \sqrt{p_b}\hat{h}_b \boldsymbol{x}_b(m)\|^2\right) \tag{6}$$

and then outputs

$$\hat{m} = \arg \max_{m \in \{1, \ldots, 2^{BTR}\}} Q\left(\boldsymbol{y}, \hat{\boldsymbol{h}}, \boldsymbol{p}, \boldsymbol{x}(m)\right). \tag{7}$$

III. Information-Theoretic Preliminaries

Due to the noisy fading estimates, the above nearest neighbour decoder is mismatched [7], [11]. The fundamental limit (in the limit for large T) of the channel for the inputs that have been generated i.i.d. (i.i.d. inputs) is the generalised outage probability [7], [12]

$$P_{\rm gout}(R) \triangleq \Pr\left\{I^{\rm gmi}(\boldsymbol{H}, \hat{\boldsymbol{H}}, \boldsymbol{p}) < R\right\} \tag{8}$$

where

$$I^{\rm gmi}(\boldsymbol{h}, \hat{\boldsymbol{h}}, \boldsymbol{p}) = \sup_{s \geq 0} \frac{1}{B} \sum_{b=1}^{B} I_b^{\rm gmi}(s, h_b, \hat{h}_b, p_b) \tag{9}$$

is the generalised mutual information (GMI) for fading $\boldsymbol{h}$, receiver estimate $\hat{\boldsymbol{h}}$ and power level $\boldsymbol{p}$, and

$$I_b^{\rm gmi}(s, h_b, \hat{h}_b, p_b) = \tag{10}$$
$$\mathbb{E}\left[\log_2 \frac{Q^s\left(Y, \hat{h}_b, p_b, X\right)}{\mathbb{E}\left[Q^s\left(Y, \hat{h}_b, p_b, X'\right)\middle| Y, h_b, \hat{h}_b, p_b\right]}\middle| h_b, \hat{h}_b, p_b\right].$$

This implies that whenever the fading and channel estimate is such that $I^{\rm gmi}(\boldsymbol{h}, \hat{\boldsymbol{h}}, \boldsymbol{p})$ is less than the data rate R, the probability of decoding error tends to one in the limit of large block length T. On the other hand, when $I^{\rm gmi}(\boldsymbol{h}, \hat{\boldsymbol{h}}, \boldsymbol{p})$ is greater than or equal to R, the probability of decoding error vanishes for increasing block length.

We are interested in characterising the behaviour of $P_{\rm gout}(R)$ at high SNR. One important figure of merit is the outage diversity or outage SNR exponent defined as

$$d \triangleq \lim_{P \to \infty} -\frac{\log P_{\rm gout}(R)}{\log P}. \tag{11}$$

Recent work [7] showed that with uniform power allocation the optimal diversity is the function of the perfect CSIR diversity and the quality of the imperfect CSIR as

$$d^{\rm u}_{\rm icsir} = \min(1, d_{\hat{e}}) \times d^{\rm u}_{\rm csir} \tag{12}$$

where superscript $^{\rm u}$ denotes the uniform power allocation, the subscript $_{\rm csir}$ denotes perfect CSIR. From [3], we have that

$$d^{\rm u}_{\rm csir} = \begin{cases} B, & \text{for Gaussian inputs} \\ d_B(R) \triangleq 1 + \left\lfloor B\left(1 - \frac{R}{M}\right)\right\rfloor, & \text{for discrete inputs} \end{cases} \tag{13}$$

where d_B is the Singleton bound [2], and where $M \triangleq$

$\log_2|\mathcal{X}|$. This result implies that if the channel estimation error is less than or equal to the inverse of the SNR, the perfect CSIR diversity is achievable. Otherwise, the imperfect CSIR diversity is smaller than the perfect CSIR diversity.

If CSIT is available, then the transmitter can adapt its transmission power to minimise the generalised outage probability. The idea is that in a very bad channel realisation, power can be saved and used when channel conditions improve. Reference [4] showed that if perfect CSI is available at both communication ends, then zero outage is possible, implying that delay-limited capacity is positive. References [5], [6] extend the results to perfect CSIR and mismatched CSIT setup. In this case, the SNR exponent is

$$d_{\text{icsit}}^{\text{p}} = d_{\text{csir}}^{\text{u}}\left(1 + d_{\text{csir}}^{\text{u}} d_{\tilde{e}}\right) \tag{14}$$

where superscript $^{\text{p}}$ denotes power control.

In practical scenarios, both CSIR and CSIT will be imperfect. It is therefore of practical interest to study mismatched CSI at both ends under a unified framework. In this work, we find the SNR exponents with imperfect CSI at both ends. In particular, the power allocation algorithm is given by the solution to the following optimisation problem

$$\begin{aligned} \text{minimize} \quad & P_{\text{gout}}(R) \\ \text{subject to} \quad & \mathbb{E}\left[\frac{1}{B}\sum_{b=1}^{B} p_b(\tilde{\boldsymbol{H}}_{n(b)})\right] \le P \\ & p_b(\tilde{\boldsymbol{H}}_{n(b)}) \ge 0, \quad b = 1,\ldots,B. \end{aligned} \tag{15}$$

Solving the above optimisation problem can be difficult in general. Given our CSIT model, the minimum-outage power allocation is difficult to find since $P_{\text{gout}}(R)$ depends on both actual channel and channel estimate. Nevertheless, we will see that despite this difficulty, studying the behaviour of the optimal solution at high SNR is possible. We will use the technique in [6] to derive the asymptotic power allocation that results in no loss in terms of outage exponent.

IV. Outage SNR Exponents

Depending on the type of CSIT (full, causal or predictive), we will have a different SNR exponent for the corresponding power allocation schemes. In particular, for full CSIT we have the following result.

Theorem 1 (Full CSIT): For full CSIT ($n(b) = B$), the outage SNR exponent with imperfect CSI at both ends is

$$d_{\text{icsi}}^{\text{p}} = \begin{cases} d_{\text{csir}}^{\text{u}} d_{\hat{e}}, & \text{if } d_{\hat{e}} \le 1 + d_{\text{csir}}^{\text{u}} d_{\tilde{e}} \\ d_{\text{csir}}^{\text{u}}\left(1 + d_{\text{csir}}^{\text{u}} d_{\tilde{e}}\right), & \text{if } d_{\hat{e}} > 1 + d_{\text{csir}}^{\text{u}} d_{\tilde{e}}. \end{cases} \tag{16}$$

This relationship holds for both Gaussian and discrete constellations.

Proof: See Appendix A for a sketch. ∎

The results highlight the trade-off on the resources spent for estimating the channel at both ends and the effectiveness of power control given a noisy CSIR. Power control is effective whenever the CSIR noise variance is much smaller than the CSIT noise variance, i.e. $d_{\hat{e}} > 1 + d_{\text{csir}}^{\text{u}} d_{\tilde{e}}$. For example, with Gaussian inputs $d_{\hat{e}}$ must be larger than $d_{\tilde{e}}$ approximately by a factor of B. On the other aspects, the condition of $d_{\hat{e}} > 1$ highlights the improvement made by power control over uniform power allocation. Power control removes the 1 in the expression $\min(1, d_{\hat{e}})$ (12). Outage events are dominated by CSIR with strong noise, i.e. for $d_{\hat{e}} \le 1 + d_{\text{csir}}^{\text{u}} d_{\tilde{e}}$. Otherwise, outage events are dominated by the mismatched CSIT.

Remark 1: CSIR has stronger influence in the outage diversity than CSIT. With perfect CSIR, a very bad CSIT has the same effect as having no power control. On the other hand, with perfect CSIT, a very bad CSIR results in the diversity approaching zero.

The result in Theorem 1 is consistent previous results. In particular, we recover the mismatched-CSIT perfect-CSIR outage exponent in [5], [6] by letting $d_{\hat{e}} \uparrow \infty$ (perfect CSIR), and the no-CSIT mismatched-CSIR [7] by letting $d_{\tilde{e}} \downarrow 0$.

Remark 2: With mismatched CSIR $\hat{H}_b$, having $d_{\tilde{e}} = \infty$ or $d_{\tilde{e}} = d_{\hat{e}}$ at the transmitter does not make any difference, as the diversity is given by $d_{\hat{e}} d_{\text{csir}}^{\text{u}}$ in both cases. Instead, with mismatched CSIT $\tilde{H}_b$, having $d_{\hat{e}} = \infty$ or $d_{\hat{e}} = d_{\tilde{e}}$ at the receiver results in different exponents. In the former case, the diversity is given by $d_{\text{csir}}^{\text{u}}(1 + d_{\text{csir}}^{\text{u}} d_{\tilde{e}})$ whereas in the latter case the diversity is given by $d_{\tilde{e}} d_{\text{csir}}^{\text{u}}$.

The above discussion assumes full imperfect CSIT knowledge for each codeword. In some cases, CSIT may only be available causally (with some delay). In this case, we have the following results.

Theorem 2 (Causal CSIT): For causal CSIT with delay τ_{d} ($n(b) = b - \tau_{\text{d}}$), the outage diversity is given by

$$d_{\text{icsi}}^{\text{p}} = \begin{cases} d_{\hat{e}} B, & \text{if } d_{\hat{e}} \le 1 \\ B, & \text{if } d_{\hat{e}} > 1 \text{ and } B - \tau_{\text{d}} \le 0 \\ \sum_{i=1}^{B} a_i, & \text{if } d_{\hat{e}} > 1 \text{ and } B - \tau_{\text{d}} > 0. \end{cases} \tag{17}$$

where

$$a_i = \begin{cases} 1, & i = 1,\ldots,\tau_{\text{d}} \\ \min\left(d_{\hat{e}}, a_{i-1} + \min(a_{i-\tau_{\text{d}}}, d_{\tilde{e}})\right), & i = \tau_{\text{d}}+1,\ldots,B. \end{cases} \tag{18}$$

for Gaussian inputs and

$$d_{\text{icsi}}^{\text{p}} = \begin{cases} d_{\hat{e}} d_B(R), & \text{if } d_{\hat{e}} \le 1 \\ \sum_{i=1}^{d_B(R)} c_i, & \text{if } d_{\hat{e}} > 1 \end{cases} \tag{19}$$

where

$$c_i = \begin{cases} 1, & i = 1,\ldots,\tau_{\text{d}} \\ \min\left(d_{\hat{e}}, c_{i-1} + \min(c_{i-\tau_{\text{d}}}, d_{\tilde{e}})\right), & i = \tau_{\text{d}}+1,\ldots,d_B(R). \end{cases} \tag{20}$$

for discrete inputs.

Proof: See Appendix A for a sketch. ∎

There are two cases for which power control cannot increase the diversity. First, when the CSIR diversity is not reliable, i.e. $d_{\hat{e}} \le 1$. Second, when the delay of obtaining CSIT is long ($\tau_{\text{d}} \ge d_{\text{csir}}^{\text{u}}$). Power control does not improve diversity for example if the delay of obtaining CSIT is greater than B for Gaussian inputs or greater than $d_B(R)$ for discrete inputs. Hence, discrete inputs have a more stringent delay requirement in obtaining causal CSIT. To achieve the perfect CSI diversity with power adaptation, the CSIR reliability and the CSIT reliability need to be above a certain threshold. It is straightforward to check that the thresholds are given by

$$d_{\hat{e}} \geq a^*_{B-1} + a^*_{B-\tau_{\rm d}} \tag{21}$$

$$d_{\tilde{e}} \geq a^*_{B-\tau_{\rm d}} \tag{22}$$

for Gaussian inputs and

$$d_{\hat{e}} \geq c^*_{d_B(R)-1} + c^*_{d_B(R)-\tau_{\rm d}} \tag{23}$$

$$d_{\tilde{e}} \geq c^*_{d_B(R)-\tau_{\rm d}} \tag{24}$$

for discrete inputs, where the symbol * indicates the coefficient with $d_{\hat{e}} \uparrow \infty$ and $d_{\tilde{e}} \uparrow \infty$. It is observed that the requirement for decoding (CSIR) reliability is much higher than the CSIT reliability to achieve the perfect CSI diversity.

If the transmitter is able to predict some future fading coefficients, then we have following results.

Theorem 3 (Predictive CSIT): For predictive CSIT ($n(b) = \min(B, b+\tau_{\rm f})$) we have that for Gaussian inputs

$$d^{\rm p}_{\rm icsi} = \begin{cases} d_{\hat{e}}B, & \text{if } d_{\hat{e}} \leq 1 \\ \sum_{b=1}^{B} \min\left(d_{\hat{e}}, 1 + d_{\tilde{e}} \min(B, b+\tau_{\rm f})\right), & \text{if } d_{\hat{e}} > 1. \end{cases} \tag{25}$$

On the other hand, for discrete inputs, the outage diversity is given by

$$d^{\rm p}_{\rm icsi} = \begin{cases} d_{\hat{e}} d_B(R), \\ \qquad \text{if } d_{\hat{e}} \leq 1 \\ \sum_{b=\lceil \frac{BR}{M} \rceil}^{B} \\ \min\left(d_{\hat{e}}, 1 + d_{\tilde{e}}\left[\min(B, b+\tau_{\rm f}) - \lceil \frac{BR}{M} \rceil + 1\right]_+\right), \\ \qquad \text{if } d_{\hat{e}} > 1. \end{cases} \tag{26}$$

Proof: See Appendix A for a sketch. ■

We observe how the predictive-CSIT power control improves the outage diversity via a recursion in the power adaptation. Note that with

$$d_{\hat{e}} \geq 1 + d^{\rm u}_{\rm csir} d_{\tilde{e}} \tag{27}$$

we essentially obtain the same diversity as in the noiseless CSIR case. If the predictive time interval $\tau_{\rm f}$ satisfies

$$\tau_{\rm f} \geq \begin{cases} B-1, & \text{for Gaussian inputs} \\ d_B(R)-1, & \text{for discrete inputs,} \end{cases} \tag{28}$$

then the diversity obtained with predictive CSIT is same as the diversity obtained with full CSIT.

For $d_{\hat{e}} \leq 1$, we note that any power control with full, causal or predictive CSIT cannot improve the outage diversity with respect to that achieved with uniform power allocation. This corresponds to the case where the CSIR is too unreliable.

V. Conclusion

We have studied the effects of imperfect CSI on the performance of data transmission over block-fading channels. In particular, we derived the outage SNR exponent as a function of CSIR and CSIT noise variances, $\sigma^2_{\hat{e}} = P^{-d_{\hat{e}}}$ and $\sigma^2_{\tilde{e}} = P^{-d_{\tilde{e}}}$ where P is the average data transmission power. We showed that noisy CSIR has more detrimental effects on the SNR exponent than noisy CSIT. The results give insight into the design of pilot-assisted channel estimation in block-fading channels. If pilot symbols from both ends are sent with power P ($d_{\hat{e}} = d_{\tilde{e}} = 1$), then the CSI feedback and the power adaptation is useless in terms of the SNR exponent. On a positive note, if the pilot signalling can be done at a power level sufficiently higher than P, then one can reap significant benefits due to power adaptation across blocks.

Appendix A
Proof Sketch

Due to space limitations, only a proof sketch is provided. From Section II, we define a new random variable $\bar{H}_b \triangleq \frac{\sqrt{2}}{\sigma_{\tilde{e}}}\tilde{H}_b$. Given H_b, $\bar{H}_b$ is a complex Gaussian random variable with mean of $\frac{\sqrt{2}}{\sigma_{\tilde{e}}}H_b$ and a scaled identity variance. Let $\gamma_b = |h_b|^2$, $\hat{\gamma}_b = |\hat{h}_b|^2$, $\bar{\gamma}_b = |\bar{h}_b|^2$, $\tilde{\gamma}_b = |\tilde{h}_b|^2$ and $\hat{\xi}_b = |\hat{e}_b|^2$. We use the following change of variables to analyse the behaviour of the system for large P: $\alpha_b = -\frac{\log \gamma_b}{\log P}$, $\hat{\alpha}_b = -\frac{\log \hat{\gamma}_b}{\log P}$, $\bar{\alpha}_b = -\frac{\log \bar{\gamma}_b}{\log P}$, $\tilde{\alpha}_b = -\frac{\log \tilde{\gamma}_b}{\log P}$, $\hat{\theta}_b = -\frac{\log \hat{\xi}_b}{\log P}$. We define: $\boldsymbol{\gamma} \triangleq [\gamma_1, \ldots, \gamma_B]$, $\hat{\boldsymbol{\gamma}} \triangleq [\hat{\gamma}_1, \ldots, \hat{\gamma}_B]$, $\bar{\boldsymbol{\gamma}} \triangleq [\bar{\gamma}_1, \ldots, \bar{\gamma}_B]$, $\tilde{\boldsymbol{\gamma}} \triangleq [\tilde{\gamma}_1, \ldots, \tilde{\gamma}_B]$ and $\hat{\boldsymbol{\xi}} \triangleq [\hat{\xi}_1, \ldots, \hat{\xi}_B]$. $\boldsymbol{\alpha}$, $\hat{\boldsymbol{\alpha}}$, $\bar{\boldsymbol{\alpha}}$, $\tilde{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\theta}}$ follow accordingly.

There are three key steps of proving the results in Theorems 1, 2 and 3. The first one, is finding the asymptotic behaviour of the considered power allocation scheme. The second one, is evaluating the asymptotic probability density function of the fading for large P. The third step, is characterising the asymptotic outage set and finding infimum solutions that give the exponent according to Varadhan's lemma [13].

A. Asymptotic Power Allocation

$\tilde{H}_b$ can be decomposed into the magnitude $|\tilde{H}_b|$ (Rayleigh distributed) and phase $\phi_{\tilde{H}_b}$ (uniformly distributed over $[-\pi, \pi)$). Since the probability density function (p.d.f.) of the phase is $1/(2\pi)$, which is constant, it can be shown that the phase distribution does not affect the asymptotic power allocation at large SNR. Then, we express the power as a function of magnitude square, $p_b = p_b(\tilde{\boldsymbol{\gamma}}_{n(b)})$.

One can show that the power allocation with constraint $\mathbb{E}\left[p_b(\tilde{\boldsymbol{\gamma}}_{n(b)})\right] \leq BP$ for all $b = 1, \ldots, B$ results in the upper bound to the outage SNR exponent; note that this violates the constraint in (15). On the other hand, one can consider a suboptimal power allocation such that $\mathbb{E}\left[p_b(\tilde{\boldsymbol{\gamma}}_{n(b)})\right] \leq P$ to obtain the lower bound to the outage SNR exponent. Let $p_b(\tilde{\boldsymbol{\gamma}}_{n(b)}) = P^{\omega_b(\tilde{\boldsymbol{\gamma}}_{n(b)})}$, then the optimal power allocation with full CSIT satisfies

$$\int_{\tilde{\boldsymbol{\gamma}} \in \mathbb{R}^B_+} P^{\omega_b(\tilde{\boldsymbol{\gamma}})} f(\tilde{\boldsymbol{\gamma}}) d\tilde{\boldsymbol{\gamma}} \dot{\leq} P \tag{29}$$

where $\dot{\leq}$ is defined as the exponential inequality , i.e. $g(P) \dot{\leq} P^k$ if $\lim_{P\to\infty} \frac{\log g(P)}{\log P} \leq k$. $\doteq$ and $\dot{\geq}$ are similarly defined. Herein we disregard $n(b)$ from the subscript of the vector $\tilde{\boldsymbol{\gamma}}$ since we have $n(b) = B$ for full CSIT. The dot inequality above holds because B is finite and is not a function of P. Using the change of variables described above (from $\tilde{\gamma}_b$ to $\tilde{\alpha}_b$) and applying Varadhan's lemma [13] as in [5], [6] yields

$$\sup_{\tilde{\boldsymbol{\alpha}} \in \mathbb{R}_+^B} \left\{ \omega_b(\tilde{\boldsymbol{\alpha}}) - \sum_{b=1}^{B} \tilde{\alpha}_b \right\} \leq 1. \tag{30}$$

Since the generalised outage probability is a monotonously non-increasing function of the SNR at high SNR [7], the exponent of the optimum allocation satisfies

$$\omega_b(\tilde{\boldsymbol{\alpha}}) = 1 + \sum_{b=1}^{B} \tilde{\alpha}_b \tag{31}$$

asymptotically for large SNR. As for the vector $\bar{\boldsymbol{\alpha}}$, the above exponent can be expressed as

$$\omega_b(\bar{\boldsymbol{\alpha}}) = 1 + B d_{\tilde{e}} + \sum_{b=1}^{B} \bar{\alpha}_b. \tag{32}$$

We can follow the same argument above and the steps in [7], [14], and show that the exponent of the optimal allocation with causal CSIT is given by

$$\omega_b\left(\bar{\boldsymbol{\alpha}}_{n(b)}\right) = 1 + \min\left(\sum_{b'=1}^{b-\tau_{\mathrm{d}}} (\bar{\alpha}_{b'} + d_{\tilde{e}}), [d_{\hat{e}} - 1]_+ \right) \tag{33}$$

where $n(b) = b - \tau_{\mathrm{d}}$. On the other hand, using the steps in [7], [14], the exponent of the optimal allocation with predictive CSIT is given by

$$\omega_b\left(\bar{\boldsymbol{\alpha}}_{n(b)}\right) = 1 + \min\left(\sum_{b'=1}^{\min(B, b+\tau_{\mathrm{f}})} (\bar{\alpha}_{b'} + d_{\tilde{e}}), [d_{\hat{e}} - 1]_+ \right) \tag{34}$$

where $n(b) = \min(B, b + \tau_{\mathrm{f}})$. Note $\min(B, b + \tau_{\mathrm{f}})$ is used because any knowledge for blocks beyond the current codeword does not improve performance.

B. Generalised Outage Probability

Evaluating $P_{\mathrm{gout}}(R)$ requires the p.d.f. of the joint random variables that defines the outage set. Let $\mathcal{O}_{\mathcal{X}}$ be the outage set for input constellation $\mathcal{X}$ and a given power allocation. We have that the outage probability expression at high SNR is

$$P_{\mathrm{gout}}(R) \doteq \int_{\mathcal{O}_{\mathcal{X}}} f(\bar{\boldsymbol{\alpha}}|\boldsymbol{\alpha}) f(\boldsymbol{\alpha}) f(\hat{\boldsymbol{\theta}}) d\bar{\boldsymbol{\alpha}} d\boldsymbol{\alpha} d\hat{\boldsymbol{\theta}} \doteq P^{-d}. \tag{35}$$

The exponent d can be found by invoking Varadhan's lemma [13] as in [6], [7] yielding

$$d = \inf_{\mathcal{O}'_{\mathcal{X}}} \left\{ \sum_{b:\ 0 \leq \alpha < d_{\tilde{e}}, \bar{\alpha}_b = \alpha_b - d_{\tilde{e}}, \hat{\theta}_b \geq d_{\hat{e}}} \alpha_b + (\hat{\theta}_b - d_{\hat{e}}) + \sum_{b:\ \alpha \geq d_{\tilde{e}}, \bar{\alpha}_b \geq 0, \hat{\theta}_b \geq d_{\hat{e}}} \bar{\alpha}_b + \alpha_b + (\hat{\theta}_b - d_{\hat{e}}) \right\}. \tag{36}$$

where

$$\mathcal{O}'_{\mathcal{X}} = \mathcal{O}_{\mathcal{X}} \cap \left\{ \bigcap_b \{0 \leq \alpha < d_{\tilde{e}}, \bar{\alpha}_b = \alpha_b - d_{\tilde{e}}, \hat{\theta}_b \geq d_{\hat{e}}\} \cup \{\alpha \geq d_{\tilde{e}}, \bar{\alpha}_b \geq 0, \hat{\theta}_b \geq d_{\hat{e}}\} \right\}. \tag{37}$$

C. Bound to the Outage Set and Finding the Infimum Solution

Deriving the exact $\mathcal{O}_{\mathcal{X}}$ based on GMI expression in (9) and (10) for a given constellation $\mathcal{X}$ is not an easy task. The main difficulty is to find the optimising s taking into account all B fading blocks. We follow the steps in reference [7] to tackle the problem by finding the upper and lower bound to GMI with the power allocation described in the previous section. Note that we refer the resulting set from the lower bound to the GMI as $\underline{\mathcal{O}}_{\mathcal{X}}$ and the resulting set from the upper bound to the GMI as $\overline{\mathcal{O}}_{\mathcal{X}}$. It follows that

$$\overline{\mathcal{O}}_{\mathcal{X}} \subseteq \mathcal{O}_{\mathcal{X}} \subseteq \underline{\mathcal{O}}_{\mathcal{X}}. \tag{38}$$

The procedures are then continued by solving the infimum problem in (36) using these bounds. The outage exponent bounds derived using $\underline{\mathcal{O}}_{\mathcal{X}}$ and $\overline{\mathcal{O}}_{\mathcal{X}}$ are tight.

REFERENCES

[1] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994.

[2] R. Knopp and P. Humblet, "On coding for block fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 1, pp. 189–205, Jan. 2000.

[3] 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, no. 1, pp. 91–114, Jan. 2006.

[4] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468–1489, Jul. 1999.

[5] T. T. Kim and G. Caire, "Diversity gains of power control with noisy CSIT in MIMO channels," *IEEE Trans. Inf. Theory*, vol. 55, pp. 1618–1626, Apr. 2009.

[6] T. T. Kim, K. D. Nguyen, and A. Guillén i Fàbregas, "Coded modulation with mismatched CSIT over MIMO block-fading channels," *IEEE Trans. Inf. Theory*, vol. 56, pp. 5631–5640, Nov. 2010.

[7] A. T. Asyhari and A. Guillén i Fàbregas, "Nearest neighbour decoding in MIMO block-fading channels with imperfect CSIR," *Submitted to IEEE Trans. Inf. Theory*, Mar. 2010.

[8] R. Knopp and G. Caire, "Power control and beamforming for systems with multiple transmit and receive antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 638–648, Oct. 2002.

[9] M. Guillaud, D. T. M. Slock, and R. Knopp, "A practical method for wireless channel reciprocity exploitation through relative calibration," in *Proc. Eighth Int. Symp. Signal Process. and Its Applications*, Sydney, Australia, Aug. 2005.

[10] G. Taricco and E. Biglieri, "Space-time decoding with imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1874–1888, Jul. 2005.

[11] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai, "On information rates for mismatched decoders," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1953–1967, Nov. 1994.

[12] A. T. Asyhari and A. Guillén i Fàbregas, "MIMO block-fading channels with mismatched CSIR," in *Proc. Int. Symp. Inf. Theory and Its Applications*, Taichung, Taiwan, Oct. 2010.

[13] A. Dembo and O. Zeitouni, "Large deviations techniques and applications," in *Applications of Mathematics*, 2nd ed. New York: Springer-Verlag, 1998, vol. 38.

[14] K. D. Nguyen, N. Letzepis, A. Guillén i Fàbregas, and L. K. Rasmussen, "Outage diversity of MIMO block-fading channels with causal channel state information," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, Jun. 2010.