MIMO Block-Fading Channels with Mismatched CSIR

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

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

***Abstract*—This paper presents the outage analysis of multiple-input multiple-output (MIMO) block-fading channels with nearest neighbour decoding and mismatched channel state information at the receiver (CSIR). Based on mismatched decoding arguments, we demonstrate the achievability of the generalised outage probability, the probability that the generalised mutual information (GMI) is less than the target data rate, and show that this probability is the fundamental limit for independent and identically distributed (i.i.d.) codebooks. We then analyse the behaviour of the generalised outage probability at high signal-to-noise ratio (SNR) regime. For both Gaussian and discrete signal codebooks, we provide a simple characterisation of the mismatched CSIR SNR exponents and derive the necessary condition on the block length to achieve those SNR exponents.**

I. Introduction

The block-fading channel is a commonly used model to study reliable communication of delay-limited applications over slowly-varying fading [1]. In this model, the transmission of the encoded message spans a finite number of fading realisations. Within a block-fading period, the channel gains remain constant, varying from block to block according to some underlying distribution.

The assumption of perfect channel state information at the receiver (CSIR) has been widely used to investigate reliable communication over block-fading channels [1]–[4]. Malkämaki and Leib [3] showed that the outage probability, i.e., the probability that the mutual information is smaller than the target rate [1], is the fundamental limit of the channel. Furthermore, optimal codes should be maximum-distance separable (MDS) on a blockwise basis, i.e., achieving the Singleton bound on the blockwise Hamming distance of the code with equality [2]–[4]. Reference [4] proved that the outage probability as a function of the signal-to-noise ratio (SNR) decays linearly in a log-log scale with a slope given by the outage diversity or SNR exponent. This SNR exponent is shown to be the Singleton bound [4] for the codes constructed over discrete alphabets, showing the optimal trade-off between the diversity and the target rate. On the other hand, Gaussian codebooks of positive rates achieve the maximum diversity.

For perfect CSIR, nearest neighbour decoding yields minimal error probability [3], [5]. The perfect CSIR assumption, however, implies that the channel estimator provides the decoder with a perfectly accurate channel estimate. This assumption is too optimistic and very difficult to guarantee in practice. The time-varying fading characteristics imply that the receiver is likely to be unable to keep track on the channel fluctuations exactly. This makes nearest neighbour decoding inherently suboptimal.

This paper studies the robustness of that nearest neighbour decoder for the transmission over multiple-input multiple-output (MIMO) block-fading channels when the receiver has access only to the noisy CSIR. We approach the problem by studying the generalised mutual information (GMI) [6], [7] as the achievable rate for nearest neighbour decoder with imperfect CSIR. In particular, using the GMI we establish the achievability of the generalised outage probability, the probability that the GMI is less than the target rate, in MIMO block-fading channels. Furthermore, we characterise the behaviour of this probability at high SNR and find that the SNR exponent is given by the simple formula

$$d_{\rm icsir} = \min(1, d_e) \times d_{\rm csir} \tag{1}$$

where $d_{\rm csir}$ is the perfect CSIR outage exponent and d_e is the channel estimation error diversity. Interestingly, this relationship holds for both Gaussian-input and discrete-input codebooks. This formula is valid for MIMO channels with any numbers of transmit and receive antennas; thus, it generalises our preliminary result for single-antenna systems in [8]. Furthermore, we prove that the SNR exponent, $d_{\rm icsir}$, is achieved by Gaussian codes with finite length as long as the length is larger than a threshold, which depends on the fading distribution. On the other hand, discrete input codes only achieve $d_{\rm icsir}$ with block length growing with the logarithm of SNR. These results validate the use of perfect CSIR code designs based on Gaussian and discrete inputs, and provide guidelines for designing reliable channel estimators that achieve the optimal SNR exponents. Furthermore, the results are applicable to a wide-range of fading distributions including Rayleigh, Rician, Nakagami-m, Nakagami-q and Weibull fading models.

II. System Model

Consider transmission over a MIMO block-fading channel with $n_{\rm r}$ receive antennas, $n_{\rm t}$ transmit antennas and B fading blocks corrupted by additive white Gaussian noise (AWGN)

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

and affected by a fading matrix, $\boldsymbol{H}_b \in \mathbb{C}^{n_r \times n_t}, b = 1, \ldots, B$. The input-output relationship of the channel is

$$\boldsymbol{Y}_b = \sqrt{\frac{\mathsf{snr}}{n_t}} \boldsymbol{H}_b \boldsymbol{X}_b + \boldsymbol{Z}_b, \quad b = 1, \ldots, B \tag{2}$$

where $\boldsymbol{Y}_b, \boldsymbol{Z}_b \in \mathbb{C}^{n_r \times L}$ and $\boldsymbol{X}_b \in \mathbb{C}^{n_t \times L}$ are the received, the noise and the transmitted signal matrices corresponding to block b, respectively; L denotes the block length. We assume that the entries of $\boldsymbol{Z}_b$ are independent and identically distributed (i.i.d.) zero-mean unit-variance complex circularly symmetric Gaussian random variables.

At the transmitting end, we consider coding schemes given by $\mathcal{M} \subset \mathbb{C}^{n_t \times N}$ of a fixed rate R and length $N = BL$, whose codewords are defined as $\boldsymbol{X} = \left[\boldsymbol{X}_1, \cdots, \boldsymbol{X}_B\right] \in \mathcal{X}^{n_t \times BL}$, where $\mathcal{X}$ denote the signal constellation. The codewords are assumed to satisfy the average input power constraint $\frac{1}{BL}\mathbb{E}[\|\boldsymbol{X}\|^2] \leq n_t$.

Each entry of the channel matrix, $h_{b,r,t}$, is assumed to have probability density function (p.d.f) given by [9]

$$p_{h_{b,r,t}}(h) = w_0 |h|^{\tau} e^{-w_1 |h - w_2|^{\varphi}} \tag{3}$$

where $w_0 > 0$, $\tau \in \mathbb{R}$, $w_1 > 0$, $w_2 \in \mathbb{C}$ and $\varphi > 1$ are constants (finite and SNR independent). This model subsumes a number of widely used fading distributions as tabulated in Table I. For Rayleigh and Rician fading channels, the above p.d.f. represents the p.d.f. of the complex random Gaussian variable. For Nakagami-m, Weibull and Nakagami-q fading channels, the above p.d.f. represents the distribution of the fading magnitude; uniformly distributed phase over $[0, 2\pi)$ is assumed. Note that Ω denotes the variance except for Weibull, where its variance is given by $\Omega^2\Gamma(1 + 2/\eta)^2$. Furthermore, we assume that the average of fading gain is normalised to one, i.e. $\mathbb{E}\left[\mathrm{tr}(\boldsymbol{H}^\dagger \boldsymbol{H})\right] = n_t n_r$.

At the receiving end, we employ channel estimators that yield accurate yet imperfect channel estimate, i.e.

$$\hat{\boldsymbol{H}}_b = \boldsymbol{H}_b + \boldsymbol{E}_b, \quad b = 1, \ldots, B \tag{4}$$

where $\hat{\boldsymbol{H}}_b \in \mathbb{C}^{n_r \times n_t}$ and $\boldsymbol{E}_b \in \mathbb{C}^{n_r \times n_t}$ are the imperfect channel estimate and the channel estimation error, respectively. The entries of $\boldsymbol{E}_b$ are assumed to follow i.i.d. Gaussian random variables with zero mean and variance σ_e^2, where

$$\sigma_e^2 = \mathsf{snr}^{-d_e}. \tag{5}$$

Thus, we have assumed a family of channel estimators for which the CSIR noise variance is decreasing with SNR. This model is widely used in space-time decoding with imperfect channel estimation at the receiver [10], [11] for which the error variance is proportional to the reciprocal of pilot SNR.

To decode the received codewords, we consider a nearest neighbour decoder, which uses the imperfect channel estimate as if it were perfect. Assuming a memoryless channel, it performs decoding by calculating the following metric

$$Q_{\boldsymbol{Y}|\boldsymbol{X},\hat{\boldsymbol{H}}}(\boldsymbol{y}|\boldsymbol{x}, \hat{\boldsymbol{H}}_b) = \frac{1}{\pi^{n_t}} e^{-\|\boldsymbol{y} - \hat{\boldsymbol{H}}_b \sqrt{\mathsf{snr}}\boldsymbol{x}\|^2} \tag{6}$$

[2]Note that $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ is the Gamma function

TABLE I
P.D.F. FOR DIFFERENT FADING DISTRIBUTION

Fading type	w_0	τ	w_1	w_2	φ
Rayleigh	$\frac{1}{\pi\Omega}$	0	$\frac{1}{\Omega}$	0	2
Rician	$\frac{1}{\pi\Omega}$	0	$\frac{1}{\Omega}$	μ	2
Nakagami-m	$\frac{m^m}{m\Omega^m\Gamma(m)}$	$2m-2$	$\frac{m}{\Omega}$	0	2
Weibull	$\frac{\eta\Omega^{-\eta}}{2\pi}$	$\eta - 2$	$\frac{1}{\Omega^\eta}$	0	η
Nakagami-q	See footnote[2]				

for each channel use. The decision is made at the end of BL channel uses for a single codeword.

Define $\alpha_{b,r,t} \triangleq -\frac{\log|h_{b,r,t}|^2}{\log \mathsf{snr}}$, $\hat{\alpha}_{b,r,t} \triangleq -\frac{\log|\hat{h}_{b,r,t}|^2}{\log \mathsf{snr}}$ and $\theta_{b,r,t} \triangleq -\frac{\log|e_{b,r,t}|^2}{\log \mathsf{snr}}$. Then, $\boldsymbol{\alpha}_b$, $\hat{\boldsymbol{\alpha}}_b$ and $\boldsymbol{\theta}_b$ are $n_r \times n_t$ matrices whose element at row r and column t is given by $\alpha_{b,r,t}$, $\hat{\alpha}_{b,r,t}$ and $\theta_{b,r,t}$, respectively, for all $r = 1, \ldots, n_r$ and $t = 1, \ldots, n_t$. This change of random variables is used to analyse the p.d.f. at high SNR regime.

III. Outage and Generalised Outage Probability

This section reviews the fundamental concepts of outage [3], [4] and introduces the concept of generalised outage in MIMO block-fading channels.

Definition 1: The information outage probability, $P_{\mathrm{out}}(R)$, is defined by the probability that the mutual information is less than the target rate, R, i.e. [3], [4]

$$P_{\mathrm{out}}(R) \triangleq \Pr\{I(\boldsymbol{H}) < R\} \tag{7}$$

where

$$I(\boldsymbol{H}) = \frac{1}{B} \sum_{b=1}^{B} I^{\mathrm{awgn}}\left(\sqrt{\frac{\mathsf{snr}}{n_t}} \boldsymbol{H}_b\right) \tag{8}$$

and $I^{\mathrm{awgn}}\left(\sqrt{\frac{\mathsf{snr}}{n_t}}\boldsymbol{H}_b\right)$ is the mutual information (in bits per channel use) for MIMO AWGN channel with a channel matrix $\sqrt{\frac{\mathsf{snr}}{n_t}}\boldsymbol{H}_b$.

Reference [3] shows that with a perfectly known CSIR, the fundamental limit of block fading channels is given by this $P_{\mathrm{out}}(R)$. As the block length grows to infinity, the smallest error probability that can be achieved is the outage probability.

When the CSIR is limited to a noisy channel estimate $\hat{\boldsymbol{H}}$ and the receiver supplies this estimate to the decoder, then the decoder is mismatched. For a fixed channel and channel estimation error, the error probability of the ensemble of random codes can be upper-bounded as [7], [12]

$$\bar{P}_e(\hat{\boldsymbol{H}}) \leq 2^{-N E_r^Q(R, \hat{\boldsymbol{H}})} \tag{9}$$

[2]Nakagami$-q$ p.d.f. is given by $\frac{1+q^2}{2\pi q\Omega} I_0\left(\frac{(1-q^4)|h|^2}{4q^2\Omega}\right) e^{-\frac{(1+q^2)^2}{4q^2\Omega}|h|^2}$. Notice that the zero-order modified Bessel function of the first kind, $I_0(x)$, can be bounded by $1 \leq I_0(x) \leq e^x$, for $x \geq 0$ [9].

where $E_r^Q(R, \hat{\boldsymbol{H}})$ is the mismatched decoding error exponent [7], [12]. Error exponent analysis [12] shows that $E_r^Q(R, \hat{\boldsymbol{H}})$ is positive whenever we have $R \leq I^{\rm gmi}(\hat{\boldsymbol{H}}) - \epsilon$, and zero otherwise, proving the achievability of $I^{\rm gmi}(\hat{\boldsymbol{H}})$ for a fixed channel and channel estimate realisations, where

$$I^{\rm gmi}(\hat{\boldsymbol{H}}) = \sup_{s>0} \frac{1}{B} \sum_{b=1}^{B} I_b^{\rm gmi}\left(\sqrt{\frac{\mathsf{snr}}{n_{\rm t}}}\hat{\boldsymbol{H}}_b, s\right) \tag{10}$$

and

$$I_b^{\rm gmi}\left(\sqrt{\frac{\mathsf{snr}}{n_{\rm t}}}\hat{\boldsymbol{H}}_b, s\right) = \mathbb{E}\left[\log_2 \frac{Q^s_{\boldsymbol{Y}|\boldsymbol{X},\hat{\boldsymbol{H}}}(\boldsymbol{Y}|\boldsymbol{X}, \hat{\boldsymbol{H}}_b)}{\mathbb{E}\left[Q^s_{\boldsymbol{Y}|\boldsymbol{X},\hat{\boldsymbol{H}}}(\boldsymbol{Y}|\boldsymbol{X}', \hat{\boldsymbol{H}}_b)\Big|\boldsymbol{Y}\right]}\right]. \tag{11}$$

Definition 2: The generalised outage probability is defined as the probability that the GMI is less than the data rate, i.e.

$$P_{\rm gout}(R) \triangleq \Pr\{I^{\rm gmi}(\hat{\boldsymbol{H}}) < R\}. \tag{12}$$

The information-theoretic meaning of $P_{\rm gout}(R)$ can be explained as follows. Using (9), the error probability averaged over all the fading and channel estimation error is upper-bounded as

$$\bar{P}_e \leq \mathbb{E}\left[2^{-NE_r^Q(R,\hat{\boldsymbol{H}})}\right]. \tag{13}$$

In the limit of large N, the above expression becomes

$$\bar{P}_e \leq \Pr\{I^{\rm gmi}(\hat{\boldsymbol{H}}) < R\} \triangleq P_{\rm gout}(R). \tag{14}$$

This analysis shows the achievability of $P_{\rm gout}(R)$ in MIMO block-fading channels. Unfortunately, general tight converse results are not available for mismatched decoding [6] which implies that there may exists codes whose error probability for large block length is lower than $P_{\rm gout}(R)$. However, as shown in [13]–[15], i.i.d. codebooks have a GMI converse, i.e., no rate larger than the GMI can be transmitted with vanishing error probability for i.i.d. codebooks.

Proposition 1: In MIMO block-fading channels, the lower bound of the error probability for i.i.d. codebooks with sufficiently large block length is given by

$$\bar{P}_e \geq P_{\rm gout}(R) \geq P_{\rm out}(R). \tag{15}$$

Proof: The inequality of $\bar{P}_e \geq P_{\rm gout}(R)$ comes out from the GMI converse in [15] for i.i.d. codebooks. The inequality of $P_{\rm gout}(R) \geq P_{\rm out}(R)$ is due to the data-processing inequality for error exponents, i.e., $E_r^Q(R, \hat{\boldsymbol{H}}) \leq E_r(R, \boldsymbol{H})$ [7], [12]. ■

Remark 1: The second inequality in the Proposition 1 is general for any decoding metric and any codebooks generation (not only restricted to i.i.d. codebooks) since the inequality is due to the data processing inequality for error exponents.

IV. MISMATCHED CSIR SNR EXPONENTS

We now characterise the behaviour of the generalised outage probability at high SNR regime with nearest neighbour decoding. The parameter of interest is captured by the SNR exponents defined as follows.

Definition 3: Define the perfect and imperfect CSIR SNR exponents as

$$d_{\rm csir} \triangleq \lim_{\mathsf{snr}\to\infty} -\frac{\log P_{\rm out}(R)}{\log \mathsf{snr}}, \tag{16}$$

$$d_{\rm icsir} \triangleq \lim_{\mathsf{snr}\to\infty} -\frac{\log P_{\rm gout}(R)}{\log \mathsf{snr}}, \tag{17}$$

where both refer to the slope of the outage (generalised outage) probability curve on a log-log scale for large SNR.

Proposition 1 suggests that $d_{\rm icsir} \leq d_{\rm csir}$. We present our main results in the following theorem, summarising the exact relationship between $d_{\rm icsir}$ and $d_{\rm csir}$ and how such exponent can be achieved using random codes of a given length L.

Theorem 1: Consider MIMO block-fading channels with the channel, the fading and the imperfect CSIR described by (2), (3) and (4), respectively. The generalised outage SNR exponent using nearest neighbour decoding (6) with imperfect CSIR is given by

$$d_{\rm icsir} = \min(1, d_e) \times d_{\rm csir} \tag{18}$$

where

$$d_{\rm csir} = \begin{cases} \left(1 + \frac{\tau}{2}\right) B n_{\rm t} n_{\rm r} & \text{for Gaussian inputs} \\ \left(1 + \frac{\tau}{2}\right) d_B(R) & \text{for discrete inputs,} \end{cases} \tag{19}$$

is the perfect CSIR outage SNR exponent and

$$d_B(R) = n_{\rm r}\left(1 + \left\lfloor B\left(n_{\rm t} - \frac{R}{\log_2|\mathcal{X}|}\right)\right\rfloor\right) \tag{20}$$

is the Singleton bound.

Furthermore, $d_{\rm icsir}$ is achievable using Gaussian random codes with block length such that

$$L \geq \begin{cases} n_{\rm t} + n_{\rm r} - 1 & \text{for } \tau = 0 \\ \left\lceil \left(1 + \frac{\tau}{2}\right) n_{\rm t} n_{\rm r} \right\rceil & \text{otherwise.} \end{cases} \tag{21}$$

On the other hand, random coded modulation schemes are only able to achieve $d_{\rm icsir}$ up to discontinuous points of $d_B(R)$ with block length growing as $L(\mathsf{snr}) = \omega \log \mathsf{snr}$ and ω satisfies

$$\omega \geq \frac{1}{\log|\mathcal{X}|} \times \left(\frac{\min(1, d_e) \times \left(1 + \frac{\tau}{2}\right) n_{\rm r}}{1 + \left\lfloor \frac{BR}{\log_2|\mathcal{X}|} \right\rfloor - \frac{BR}{\log_2|\mathcal{X}|}}\right). \tag{22}$$

Proof: (sketch) For the converse proof, we use the generalised outage probability for long i.i.d. codebooks (Proposition 1). Firstly, the generalised outage probability is derived from the GMI expression in (10) for both Gaussian and discrete inputs. Secondly, the asymptotic behaviour in the limit of large SNR is characterised resulting in a generalised outage set. Thirdly, large deviation analysis is used to derive the SNR exponents.

Using bounding techniques, we are able to express the generalised outage set in terms of the entries of $\boldsymbol{\alpha}_b$ and $\boldsymbol{\theta}_b$. The GMI upper bound is found by scaling the positive terms in the GMI expression appropriately. The GMI lower bound is found by finding an appropriate suboptimal s that is close to optimal for large SNR only. Let $\mathcal{O}$ be the large-SNR outage set expressed in terms of $\alpha_{b,r,t}$ and $\theta_{b,r,t}$ that may represent the outer (due to the GMI lower bound) and the inner (due to GMI upper bound) bounds to the generalised outage set. $\mathcal{O}$ depends on the required data rate R and the input probability distribution $P_X(x)$. Since $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$ are independent, upper and lower bounds on $P_{\text{gout}}(R)$ are obtained by the following integral

$$\int_{\mathcal{O}} p_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) p_{\boldsymbol{\theta}}(\boldsymbol{\theta}) d\boldsymbol{\alpha} d\boldsymbol{\theta}. \tag{23}$$

Then, by inserting (23) to (17) and applying Varadhan's lemma [16], it can be shown that for the fading models in (3), the upper and the lower bounds to d_{icsir} take the following form

$$\inf_{\mathcal{O} \cap \{\boldsymbol{\alpha} \succeq 0, \boldsymbol{\theta} \succeq d_e\}} \left\{ \left(1 + \frac{\tau}{2}\right) \sum_{b=1}^{B} \sum_{r=1}^{n_{\text{r}}} \sum_{t=1}^{n_{\text{t}}} (\alpha_{b,r,t} + \theta_{b,r,t} - d_e) \right\} \tag{24}$$

where upper/lower bounds depends on whether $\mathcal{O}$ represents the inner/outer bounds of the generalised outage set. It seems that from (24), the SNR exponent d_{icsir} does not depend on the phases of $\boldsymbol{H}$ and $\boldsymbol{E}$. However, it can be shown that these phases affect the generalised outage set $\mathcal{O}$. These details are omitted due to space limitation.

Achievability results are proven using (9). By expansion, it can be shown that for Gaussian inputs

$$\lim_{\mathsf{snr} \to \infty} E_r^Q(R, \hat{\boldsymbol{H}}) \tag{25}$$

$$\geq \sup_{0 \leq \rho \leq 1} \frac{\rho}{B} \sum_{b=1}^{B} \log_2 \left(e^{-1} \det \left(\boldsymbol{I}_{n_{\text{r}}} + \hat{s} \frac{\mathsf{snr}}{n_{\text{t}}} \hat{\boldsymbol{H}}_b \hat{\boldsymbol{H}}_b^{\dagger} \right) \right) - \rho R$$

$$\geq \sup_{0 \leq \rho \leq 1} \frac{\rho}{B} \sum_{b=1}^{B} \log_2 \left(e^{-1} \left(1 + \hat{s} \frac{\mathsf{snr}}{n_{\text{t}}} \|\hat{\boldsymbol{H}}_b\|^2 \right) \right) - \rho R \tag{26}$$

where the optimal s for large SNR (obtained from the first order derivative) is given by

$$\hat{s} \doteq \frac{1}{n_{\text{r}} + \mathsf{snr} \sum_{b=1}^{B} \|\boldsymbol{E}_b\|^2} \tag{27}$$

and $\doteq$ is defined similarly in [17]. Note that the inequality in (26) is found by expanding the determinant in (25) in terms of its eigenvalues and removing some positive terms. This transformation is to facilitate the use of the joint probability of the entries of $\hat{\boldsymbol{H}}_b$ to evaluate the generalised outage probability. We first use $\|\hat{\boldsymbol{H}}_b\|^2$ in (26) and find the relationship with $\|\boldsymbol{H}_b\|^2$ and $\|\boldsymbol{E}_b\|^2$. Then, we perform asymptotic analysis on (26) and use change of random variables $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$. We next evaluate the expectation in (13) by integral over the joint probability of $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$. Then, by solving this integral and applying the Varadhan's lemma [16], it is possible to show the second inequality in (21) for any channel parameter τ.

Using similar steps as in Gaussian inputs, we can analyse the achievability of random codes constructed over discrete alphabets. The discrete signal random codes are shown to be able to achieve the d_{icsir} provided that block length growth as $L(\mathsf{snr}) = \omega \log \mathsf{snr}$ where $\omega \to \infty$. This is so because $E_r^Q(R, \hat{\boldsymbol{H}}_b)$ in (9) for discrete input is upper-bounded by an

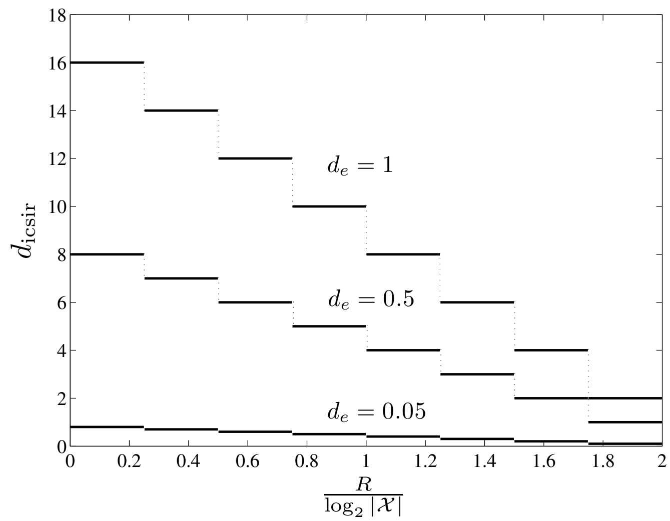

Fig. 1. d_{icsir} for discrete-input MIMO block-fading channel, $\tau = 0$, $B = 4$, $n_{\text{t}} = 2$ and $n_{\text{r}} = 2$.

SNR-independent quantity related to the cardinality of the signal set, $\log_2 |\mathcal{X}|$, and the number of transmit antennas, n_{t}; for Gaussian inputs $E_r^Q(R, \hat{\boldsymbol{H}}_b)$ grows as $\log \mathsf{snr}$.

Inequality in (26) leads to a looser achievability bound for Gaussian inputs. A tighter bound is obtained using (25) and transforming the determinant over $\hat{\boldsymbol{H}}_b \hat{\boldsymbol{H}}_b^{\dagger}$ and $\|\boldsymbol{E}_b\|^2$ as a function of their eigenvalues. We are no longer able to express the $E_r^Q(R, \hat{\boldsymbol{H}})$ lower bound in terms of $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$. The key step here is using the joint probability of those eigenvalues by invoking a conditional p.d.f. of $\hat{\boldsymbol{H}}$ given $\boldsymbol{E}$, i.e. for a given $\boldsymbol{E}$, $\hat{\boldsymbol{H}}$ has the same distribution of $\boldsymbol{H}$ with the same variance and the mean shifted by $\boldsymbol{E}$. Following the techniques in [9], for $\tau \neq 0$, the shift of mean makes the eigenvalue transformation intractable because the random matrix $\boldsymbol{E}$ affects $\hat{h}$ in both linear and exponential terms in (3). However, for $\tau = 0$, the transformation is possible since the shift of mean only appears in the exponential term. This transformation leads to a tighter block length threshold as shown in the first inequality in (21). ∎

V. Numerical Results and Discussion

The optimal SNR exponents for any coding scheme are achieved when $d_e \geq 1$. Using the outage converse bound in [17] and Proposition 1, we show that the converse on the SNR exponent is strong since $P_{\text{gout}}(R)$ has the same exponential decay in SNR as $P_{\text{out}}(R)$.

The results emphasises the role of channel estimation error SNR exponent d_e for determining the generalised outage exponent. Even though the obtained CSIR is noisy, we are still able to achieve the perfect CSIR SNR exponents provided that $d_e \geq 1$. In the limit of large SNR, the error level is most likely to be much less than the reciprocal of the SNR level. This provides a more precise numerical quantification of the channel estimator reliability for high SNR than [18]. If $d_e < 1$, the achieved SNR exponent scales linearly with d_e and approaching zero for $d_e \downarrow 0$. Fig. 1 illustrates this effect in a discrete-input block-fading channel with $B = 4$, $n_\mathrm{t} = 2$, $n_\mathrm{r} = 2$ and $\tau = 0$ (Rayleigh, Rician and Nakagami-q fading).

The achievability results imply that with a reliable channel estimator, $d_e \geq 1$, we can construct random codes of Gaussian constellations with finite block length to achieve the full diversity offered by the channel as long as the block length is greater than the threshold in (21). The value of this threshold depends on the number of antennas, n_t and n_r, and the fading parameter τ, which characterises the fading type used for analysis. The finite length condition of (21) for $\tau = 0$ is same as those in [9], [17], where perfect CSIR was used in both literature. For $\tau \neq 0$, the condition in (21) is looser than the result in [9] since we use a looser bound in (26) and the probability densities of the entries of both fading and channel estimation error matrices. Tighter bound is possible using (25) and the probability densities of the eigenvalues of fading and channel estimation error matrices. However, this tighter bound is difficult to analyse due to the shift of mean in the linear term of the p.d.f. which makes the eigenvalue transformation intractable (see the proof of Theorem 1). Nevertheless, we conjecture that the condition of the block length in [9] to obtain the optimal diversity is still applicable in the mismatched CSIR scenario.

On the other hand, for a strictly finite block length random coded modulation schemes are unable to achieve the full diversity since the random coding exponent $E_r^Q(R, \hat{\boldsymbol{H}}_b)$ for discrete input is upper-bounded by an SNR-independent quantity related to the cardinality of the signal set, $\log_2 |\mathcal{X}|$, and n_t. Thus, we require that the block length grows with the logarithm of SNR at certain growth rate ω shown in (22). Consider fixed B, d_e, n_r, τ and $\mathcal{X}$. Then, given data rate R, a finite ω can be computed to yield the optimal diversity. Now, let K be a positive integer and $\frac{BR}{\log_2 |\mathcal{X}|} = K - \epsilon$, with $0 < \epsilon < 1$. Varying R results in changing ϵ if the rest of parameters are fixed. Then, we may write (22) to

$$\omega \geq \frac{1}{\log |\mathcal{X}|} \times \left(\frac{\min(1, d_e) \times \left(1 + \frac{\tau}{2}\right) n_\mathrm{r}}{\epsilon} \right). \tag{28}$$

In the limit of $\epsilon \to 0$, we require that $\omega \to \infty$ to achieve the optimal diversity. Thus, $\omega \to \infty$ is needed to ensure that the optimal diversity is achievable for all rates up to discontinuous points of $d_B(R)$. If ω is strictly finite, then we may achieve the optimal diversity for limited range of R.

We have analysed the SNR exponents for imperfect CSIR. In general, imperfect CSIR may be obtained via separate channel-data detection or joint channel-data detection. Separate channel-data detection is simpler and widely used in practice where we have training symbols (pilot) and an independent channel estimator. Joint channel-data detection has larger complexity although it is promising in terms of efficiency. For a specific channel estimator, the price of obtaining a reliable channel estimation can be explained as follows. Consider space-time coding with maximum-likelihood (ML) channel estimation [10]. Then, it can be shown that d_e is proportional to the pilot power. Larger pilot power implies

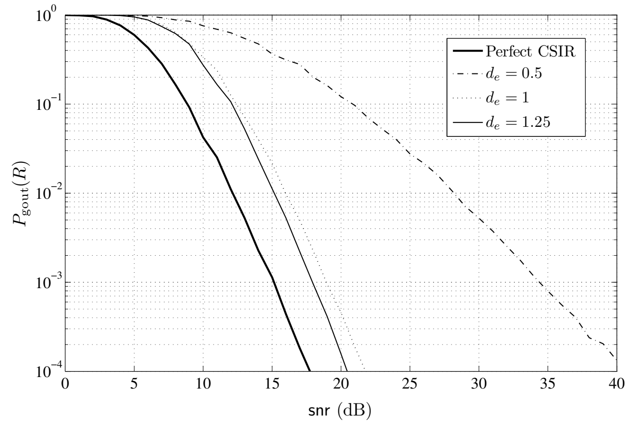

Fig. 2. $P_\mathrm{gout}(R)$ as a function of snr for Gaussian codebooks in Rayleigh fading, with $B = 2$, $n_\mathrm{t} = 2$, $n_\mathrm{r} = 1$ and $R = 2$ bits/channel use.

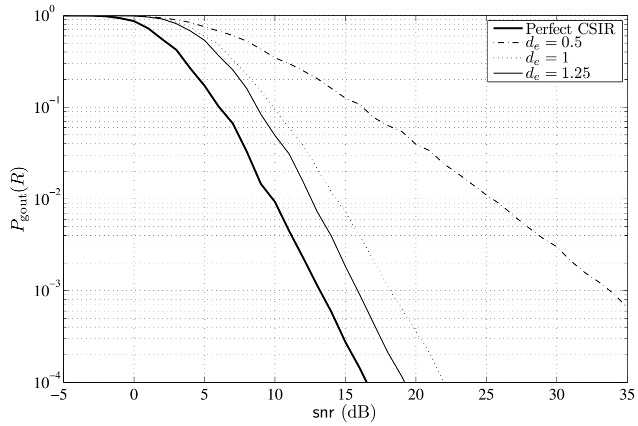

Fig. 3. $P_\mathrm{gout}(R)$ as a function of snr for BPSK codebooks in Rayleigh fading, with $B = 2$, $n_\mathrm{t} = 2$, $n_\mathrm{r} = 1$ and $R = 1$ bits/channel use.

larger d_e. Hence, for ML channel estimation, the price for obtaining the optimal SNR exponent is in the pilot power which does not contain any information data. Nevertheless, the limiting condition, $\min(1, d_e)$, ensures that the optimal diversity is achievable by designing pilot power so that $d_e = 1$.

Figs. 2 and 3 demonstrate the simulation results of the generalised outage probability for Gaussian and binary phase-shift keying (BPSK) inputs, respectively, over a MIMO Rayleigh block-fading channel with $B = 2$, $n_\mathrm{t} = 2$ and $n_\mathrm{r} = 1$. The data rates are specified as follows: $R = 2$ bits/channel use for Gaussian input, and $R = 1$ bits/channel use for BPSK input. For both Gaussian and BPSK inputs, the optimisation over s in (11) is performed numerically using the concavity property of the GMI in s, which can be proved easily using Hölder's inequality [19]. From the curves, the perfect CSIR slope is shown to be 4 and 3 for Gaussian and BPSK inputs, respectively. The numerical results are consistent with Theorem 1, i.e. the large-SNR mismatched CSIR slope becomes steeper with increasing d_e and is parallel to the perfect CSIR outage curve for $d_e \geq 1$. For $d_e \geq 1$, the increase in d_e does not increase the large-SNR slope as the curves become parallel for this range. However, the increase in d_e improves the outage performance as we may obtain same outage probability with lower SNR. For random coding of Gaussian and discrete inputs with large block length, this translates into the improvement of coding gain as d_e increases. Therefore, accounting for the imperfectness of CSIR, we should consider the goodness of the channel estimation scheme used to estimate the CSIR in the process of designing good codes.

Note that the results stated in Theorem 1 is valid for the fading model in (3). We notice that the only parameter that determines the SNR exponent is τ. Table I provides the values of τ for several fading distributions that are widely used in modelling the channel response of wireless communications.

VI. Conclusions

We have analysed the outage behaviour of nearest neighbour decoding with imperfect CSIR in MIMO block-fading channels. Due to the data-processing inequality for error exponents, the generalised outage probability is larger than the outage probability of the perfect CSIR case. We have further analysed the generalised outage probability in the large-SNR regime and we have derived the SNR exponents for both Gaussian and discrete inputs. We have shown that for both inputs, the SNR exponent is given by the perfect CSIR SNR exponent scaled by the minimum of channel estimation error diversity and one. Therefore, the robustness of nearest neighbour decoder can be maintained by having channel estimators that ensure the estimation error diversity equal to or larger than one. Furthermore, we provide a useful characterisation of the block length requirement for random codes to achieve the mismatched CSIR SNR exponents.

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] E. Malkamäki and H. Leib, "Coded diversity on block-fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 771–781, Mar. 1999.

[4] 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.

[5] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[6] 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.

[7] G. Kaplan and S. Shamai, "Information rates and error exponents of compound channels with application to antipodal signaling in a fading environment," *AEÜ Archiv für Elektronik und Übertragungstechnik*, vol. 47, no. 4, pp. 228–239, 1993.

[8] A. T. Asyhari and A. Guillén i Fàbregas, "Nearest neighbour decoding in block-fading channels with imperfect CSIR," in *Proc. IEEE Inf. Theory Workshop*, Taormina, Italy, Oct. 2009.

[9] L. Zhao, W. Mo, Y. Ma, and Z. Wang, "Diversity and multiplexing tradeoff in general fading channels," in *Proc. Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, Mar. 2006.

[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] E. Biglieri, *Coding for Wireless Channels*. New York, NY: Springer, 2005.

[12] A. Guillén i Fàbregas, A. Martinez, and G. Caire, "Bit-interleaved coded modulation," *Foundations and Trends on Communications and Information Theory*, vol. 5, no. 1-2, pp. 1–153, 2008.

[13] A. Ganti, A. Lapidoth, and I. E. Telatar, "Mismatched decoding revisited: general alphabets, channels with memory, and the wide-band limit," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2315–2328, Nov. 2000.

[14] A. Lapidoth, "Nearest neighbor decoding for additive non-Gaussian noise channels," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1520–1529, Sep. 1996.

[15] H. Weingarten, Y. Steinberg, and S. Shamai, "Gaussian codes and weighted nearest neighbor decoding in fading multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1665–1686, Aug. 2004.

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

[17] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[18] A. Lapidoth and S. Shamai, "Fading channels: how perfect need "perfect side information" be?" *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1118–1134, May 2002.

[19] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*. Cambridge University Press, 1934.