

# ON THE DIVERSITY ORDER OF VECTOR PERTURBATION PRECODING WITH IMPERFECT CHANNEL STATE INFORMATION

Joakim Jaldén, Johannes Maurer, and Gerald Matz

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology  
Gusshausstrasse 25/389, A-1040 Vienna, Austria (Europe)  
phone.: +43 1 58801 38958; fax: +43 1 58801 38999; E-mail: joakim.jalden@nt.tuwien.ac.at

## ABSTRACT

We consider vector perturbation precoding over a quasi-static MIMO channel under the assumption of imperfect channel state information (CSI). This is accomplished via a high SNR analysis, specifically targeting the overall system diversity order and the identification of typical errors. The effects of long-term and short-term power constraints, or power allocation policies, are investigated. Our results indicate that under realistic assumptions regarding the channel estimation error the system is mainly interference limited and as such, the particular power constraint does not significantly affect the asymptotic behavior of the error probability. This is in sharp contrast to the case of perfect CSI.

## 1. INTRODUCTION

In [1, 2], Hochwald, Peel and Swindlehurst introduced a vector perturbation technique for communication over multiple input-multiple output (MIMO) channels. The proposed scheme employs channel inversion at the transmitter preceded by vector perturbation of the data vector in order to reduce the transmit power. One of the key features of this approach is that the optimal receiver processing for the individual data streams (or data symbols) can be performed in a non-cooperative fashion, thus making the approach ideally suited for broadcast scenarios. In essence, the equivalent channel seen by any particular receiver becomes a non-fading additive white Gaussian noise (AWGN) channel and the receiver processing amounts to a power normalization and a modulo operation, followed by symbol decisions and decoding of the error correcting code (ECC).

In this work we study the effects of imperfect channel state information (CSI) on the performance of such a system in the high signal-to-noise ratio (SNR) regime. Specifically, we derive the diversity order of vector perturbation precoding for two types of power constraints (or power allocation policies) at the transmitter (cf. [3,4]): 1) a long-term constraint where the average transmit power is bounded; 2) a more stringent short-term constraint where the transmit power constraint is enforced for each channel realization. We consider only delay-constrained systems operating over quasi-static MIMO channel [5], where the ECC is not allowed to span several channel realizations. This is equivalent to the assumption made in [6] but differs from [1, 2] where the ergodic setting was considered. In fact, we shall only explicitly discuss the uncoded case although the results (regarding the system diversity) apply also when channel coding is used. An essential conclusion from our main result is that there is little to be gained by using a long-term power constraint for the case of imperfect CSI.

This work was supported by the STREP project MASCOT (IST-026905) within the Sixth Framework Programme of the European Commission.

The outline of the paper is as follows. The system model and power constraints are described in Section 2 and vector perturbation precoding is reviewed in Section 3. Our main contribution, a Theorem quantifying the diversity order of vector perturbation precoding with imperfect CSI, is presented and proved in 4. In Section 5 the results are illustrated by means of numerical simulations and conclusions are provided in Section 6.

## 2. SYSTEM MODEL AND POWER CONSTRAINTS

Consider the standard flat fading  $K \times M$  MIMO model given by [5]

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{w}, \quad (1)$$

where  $\mathbf{x} \in \mathbb{C}^M$  and  $\mathbf{r} \in \mathbb{C}^K$  denote the transmit and receive vector, respectively,  $\mathbf{H} \in \mathbb{C}^{K \times M}$  is the MIMO channel matrix, and  $\mathbf{w} \in \mathbb{C}^K$  is additive white Gaussian noise, i.e.,  $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$ . We assume  $K \geq M$  and i.i.d. Rayleigh fading, i.e.,  $\text{vec}(\mathbf{H}) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ . We will interpret (1) as system where a transmitter with  $M$  antennas communicates with  $K$  non-colocated single-antenna receivers, although this interpretation is no prerequisite for our analysis. It is furthermore assumed that (imperfect) CSI is available to the transmitter, given by

$$\hat{\mathbf{H}} = \mathbf{H} + \mathbf{E}. \quad (2)$$

Here,  $\mathbf{E}$  models the CSI mismatch whose nature and statistical characterization will be made precise in Section 4.1.

Two different power constraints are considered in what follows. The first one is an *average (long-term) power constraint* given by

$$\mathbb{E} \{ \|\mathbf{x}\|^2 \} = P. \quad (3)$$

This constraint allows for flexible channel-adaptive instantaneous power allocation. In contrast, with the much stricter *short-term power constraint*

$$\mathbb{E} \{ \|\mathbf{x}\|^2 \mid \hat{\mathbf{H}} \} = P, \quad (4)$$

the power budget has to be met for every channel realization. Note that under the short-term constraint the long-term power constraint is also satisfied. Hence, the average signal-to-noise ratio for both cases equals  $\rho \triangleq P/\sigma^2$ . We will in what follows focus on the high SNR behavior of the vector perturbation technique.

## 3. REVIEW OF VECTOR PERTURBATION

We next briefly summarize the main ideas of vector perturbation precoding (see also Fig. 1). The purpose of this section is mainly to introduce notation required in the subsequent analysis and we refer the reader to [2] for a more complete discussion of vector perturbation precoding.

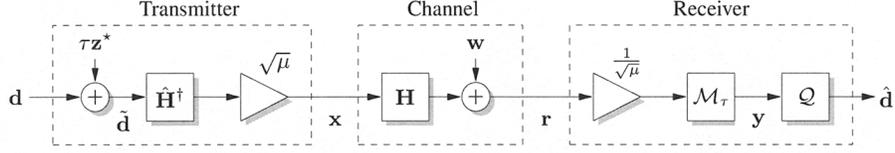


Fig. 1. Overview of transceiver chain including transmitter, channel, and receiver.

### 3.1. Transmit Processing

For a given data vector  $\mathbf{d} \in \mathcal{A}^K$ , where  $\mathcal{A}$  is a symbol alphabet, the transmit vector is formed according to (see [2] for further details)

$$\mathbf{x} = \sqrt{\mu} \hat{\mathbf{H}}^\dagger (\mathbf{d} + \tau \mathbf{z}^*), \quad (5)$$

where  $\mu$  is a normalization factor ensuring that the given power constraint is satisfied (see below),  $\hat{\mathbf{H}}^\dagger \triangleq \hat{\mathbf{H}}^H (\hat{\mathbf{H}} \hat{\mathbf{H}}^H)^{-1}$  denotes the right pseudoinverse of  $\hat{\mathbf{H}}$ ,  $\tau > 0$  is an appropriately chosen constant, and  $\mathbf{z} \in \mathbb{G}^K$  ( $\mathbb{G} = \mathbb{Z} + i\mathbb{Z}$ ) is a perturbation vector consisting of Gaussian integers. The perturbation vector is chosen according to

$$\mathbf{z}^* \triangleq \arg \min_{\mathbf{z} \in \mathbb{G}^K} \|\hat{\mathbf{H}}^\dagger (\mathbf{d} + \tau \mathbf{z})\|^2. \quad (6)$$

In what follows, we will refer to

$$\Gamma \triangleq \|\hat{\mathbf{H}}^\dagger (\mathbf{d} + \tau \mathbf{z}^*)\|^2 \quad (7)$$

as the *unnormalized instantaneous transmit power* (the instantaneous transmit power is given by  $\|\mathbf{x}\|^2$ ). This quantity will play a central role in our diversity analysis. We note that  $\mathbf{z}^*$  in (6) provides the smallest possible unnormalized instantaneous transmit power and may be efficiently obtained using the sphere decoder, at least for moderate values of  $K$ .

To satisfy the long-term power constraint (3), the scaling factor is chosen according to  $\mu = \gamma_L^{-1} P$  where (cf. (5))<sup>1</sup>

$$\gamma_L \triangleq \mathbb{E} \{\Gamma\} = \mathbb{E} \left\{ \|\hat{\mathbf{H}}^\dagger (\mathbf{d} + \tau \mathbf{z}^*)\|^2 \right\}. \quad (8)$$

Similarly, for the short-term constraint  $\mu = \gamma_S^{-1} P$  with

$$\gamma_S \triangleq \mathbb{E} \left\{ \Gamma \mid \hat{\mathbf{H}} \right\} = \mathbb{E} \left\{ \|\hat{\mathbf{H}}^\dagger (\mathbf{d} + \tau \mathbf{z}^*)\|^2 \mid \hat{\mathbf{H}} \right\}. \quad (9)$$

Note that  $\gamma_S$  depends on the channel realization and must be communicated somehow to the receivers. This potentially problematic side-effect of the short-term power constraint is well known (cf. [2]).

### 3.2. Receiver Processing

Assuming for a moment perfect CSI as in [2] ( $\hat{\mathbf{H}} = \mathbf{H}$ ), inserting (5) into (1) yields

$$\mathbf{r} = \sqrt{\mu} (\mathbf{d} + \tau \mathbf{z}^*) + \mathbf{w}. \quad (10)$$

Thus, after eliminating the scaling factor  $\mu$ , the receivers can easily detect the data by applying a modulo operation (denoted  $\mathcal{M}_\tau\{\cdot\}$ ) to remove  $\tau \mathbf{z}^*$ , followed by simple quantization with respect to  $\mathcal{A}$  (denoted  $\mathcal{Q}\{\cdot\}$ ):

$$\hat{\mathbf{d}} = \mathcal{Q} \left\{ \mathcal{M}_\tau \left\{ \frac{1}{\sqrt{\mu}} \mathbf{r} \right\} \right\}. \quad (11)$$

All operations in (11) are to be understood component-wise and do not require joint/cooperative receiver processing. We shall in the following assume that the receiver processing (11) is used also in the case of imperfect CSI where (10) does no longer hold.

<sup>1</sup>We note that in certain situations the expectation in (8) may not exist (cf. [2]). In the context of this paper, this only happens for  $M = K = 1$ , which we disregard in the following.

## 4. DIVERSITY WITH IMPERFECT CSI

### 4.1. Channel Uncertainty

The error matrix  $\mathbf{E}$  is assumed white, Gaussian and independent of  $\mathbf{H}$ , i.e.,  $\text{vec}(\mathbf{E}) \sim \mathcal{CN}(\mathbf{0}, \eta \mathbf{I})$  where  $\eta$  denotes the error variance. Since our analysis focuses on the high SNR regime, we model only the asymptotic behavior of  $\eta$ . We will see below that such a coarse model is indeed sufficient for characterizing the diversity order of vector perturbation precoding. In the following, we model CSI accuracy by

$$\eta \doteq \rho^{-\alpha} \quad (12)$$

where  $\alpha \geq 0$  is a model parameter. Here,  $g(\rho) \doteq h(\rho)$  is used to denote equality to the first order in the exponent in the sense of

$$\lim_{\rho \rightarrow \infty} \frac{\log g(\rho)}{\log \rho} = \lim_{\rho \rightarrow \infty} \frac{\log h(\rho)}{\log \rho}. \quad (13)$$

The symbols  $\dot{\leq}$  and  $\dot{\geq}$  are defined analogously. Two important special cases of (12) are as follows:

1.  $\alpha = 1$ : This models a reciprocal scenario where the channel is estimated at the transmitter using pilot signaling in the reverse link, assuming that the pilot SNR in the reverse link is proportional to the SNR in the forward link. In such a scenario, the channel estimation error decays inversely proportional to the SNR.
2.  $\alpha = 0$ : This models a scenario where the transmitter uses outdated CSI (e.g., due to delay in a system with CSI feedback). In such a scenario the channel accuracy will not improve with SNR but is rather determined by the channel coherence time.

The case of perfect CSI is obtained for  $\alpha \rightarrow \infty$ .

### 4.2. Main Result

A (symbol) error is declared whenever the output  $\hat{\mathbf{d}}$  of the receiver in (11) does not equal the transmitted data vector  $\mathbf{d}$ . The diversity order of the system is defined according to

$$d \triangleq - \lim_{\rho \rightarrow \infty} \frac{\ln \Pr\{\hat{\mathbf{d}} \neq \mathbf{d}\}}{\ln \rho}, \quad (14)$$

i.e. as the negative slope (in a log-log scale) of the probability of error with respect to SNR. The main contribution of this work, stated in the following Theorem, is the evaluation of the diversity order of vector perturbation precoding with imperfect CSI under both the long-term and short-term power constraints.

**Theorem 1** *The diversity order of vector perturbation precoding under the long-term power constraint is given by*

$$d = d_L = \alpha M, \quad (15)$$

*whereas under the short-term power constraint the diversity equals*

$$d = d_S = \min(\alpha M, M). \quad (16)$$

There are several important observations to be made from this result. For perfect CSI, it is known that the maximum diversity order under the short-term power constraint equals  $M$  [6] while the long-term power constraint leads to an exponential decay of error probability (i.e. infinite diversity). These results are recovered as special cases of Theorem 1 for  $\alpha \rightarrow \infty$ .

Imperfect CSI, however, is seen to reduce the diversity order significantly. Thus, Theorem 1 reveals that perfect CSI is a critical assumption in the sense that even slight CSI imperfections can cause the system to become interference limited in the high SNR regime. In fact, for the practically most relevant range  $\alpha \in [0, 1]$  the system becomes interference limited under both power constraints (this is further discussed in Section 4.4). Note that the effect of interference is more severe under the long-term power constraint. As an example, for  $\alpha = 1$  we have  $d_S = M$  as in the case of perfect CSI, whereas  $d_L = M$  as opposed to the infinite diversity with perfect CSI; in this case, there is also  $d_S = d_L$ , which actually holds for all  $\alpha \in [0, 1]$ . Hence, a somewhat unexpected conclusion from Theorem 1 is that within the range  $\alpha \in [0, 1]$  the diversity order of the system does not depend on which power constraint is used. This is true in particular for the two special cases discussed in Section 4.1. In the extreme case of  $\alpha = 0$ , there is  $d_S = d_L = 0$ , i.e., the error probability saturates in the high SNR regime; this is a consequence of the fact that here the CSI accuracy does not improve with increasing SNR.

### 4.3. Proof of Theorem 1

**Typical and Dominating Error Events.** We next establish the claims of Theorem 1 and provide a rigorous analysis of the limiting behavior of the system considered. We start by analyzing the input to quantizer (cf. (11) and Fig. 1),

$$\mathbf{y} \triangleq \mathcal{M}_\tau \left\{ \frac{1}{\sqrt{\mu}} \mathbf{r} \right\}.$$

Combining (1), (2) and (5) yield

$$\begin{aligned} \mathbf{y} &= \mathcal{M}_\tau \left\{ \mathbf{H}\hat{\mathbf{H}}^\dagger (\mathbf{d} + \tau \mathbf{z}^*) + \frac{1}{\sqrt{\mu}} \mathbf{w} \right\} \\ &= \mathcal{M}_\tau \left\{ (\hat{\mathbf{H}} - \mathbf{E})\hat{\mathbf{H}}^\dagger (\mathbf{d} + \tau \mathbf{z}^*) + \frac{1}{\sqrt{\mu}} \mathbf{w} \right\} \\ &= \mathcal{M}_\tau \left\{ \mathbf{d} + \tau \mathbf{z}^* - \mathbf{E}\hat{\mathbf{H}}^\dagger (\mathbf{d} + \tau \mathbf{z}^*) + \frac{1}{\sqrt{\mu}} \mathbf{w} \right\} \\ &= \mathcal{M}_\tau \left\{ \mathbf{d} - \mathbf{E}\hat{\mathbf{H}}^\dagger (\mathbf{d} + \tau \mathbf{z}^*) + \frac{1}{\sqrt{\mu}} \mathbf{w} \right\}. \end{aligned}$$

Thus, the useful part of the signal consists of  $\mathbf{d}$  while the noise-plus-interference term (before modulo operation) is given by

$$\mathbf{v} \triangleq -\mathbf{E}\hat{\mathbf{H}}^\dagger \tilde{\mathbf{d}} + \frac{1}{\sqrt{\mu}} \mathbf{w}, \quad (17)$$

where  $\tilde{\mathbf{d}} \triangleq (\mathbf{d} + \tau \mathbf{z}^*)$ . It is straightforward to show that conditioned on  $\hat{\mathbf{H}}$  and  $\tilde{\mathbf{d}}$ ,  $\mathbf{v}$  is i.i.d. zero mean Gaussian, i.e.

$$f(\mathbf{v} | \hat{\mathbf{H}}, \tilde{\mathbf{d}}) = \frac{1}{(\pi\nu)^K} \exp(-\nu^{-1} \|\mathbf{v}\|^2).$$

Here,  $\nu$  denotes the sum of interference power and noise variance,

$$\nu = \eta \Gamma + \sigma^2 \mu^{-1}, \quad (18)$$

where we used that  $\Gamma = \|\hat{\mathbf{H}}^\dagger \tilde{\mathbf{d}}\|^2$  (cf. (7)). Thus, the statistics of the noise-plus-interference term power is determined by the unnormalized instantaneous transmit power  $\Gamma$ . Note also that under the short-term power constraint,  $\mu$  depends on the channel realization.

The typical error events, in the sense of [5], are caused by particularly large realizations of  $\nu$ . Loosely speaking, for  $\nu \gg 1$  errors

are almost certain while for  $\nu \ll 1$  errors are rare. This is formalized by the following lemma, stated without the proof of which is tedious but straightforward (cf. [5, Exercise 3.3]).

**Lemma 1** For any  $\epsilon > 0$ , the diversity order  $d$  in (14) satisfies

$$-\lim_{\rho \rightarrow \infty} \frac{\ln \Pr\{\nu \geq \rho^{-\epsilon}\}}{\ln \rho} \leq d \leq -\lim_{\rho \rightarrow \infty} \frac{\ln \Pr\{\nu \geq \rho^\epsilon\}}{\ln \rho}. \quad (19)$$

It can furthermore be shown that the lower and upper bounds in (19) are continuous (and equal) at  $\epsilon = 0$ , and hence

$$d = -\lim_{\rho \rightarrow \infty} \frac{\ln \Pr\{\nu \geq 1\}}{\ln \rho}. \quad (20)$$

We do not provide more details here since this would obscure the main points of the remaining proof. In any case, (20) indicates that the diversity order is determined by the probability that the noise-plus-interference power exceeds a certain threshold (cf. [5, Chapter 3]). Next, note that (cf. (18))

$$\Pr\{\nu \geq 1\} \leq \Pr\{\eta\Gamma \geq \frac{1}{2}\} + \Pr\{\sigma^2 \mu^{-1} \geq \frac{1}{2}\} \quad (21a)$$

$$\Pr\{\nu \geq 1\} \geq \Pr\{\eta\Gamma \geq 1\} \quad (21b)$$

$$\Pr\{\nu \geq 1\} \geq \Pr\{\sigma^2 \mu^{-1} \geq 1\}, \quad (21c)$$

where (21a) is obtained by applying the union bound and (21b), (21c) follow from the nonnegativity of  $\sigma^2 \mu^{-1}$  and  $\eta\Gamma$ . By (20) and (21) it follows that

$$d = \min\{d_I, d_N\} \quad (22)$$

where

$$d_I \triangleq -\lim_{\rho \rightarrow \infty} \frac{\ln \Pr\{\eta\Gamma \geq \kappa\}}{\ln \rho} \quad (23)$$

and

$$d_N \triangleq -\lim_{\rho \rightarrow \infty} \frac{\ln \Pr\{\sigma^2 \mu^{-1} \geq \kappa\}}{\ln \rho}. \quad (24)$$

In the above,  $\kappa > 0$  is arbitrary (i.e. the actual value of  $\kappa$  does not affect the limit). Interpreting (23) as the probability of error due to (self) interference and (24) as the probability of error due to noise, (22) simply states that the diversity order is determined by the dominant source of receiver errors (interference or noise). It is also noteworthy that only (24) depends on the power constraint through  $\mu^{-1} = \gamma \rho^{-1}$ , whereas (23) is actually independent of the transmit power. Hence, increasing transmit power won't improve performance in cases where the system is interference limited.

**Interference Diversity — Upper Bound.** We first compute an upper bound for (23) via a lower bound on the unnormalized instantaneous transmit power  $\Gamma$ . To this end, define

$$c \triangleq \min_{a \in \mathcal{A}} |a|^2.$$

For simplicity, we restrict to the case  $c > 0$ , although a similar argument can be made without this assumption. By the definition of the pseudo-inverse it follows that  $\hat{\mathbf{h}}_i^H \hat{\mathbf{H}}^\dagger = \mathbf{e}_i^H$  where  $\hat{\mathbf{h}}_i^H$  denotes the  $i$ th row of  $\hat{\mathbf{H}}$  and  $\mathbf{e}_i$  is the  $i$ th natural basis vector. Thus, for any  $i = 1, \dots, K$ ,

$$c \leq |\tilde{d}_i|^2 = |\mathbf{e}_i^H \tilde{\mathbf{d}}|^2 = |\hat{\mathbf{h}}_i^H \hat{\mathbf{H}}^\dagger \tilde{\mathbf{d}}|^2 \leq \|\hat{\mathbf{h}}_i\|^2 \|\hat{\mathbf{H}}^\dagger \tilde{\mathbf{d}}\|^2$$

where  $\tilde{d}_i$  denotes the  $i$ th element of  $\tilde{\mathbf{d}}$ . It follows that

$$\Gamma = \|\hat{\mathbf{H}}^\dagger \tilde{\mathbf{d}}\|^2 \geq \frac{c}{\|\hat{\mathbf{h}}_i\|^2}. \quad (25)$$

According to (25), a short column vector of  $\hat{\mathbf{H}}$  is sufficient to make  $\Gamma$  large. Since any choice of  $i$  will work in the following, we simply set  $i = 1$  and obtain

$$\Pr\{\eta\Gamma \geq \kappa\} \geq \Pr\{\|\hat{\mathbf{h}}_1\|^2 \leq c\eta\kappa^{-1}\}, \quad (26)$$

Since  $\hat{\mathbf{h}}_1$  is i.i.d. Gaussian with variance  $1 + \eta$  (cf. (2)), we obtain

$$\Pr\{\|\hat{\mathbf{h}}_1\|^2 \leq c\eta\kappa^{-1}\} = \Pr\{\xi_M \leq c(1 + \eta)^{-1}\eta\kappa^{-1}\}.$$

Here,  $\xi_M$  is  $\chi$ -square distributed random variable with  $2M$  degrees of freedom that satisfies [5]

$$\Pr\{\xi_M \leq \rho^{-1}\} \doteq \rho^{-M}. \quad (27)$$

By using the asymptotic expression for  $\eta$  in It follows from (12) that

$$c(1 + \eta)^{-1}\eta\kappa^{-1} \doteq \rho^{-\alpha},$$

which in combination with (27) yields

$$\Pr\{\xi_M \leq c(1 + \eta)^{-1}\eta\kappa^{-1}\} \doteq \rho^{-\alpha M}.$$

In conjunction with (26) and (23), this implies

$$d_I \leq \alpha M. \quad (28)$$

**Interference Diversity — Lower Bound.** Inspired by [6], we consider  $\hat{\mathbf{H}}^\dagger$  as the generator matrix of a  $K$ -dimensional lattice in  $\mathbb{C}^M$ . The *covering radius* of this lattice is defined as

$$\zeta(\hat{\mathbf{H}}^\dagger) \triangleq \max_{\mathbf{a} \in \mathbb{C}^M} \min_{\mathbf{z} \in \mathbb{G}^M} \|\hat{\mathbf{H}}^\dagger(\mathbf{a} - \mathbf{z})\|.$$

Introducing the covering radius is useful since it provides the following convenient upper bound on  $\Gamma$ :

$$\Gamma = \|\hat{\mathbf{H}}^\dagger \hat{\mathbf{d}}\|^2 = \min_{\mathbf{z} \in \mathbb{G}^M} \|\hat{\mathbf{H}}^\dagger(\hat{\mathbf{d}} + \tau\mathbf{z})\|^2 \leq \tau^2 \zeta^2(\hat{\mathbf{H}}^\dagger). \quad (29)$$

Using a result of Banaszczyk [7, Theorem 2.2], it follows that

$$m_1(\hat{\mathbf{H}}^H) \zeta(\hat{\mathbf{H}}^\dagger) \leq K. \quad (30)$$

where  $m_1(\hat{\mathbf{H}}^H)$  denotes the *first successive minima*, or equivalently the shortest lattice vector, of the dual lattice generated by  $\hat{\mathbf{H}}^H$ , i.e.

$$m_1(\hat{\mathbf{H}}^H) \triangleq \min_{\mathbf{z} \in \mathbb{G}^K \setminus \{0\}} \|\hat{\mathbf{H}}^H \mathbf{z}\|.$$

We note that [7, Theorem 2.2] applies to real-valued lattices, but (30) follows directly by considering one complex dimension as two real valued dimensions. Combining (29) and (30) yields

$$\Gamma = \|\hat{\mathbf{H}}^\dagger \hat{\mathbf{d}}\|^2 \leq \frac{(\tau K)^2}{m_1^2(\hat{\mathbf{H}}^H)}. \quad (31)$$

The implication of (31) is that  $\Gamma$  cannot be large unless there is a short non-zero vector in the lattice generated by  $\hat{\mathbf{H}}^H$ . Note however that the existence of such a short non-zero vector is not sufficient for  $\Gamma$  to be large. Using (31), we obtain

$$\begin{aligned} \Pr\{\eta\Gamma \geq \kappa\} &\leq \Pr\{m_1^2(\hat{\mathbf{H}}^H) \leq (\tau K)^2 \eta \kappa^{-1}\} \\ &= \Pr\{m_1^2(\mathbf{U}) \leq (\tau K)^2 (1 + \eta)^{-1} \eta \kappa^{-1}\} \end{aligned} \quad (32)$$

where  $\mathbf{U} = (1 + \eta)^{-1} \hat{\mathbf{H}}$  has i.i.d. unit-variance Gaussian elements. By noting that

$$\Pr\{m_1^2(\mathbf{U}) \leq \rho^{-1}\} \doteq \rho^{-M},$$

which follows from [6, Lemma 3], and using (cf. (12))

$$(\tau K)^2 (1 + \eta)^{-1} \eta \kappa^{-1} \doteq \rho^{-\alpha},$$

it follows that

$$\Pr\{m_1^2(\mathbf{U}) \leq (\tau K)^2 (1 + \eta)^{-1} \eta\} \doteq \rho^{-\alpha M},$$

which implies  $d_I \geq \alpha M$  and, hence, when combined with (28), that

$$d_I = \alpha M. \quad (33)$$

The implication of (33) and (22) is that in the presence of imperfect CSI, the diversity can be at most  $\alpha M$ , regardless of the power constraint. This is simply a consequence of the self-interference caused by attempting to invert the channel based on imperfect CSI.

**Noise Diversity Order.** We next sketch the computation of (24) under the two power constraints.

For the *long-term power constraint*,  $\mu^{-1} = \gamma_L P^{-1}$  is non-random (cf. (8)), it follows that

$$\Pr\{\sigma^2 \mu^{-1} \geq \kappa\} = \Pr\{\gamma_L \geq \kappa \rho\} = 0$$

for sufficiently large  $\rho$ . As a consequence,  $d_N = \infty$ , and the overall diversity is given by (15) (c.f. (24) and (22)). In other words, it follows that the system is always interference limited, regardless of the value of  $\alpha$ .

For the *short-term power constraint*, we have  $\mu^{-1} = \gamma_S P^{-1}$  with  $\gamma_S = \mathbb{E}\{\Gamma|\hat{\mathbf{H}}\}$  (cf. (9)), and thus

$$\Pr\{\sigma^2 \mu^{-1} \geq \kappa\} = \Pr\{\gamma_S \geq \kappa \rho\}.$$

Upper and lower bounds on the random variable  $\gamma_S$  are obtained by noting that (25) and (31) are independent of  $\hat{\mathbf{d}}$  and hence

$$\frac{c(1 + \eta)}{\xi_M} = \frac{c}{\|\hat{\mathbf{h}}_1\|^2} \leq \gamma_S \leq \frac{(\tau K)^2}{m_1^2(\hat{\mathbf{H}}^H)} = \frac{(\tau K)^2 (1 + \eta)}{m_1^2(\mathbf{U})}. \quad (34)$$

As a consequence,

$$\Pr\{\gamma_S \geq \kappa \rho\} \geq \Pr\{\xi_M \leq c(1 + \eta)^{-1} \rho^{-1} \kappa^{-1}\}, \quad (35a)$$

$$\Pr\{\gamma_S \geq \kappa \rho\} \leq \Pr\{m_1^2(\mathbf{U}) \leq K(1 + \eta)^{-1} \rho^{-1} \kappa^{-1}\}. \quad (35b)$$

Comparing with (26) and (32) we see that exactly the same arguments as for the interference diversity can be used, the only difference being that the upper bounds in the events on the right-hand side of (35) scale as  $(1 + \eta)^{-1} \rho^{-1} \doteq \rho^{-1}$ . This immediately yields

$$d_N = M. \quad (36)$$

The result of (16) follows and completes the proof of Theorem 1. We note that (36) is equivalent to the result of [6] and could alternatively be obtained from [6, Theorem 3].

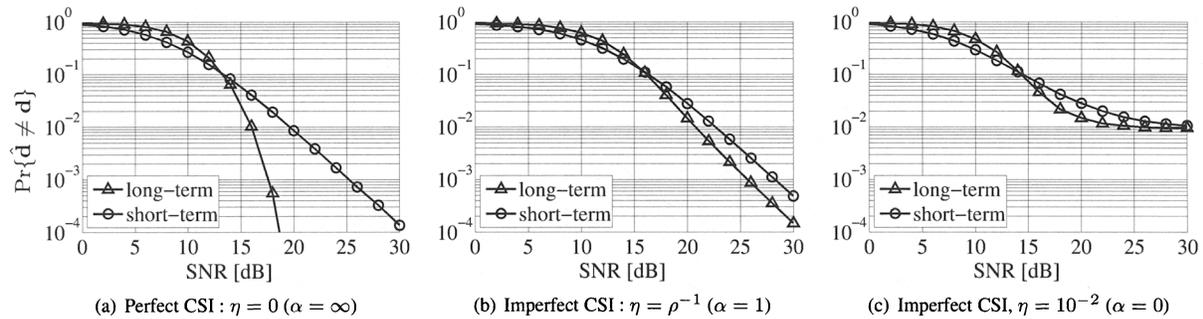


Fig. 2. Probability of error versus SNR,  $\rho$ , for a  $2 \times 2$ , 4-QAM, system with varying degrees of CSI.

#### 4.4. Discussion and Interpretation

It is apparent from the proof of Theorem 1 that in the high SNR regime, receiver errors are mainly caused by the event that the unnormalized instantaneous transmit signal  $\Gamma$  is atypically large. The only exception to this general observation is the case  $\alpha > 1$  (i.e., extremely accurate CSI) under the short-term power constraint, where indeed noise becomes the limiting factor. Thus, for the practically most interesting range  $0 \leq \alpha \leq 1$ , the diversity order is completely determined by the statistics of the random variable  $\Gamma$ . In particular, whenever  $\Gamma$  is on the order of  $\rho^\alpha$ , the accuracy of the CSI is simply too poor to successfully invert the channel and thus errors caused by self-interference become the limiting event. This fact is independent of the power constraint, which implies that using a long-term adaptive power allocation strategy that distributes the average power unevenly over several channel realizations cannot improve diversity. This is intuitive since increasing the instantaneous transmit power will proportionally increase the self interference.

If  $\Gamma \geq \rho^\alpha$ , any attempt to transmit is likely to result in errors. Inspired by this point, we developed a vector perturbation precoder that avoids transmissions by discarding the data whenever  $\Gamma$  is large (see the companion paper [8]). Even though the receivers are very likely to produce errors whenever the data is not transmitted, the gap in error probability performance to the precoding schemes discussed herein is small and in addition is paid off by smaller average transmit power. Another key feature of the approach in [8] is that it allows for a fixed scaling factor  $\mu$  that is known in advance by the receivers while still satisfying an instantaneous (peak) power constraint at the transmitter.

#### 5. SIMULATIONS

In order to illustrate Theorem 1, we provide numerical simulations for a  $2 \times 2$  system with 4-QAM. Two cases are considered for comparison. Fig. 2(a) shows symbol error rate versus average SNR for the case of perfect CSI ( $\eta = 0$ ). It is seen that the diversity order is indeed infinite under the long-term power constraint and equals  $d = 2$  under the short-term constraint. While the diversity under the short-term constraint is seen to remain unchanged for the case of imperfect CSI with  $\alpha = 1$  (Fig. 2(b)), it equals  $d = 2$  also under the long-term constraint. Indeed there is little performance difference in this case. Finally, Fig. 2(c) illustrates a case where the CSI does not improve with increasing SNR. It is seen that the system experiences an error floor at high SNR (i.e., the diversity order is 0), as predicted by Theorem 1.

#### 6. CONCLUSIONS

We have presented a diversity analysis for vector perturbation precoding according to [1, 2], however, operating with imperfect CSI. The main conclusion from this analysis is that such a system, which is essentially based on channel inversion, becomes very sensitive to channel estimation errors. As indicated by Theorem 1, the specific power constraint employed does not make a difference in this respect unless the channel is very accurately known. In essence, the CSI error needs to decay faster than the reciprocal of the SNR in order for the system to benefit from a long-term power constraint as opposed to a short-term power constraint. Unfortunately, this is not a realistic assumption in most scenarios. Thus, as a general rule of thumb, there is little to gain by imposing long-term power constraints and using adaptive power allocation in such a scenario.

#### 7. REFERENCES

- [1] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multi-antenna multiuser communication-Part I: Channel inversion and regularization," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 195–202, Jan 2005.
- [2] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multi-antenna multiuser communication-Part II: Perturbation," *IEEE Transactions on Communications*, vol. 53, no. 3, pp. 537–544, Mar 2005.
- [3] B. Hochwald and S. Vishwanath, "Space-time multiple access: Linear growth in sum rate," in *Proc. Allerton Conf. Communications, Control, Computing*, Monticello, IL, Oct. 2002.
- [4] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1468–1489, July 1999.
- [5] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.
- [6] M. Taherzadeh, A. Mobasher, and A. K. Khandani, "Communication over MIMO broadcast channels using lattice-basis reduction," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4567–4582, Dec 2007.
- [7] W. Banaszczyk, "New bounds in some transference theorems in the geometry of numbers," *Matematische Annalen*, vol. 296, no. 1, pp. 625–635, Dec 1993.
- [8] J. Maurer, J. Jaldén, and G. Matz, "Transmit outage precoding with imperfect channel state information under an instantaneous power constraint," in *Proc. IEEE SPAWC'08*, July 2008.