# TRANSMIT OUTAGE PRECODING WITH IMPERFECT CHANNEL STATE INFORMATION UNDER AN INSTANTANEOUS POWER CONSTRAINT

*Johannes Maurer, Joakim Jaldén, 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 38968; fax: +43 1 58801 38999; E-mail: jmaurer@nt.tuwien.ac.at

**ABSTRACT**

In this paper, we consider a multiple antenna broadcast scenario with multiple non-cooperative users under the assumption of imperfect channel state information. Under an instantaneous transmit power constraint, vector perturbation precoding requires that the users be continuously informed about the adaptive power scaling factor used at the transmit side. To overcome this limitation, we propose a new precoding scheme that uses a fixed power scaling and avoids transmitting when the available power is not sufficient to perform channel equalization at the transmit side, an event referred to as *outage*. We present a performance analysis and optimization of this scheme and provide numerical comparisons with classical vector perturbation.

## 1. INTRODUCTION

### 1.1. Motivation

We consider a wireless broadcast scenario in which a base station uses multiple antennas to transmit simultaneously to multiple users. In this context, vector perturbation based precoding [1–3] is a promising technique since it enables the users to perform optimal detection in a non-cooperative and low-complexity manner. These advantages are achieved by performing a channel inversion at the base station, preceded by a vector perturbation (VP) of the transmit vector in order to reduce the transmit power [2] . The receivers can easily detect their data symbols by performing a modulo operation on the receive signal, followed by a simple quantization. Throughout this paper, we will refer to this scheme as *classical VP precoding*.

To meet a given transmit power constraint, the base station has to apply adaptive power scaling. The power scaling factor has to be known by the receivers for the purpose of detection, which presents a serious obstacle for practical implementation:

- Under a strict instantaneous power constraint, the power scaling factor depends on the channel realization and on the data vector [2]. Communicating the scaling factor to the receivers would require auxiliary transmissions that cause a severe overhead. Furthermore, the auxiliary transmission mode must be *different* from VP precoding (the latter needs the scaling factor already in place), thereby necessitating more sophisticated receiver processing. As a consequence, the advantages of VP precoding would be lost.
- In a block fading scenario, a short-term power constraint can be imposed within each block. Here, the power scaling factor will be averaged over all possible data vectors and remains constant for the duration of each block (i.e., before the channel changes). However, the scaling factor still has to be transmitted to the receivers, causing similar problems as with the instantaneous power constraint (the overhead will depend on the channel coherence time).
- Using a long-term power constraint [2] results in a constant power scaling factor obtained by averaging with respect to the data *and* the channel. This scaling factor can be computed by the receivers in advance. However, such a long-term power constraint requires sufficiently fast fading and has the drawback of potentially very large instantaneous transmit power.

### 1.2. Contributions

In this paper, we propose a novel VP-based precoding scheme that satisfies the instantaneous power constraint while circumventing the need for an adaptive power scaling factor. The main idea of our scheme is to use a non-adaptive (i.e., fixed) scaling factor and abandon a transmission if the resulting transmit power exceeds the power constraint. Transmitting nothing almost surely causes errors but occurs only sporadically. Furthermore, classical precoding is likely to produce an error in these situations as well due to noise enhancement caused by large scaling factors. Thus, the new scheme can be interpreted as avoiding transmission at all (and thereby saving power) if it deems receiver errors likely.

A second important aspect we consider in this paper is imperfect channel state information (CSI) at the base station, which can have a significant effect on precoding performance. Hence, we will study the impact of CSI accuracy on the precoding performance both analytically and numerically via the bit error rate (BER) of our proposed scheme. In fact, it turns out that the performance degradation of the new precoding scheme caused by imperfect CSI is less pronounced than that of classical precoding. We note that a more detailed discussion of the performance of classical VP precoding under imperfect CSI is provided in a companion paper [4].

The rest of the paper is organized as follows. In Section 2, we present the system model, discuss different power constraints, and review classical VP precoding. The novel precoding scheme is presented in Section 3 and its performance is analyzed and optimized in Section 4. Simulation results and conclusions are provided in Sections 5 and 6, respectively.

## 2. SYSTEM MODEL AND CLASSICAL VP PRECODING

### 2.1. System Model

We consider a multi-user communications system operating in the downlink (see e.g. [1, 3]). The base station is equipped with $M$ transmit antennas and there are $K \leq M$ users, each with a single

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

receive antenna. We assume that the user antennas are not co-located so that user cooperation is not possible.

Let $\mathbf{x}[n] \triangleq (x_1[n] \ldots x_M[n])^T$ denote the transmit vector at symbol time $n$. We assume i.i.d. Rayleigh flat fading MISO channels $\mathbf{h}_k[n]$, $k = 1, \ldots, K$, from the base station to the individual users. The $k$th user receives $r_k[n] = \mathbf{h}_k^T[n]\mathbf{x}[n] + w_k[n]$ with the additive noise $w_k[n] \sim \mathcal{CN}(0, \sigma^2)$. By collecting the receive values of the $K$ users in a receive vector $\mathbf{r}[n] \triangleq (r_1[n] \ldots r_K[n])^T$, the overall channel input-output relation can be written as

$$\mathbf{r}[n] = \mathbf{H}[n]\mathbf{x}[n] + \mathbf{w}[n], \quad \text{with } \mathbf{H}[n] = \big(\mathbf{h}_1[n] \ldots \mathbf{h}_K[n]\big)^T. \quad (1)$$

The elements of the $K \times M$ channel matrix $\mathbf{H}[n]$ are assumed i.i.d. $\mathcal{CN}(0,1)$ and the $K \times 1$ noise vector is given by $\mathbf{w}[n] \triangleq (w_1[n] \ldots w_K[n])^T$. Under a block fading model [5], the channel matrix remains constant over $T$ successive time instants.

### 2.2. Transmit Power Constraints

We next state two transmit power constraints (see e.g. [6, 7]) since they are crucial for precoding performance (cf. also [4]).

Under a *long-term power constraint*, the transmit power averaged over all channel and data realizations must not exceed a certain prescribed power $P$, i.e., $\mathrm{E}\big\{\|\mathbf{x}[n]\|^2\big\} \leq P$ (here, $\mathrm{E}\{\cdot\}$ denotes expectation). In an ergodic setting, this condition is equivalent to

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \|\mathbf{x}[n]\|^2 \leq P, \quad (2)$$

which is the power constraint considered in [1].

The more stringent and practically more relevant *instantaneous power constraint* requires that the actual transmit power in each time instant is never larger than $P$, i.e., for all $n$

$$\|\mathbf{x}[n]\|^2 \leq P. \quad (3)$$

It is straightforward to see that the long-term power constraint (2) is automatically satisfied if (3) holds true. In most of the remainder of this paper, we will focus on the instantaneous power constraint (3).

### 2.3. Channel State Information

We assume that the base station has imperfect CSI constituted by a noisy version of the channel matrix $\mathbf{H}[n]$:

$$\hat{\mathbf{H}}[n] = \mathbf{H}[n] + \mathbf{E}[n]\,. \quad (4)$$

The CSI accuracy is characterized by the error matrix $\mathbf{E}[n]$ which is assumed to be independent of $\mathbf{H}[n]$. For simplicity, we model $\mathbf{E}[n]$ as complex Gaussian with i.i.d. elements, i.e.,

$$(\mathbf{E}[n])_{k,m} \sim \mathcal{CN}(0, \eta) \quad \text{with} \quad \eta \triangleq \beta\rho^{-\alpha}, \quad (5)$$

where $\rho \triangleq P/\sigma^2$ denotes the receive SNR and the parameters $\alpha \geq 0$ and $\beta > 0$ determine CSI accuracy. We note that the elements of the CSI matrix $\hat{\mathbf{H}}[n]$ are hence i.i.d. $\mathcal{CN}(0, 1 + \eta)$. The special case of perfect CSI at the base station corresponds to $\alpha \to \infty$. Two practically more relevant special cases are discussed in the following.

In *reciprocal systems*, the downlink and the uplink channel are equal. This allows the base station to estimate the channel matrix (e.g., via pilots sent by the users). In this case, $\mathbf{E}[n]$ models the errors resulting from channel estimation whose accuracy depends e.g. on the noise level and the pilot power. If the pilot power is proportional to $P$, this scenario is described by (5) with $\alpha = 1$.

In *non-reciprocal systems*, CSI feedback is required which entails two kinds of errors: (i) long feedback delay can cause the CSI to be outdated (this depends on the channel coherence time); (ii) finite-rate feedback requires CSI quantization and thus suffers from quantization noise. Note that here the CSI accuracy is independent of SNR which amounts to $\alpha = 0$ in (5).

### 2.4. Classical VP Precoding

**Base station processing.** In what follows, we suppress the index $n$ for convenience. At each time instant, the base station intends to transmit $K$ complex symbols $d_k \in \mathcal{A}$ simultaneously to the $K$ users. Here, $\mathcal{A}$ denotes the symbol alphabet. We assume $\mathrm{E}\big\{|d_k|^2\big\} = 1$ and define the $K \times 1$ data vector $\mathbf{d} \triangleq [d_1 \ldots d_K]^T$ that carries $K \log_2|\mathcal{A}|$ bits that are assumed to be independent and uniformly distributed.

The transmit signal of classical VP precoding [2] under imperfect CSI is obtained from $\mathbf{d}$ according to

$$\mathbf{x} = \sqrt{\frac{P}{\Gamma(\hat{\mathbf{H}}, \mathbf{d})}}\, \hat{\mathbf{H}}^\dagger(\mathbf{d} + \tau\mathbf{z}^\star). \quad (6)$$

Here, $\hat{\mathbf{H}}^\dagger = \hat{\mathbf{H}}^H(\hat{\mathbf{H}}\hat{\mathbf{H}}^H)^{-1}$ denotes the right pseudo inverse, $\tau$ is a fixed real-valued scaling factor, and $\mathbf{z}^\star \in \mathbb{G}^K$ is the perturbation vector whose elements are Gaussian integers[1]. Furthermore, the power scaling factor that has to be chosen adaptively to fulfill the instantaneous power constraint (3) equals

$$\Gamma(\hat{\mathbf{H}}, \mathbf{d}) = \|\hat{\mathbf{H}}^\dagger(\mathbf{d} + \tau\mathbf{z}^\star)\|^2,$$

which is seen to depend explicitly on the data and the channel (estimate). In order to maximize the receive SNR, the perturbation vector is chosen as

$$\mathbf{z}^\star = \arg\min_{\mathbf{z} \in \mathbb{G}^K} \big\|\hat{\mathbf{H}}^\dagger(\mathbf{d} + \tau\mathbf{z})\big\|^2, \quad (7)$$

The minimization problem in (7) is an integer least-squares problem whose complexity in general is exponential in the number of users $K$. A promising algorithm to solve (7) is the sphere decoding algorithm [2, 8] (in this context then also referred to as *sphere encoding* [2]), which, however, is still exponentially complex [9] (in the worst case and on average).

**Receiver processing.** Under the assumption of perfect knowledge of $\Gamma(\hat{\mathbf{H}}, \mathbf{d})$ at all receivers and perfect channel state information at the base station (i.e., $\hat{\mathbf{H}} = \mathbf{H}$), inserting (6) into (1) leads to

$$\tilde{r}_k \triangleq \sqrt{\frac{\Gamma(\mathbf{H}, \mathbf{d})}{P}}\, r_k = d_k + \tau z_k^* + \sqrt{\frac{\Gamma(\mathbf{H}, \mathbf{d})}{P}}\, w_k. \quad (8)$$

This amounts to an additive white Gaussian noise channel with data symbols from the extended alphabet $\mathcal{A} + \tau\,\mathbb{G}$ (i.e. all $\tau$-scaled integer translates of $\mathcal{A}$). If $\tau$ is selected properly (cf. [2]), each user can perform detection by applying a complex-valued modulo-$\tau$ operation (denote $\mathcal{M}_\tau\{\cdot\}$) to $\tilde{r}_k$, followed by quantization with respect to the symbol alphabet $\mathcal{A}$ (denoted $\mathcal{Q}\{\cdot\}$), i.e.,

$$\hat{d}_k = \mathcal{Q}\left\{\mathcal{M}_\tau\{\tilde{r}_k\}\right\}. \quad (9)$$

For this kind of simple receiver processing, it is critical that the users have accurate knowledge of the power scaling factor. This requires auxiliary transmissions that create a significant undesired overhead. Furthermore, the auxiliary transmission mode can not be based on

[1]The set of Gaussian integers $\mathbb{G} = \mathbb{Z} + j\mathbb{Z}$ comprises all complex numbers with integer real and imaginary parts.

VP precoding and hence some other (broadcast) scheme is required for communicating $\Gamma(\hat{\mathbf{H}},\mathbf{d})$. This immediately causes higher complexity at the receivers as they are required to operate in at least two different modes. Auxiliary transmissions can be avoided if the long-term power constraint (2) is imposed since here the power scaling factor in (6) is replaced with its mean

$$\bar{\gamma} = \mathrm{E}\big\{\Gamma(\hat{\mathbf{H}},\mathbf{d})\big\} = \mathrm{E}\big\{\|\hat{\mathbf{H}}^\dagger(\mathbf{d}+\tau\mathbf{z}^\star)\|^2\big\}.$$

However, in this case the transmit amplifiers must be able to deal with the potentially very large instantaneous transmit power (given by $P\Gamma(\hat{\mathbf{H}},\mathbf{d})/\bar{\gamma}$, which can be much larger than $P$).

## 3. TRANSMIT OUTAGE PRECODING

We next introduce the novel precoding scheme as a modification of classical VP precoding that uses a *fixed* power scaling factor that depends neither on the channel realization nor on the transmit data. The transmit signal is given by

$$\mathbf{x} = \begin{cases} \sqrt{\dfrac{P}{\gamma}}\,\hat{\mathbf{H}}^\dagger(\mathbf{d}+\tau\mathbf{z}^\star) & \text{for } \Gamma(\hat{\mathbf{H}},\mathbf{d}) \le \gamma, \\ \mathbf{0} & \text{for } \Gamma(\hat{\mathbf{H}},\mathbf{d}) > \gamma, \end{cases} \tag{10}$$

where the scaling factor $\gamma$ is a design parameter known in advance by the receivers. The precoding scheme (10) is straightforwardly shown to satisfy the instantaneous power constraint (3). The receiver processing is the same as in (8), (9) with $\gamma$ replacing $\Gamma(\hat{\mathbf{H}},\mathbf{d})$.

The main ideas underlying (10) are as follows:

- Under "good" channel conditions ($\Gamma(\hat{\mathbf{H}},\mathbf{d}) \le \gamma$), we transmit the VP precoded data using fixed power scaling, thereby avoiding the need to repeatedly inform the receivers about the scaling factor. The instantaneous transmit power here equals $P\Gamma(\hat{\mathbf{H}},\mathbf{d})/\gamma \le P$ and the noise power for perfect CSI in (8) equals $\gamma/\rho$. This is markedly different from classical VP precoding where the instantaneous transmit power equals $P$ and the noise power for perfect CSI equals $\Gamma(\hat{\mathbf{H}},\mathbf{d})/\rho$.
- For "bad" channels ($\Gamma(\hat{\mathbf{H}},\mathbf{d}) > \gamma$), we voluntarily declare a *transmit outage event* (denoted $\mathcal{O}_\gamma$) and discard the data by sending nothing (i.e., $\mathbf{x}=\mathbf{0}$). In this case, the adaptive power scaling of classical VP precoding degrades the instantaneous receive SNR (cf. (8)) and thus very likely results in detection errors. Avoiding transmission is also likely to produce errors but at least saves transmit power.

Note that the term outage here is not to be understood in the information theoretic sense (where it indicates that the channel does not support a certain rate). Rather, it means that the available transmit power is not sufficient for inverting the channel. The choice of $\gamma$ influences the probability $\Pr\{\mathcal{O}_\gamma\}$ of a transmit outage event but also the instantaneous receive SNR and thereby the BER of non-outage transmissions. A small outage probability requires large $\gamma$, whereas small non-outage BER requires small $\gamma$. This trade-off will be discussed further in Section 4. In this context we note that the optimal perturbation vector $\mathbf{z}^\star$ in (7) can be interpreted as minimizing the probability of an outage event and thereby maximizing the probability of correct decisions at the receivers.

A simple modification of (10) is to transmit in the outage case

$$\mathbf{x} = \sqrt{\frac{P}{\Gamma(\hat{\mathbf{H}},\mathbf{d})}}\hat{\mathbf{H}}^\dagger(\mathbf{d}+\tau\mathbf{z}^\star) \quad \text{for } \Gamma(\hat{\mathbf{H}},\mathbf{d}) > \gamma\,,$$

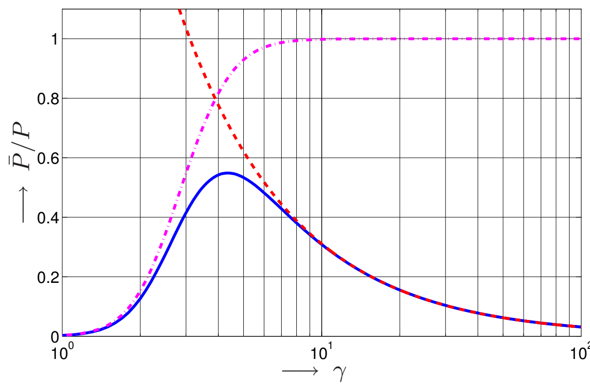



**Fig. 1**. Normalized average transmit power (including upper bounds) of transmit outage precoding versus scaling factor $\gamma$ for $M=6$ antennas, $K=6$ users, 4-QAM symbol alphabet, and perfect CSI.

while the receivers continue to use $\gamma$ instead of the instantaneous scaling factor $\Gamma(\hat{\mathbf{H}},\mathbf{d})$ (which is unknown to them). This is a somewhat more moderate way of dealing with transmission outages that certainly improves BER performance. However, simulation results indicate that the BER improvement is only small and is paid off by a significant increase of the average transmit power.

**Average Transmit Power.** The average transmit power used by transmit outage precoding according to (10) equals

$$\begin{aligned} \bar{P} &\triangleq \mathrm{E}\big\{\|\mathbf{x}\|^2\big\} = \mathrm{E}\big\{\|\mathbf{x}\|^2\big|\bar{\mathcal{O}}_\gamma\big\}\ \Pr\{\bar{\mathcal{O}}_\gamma\} \\ &= \frac{P}{\gamma}\mathrm{E}\big\{\Gamma(\hat{\mathbf{H}},\mathbf{d})\big|\bar{\mathcal{O}}_\gamma\big\}\Pr\{\bar{\mathcal{O}}_\gamma\} \\ &\le P\min\Big\{\frac{\bar{\gamma}}{\gamma},\Pr\{\bar{\mathcal{O}}_\gamma\}\Big\}, \end{aligned}$$

where $\bar{\mathcal{O}}_\gamma$ denotes the non-outage event $\Gamma(\hat{\mathbf{H}},\mathbf{d}) \le \gamma$ (i.e., the complement of $\mathcal{O}_\gamma$) and we used the fact that $\mathrm{E}\big\{\Gamma(\hat{\mathbf{H}},\mathbf{d})\big|\bar{\mathcal{O}}_\gamma\big\}$ is simultaneously upper bounded by $\gamma$ and $\bar{\gamma}$.

For large $\gamma$, $\mathrm{E}\big\{\Gamma(\hat{\mathbf{H}},\mathbf{d})\big|\bar{\mathcal{O}}_\gamma\big\}$ approaches $\bar{\gamma}$ and $\Pr\{\bar{\mathcal{O}}_\gamma\}$ approaches 1 such that $\bar{P}\approx P\bar{\gamma}/\gamma$. For small $\gamma$, $\mathrm{E}\big\{\Gamma(\hat{\mathbf{H}},\mathbf{d})\big|\bar{\mathcal{O}}_\gamma\big\}\approx\gamma$ such that $\bar{P}\approx P\Pr\{\bar{\mathcal{O}}_\gamma\}$. Hence, picking $\gamma$ too small will decrease the average transmit power due to an increasing number of outages whereas for too large choices of $\gamma$ we will transmit almost all the time but with very small power. Note that a small average power is not necessarily suboptimal. An example of average transmit power versus $\gamma$ is shown in Fig. 1 for a $6\times 6$ system with 4-QAM and perfect CSI. We conclude from the plot that in this case transmit outage precoding on average spends only 55% of the maximum power $P$. Note that $\hat{\mathbf{H}}$ has the same distribution as $\sqrt{1+\eta}\mathbf{H}$, which implies that $\mathrm{E}\big\{\Gamma(\hat{\mathbf{H}},\mathbf{d})\big|\bar{\mathcal{O}}_\gamma\big\} = (1+\eta)\mathrm{E}\big\{\Gamma(\mathbf{H},\mathbf{d})\big|\bar{\mathcal{O}}_\gamma\big\}$. Hence, the average power with imperfect CSI is increased by $(1+\eta)$ as compared to perfect CSI. This increase is negligible at moderate to high SNR.

## 4. PERFORMANCE ANALYSIS AND OPTIMIZATION

In this section, we discuss the performance of the proposed precoding scheme and how to optimally choose the parameter $\gamma$ in (10).

### 4.1. Bit Error Probability

The probability of a bit error $\mathcal{E}_b$, $P_{\mathcal{E}}(\gamma) \triangleq \Pr\{\mathcal{E}_b\}$, of transmit outage precoding can be written as

$$P_{\mathcal{E}}(\gamma) = \Pr\{\mathcal{E}_b|\mathcal{O}_\gamma\}\Pr\{\mathcal{O}_\gamma\} + \Pr\{\mathcal{E}_b|\bar{\mathcal{O}}_\gamma\}\big(1-\Pr\{\mathcal{O}_\gamma\}\big). \tag{11}$$

**Transmit Outage.** In the case of a transmit outage, the receivers see only noise, i.e., $\tilde{r}_k = \sqrt{\frac{\gamma}{P}} w_k$ (cf. (8)). This implies that $\Pr\{\mathcal{E}_b|\mathcal{O}_\gamma\} = 1/2$. The outage probability can be developed as

$$\begin{aligned}\Pr\{\mathcal{O}_\gamma\} &= \Pr\{\Gamma(\hat{\mathbf{H}}, \mathbf{d}) > \gamma\} = \Pr\{\|\hat{\mathbf{H}}^\dagger(\mathbf{d} + \tau\mathbf{z}^\star)\|^2 > \gamma\} \\ &= \Pr\{\|\mathbf{H}^\dagger(\mathbf{d} + \tau\mathbf{z}^\star)\|^2 > (1+\eta)\gamma\}\end{aligned}$$

where we used the fact that $\hat{\mathbf{H}}$ has the same distribution as $\mathbf{H}\sqrt{1+\eta}$. By letting $g(\xi)$ denote the complementary cumulative distribution function of $\Gamma(\mathbf{H}, \mathbf{d})$, i.e., $g(\xi) \triangleq \Pr\{\|\mathbf{H}^\dagger(\mathbf{d} + \tau\mathbf{z}^\star)\|^2 > \xi\}$, the outage probability can be written as

$$\Pr\{\mathcal{O}_\gamma\} = g\big((1+\eta)\gamma\big). \tag{12}$$

**No Transmit Outage.** We next turn to $\Pr\{\mathcal{E}_b|\bar{\mathcal{O}}_\gamma\}$. For the case of no outage, the receivers see

$$\begin{aligned}\tilde{\mathbf{r}} &\triangleq \sqrt{\frac{\gamma}{P}}\mathbf{r} = \sqrt{\frac{\gamma}{P}}(\mathbf{H}\mathbf{x} + \mathbf{w}) = \mathbf{H}\hat{\mathbf{H}}^\dagger(\mathbf{d}+\tau\mathbf{z}^\star) + \sqrt{\frac{\gamma}{P}}\mathbf{w} \\ &= (\hat{\mathbf{H}} - \mathbf{E})\hat{\mathbf{H}}^\dagger(\mathbf{d} + \tau\mathbf{z}^\star) + \sqrt{\frac{\gamma}{P}}\mathbf{w} = \tilde{\mathbf{d}} + \mathbf{v}\end{aligned}$$

where $\tilde{\mathbf{d}} \triangleq \mathbf{d}+\tau\mathbf{z}^\star$ and $\mathbf{v} \triangleq -\mathbf{E}\hat{\mathbf{H}}^\dagger\tilde{\mathbf{d}}+\sqrt{\frac{\gamma}{P}}\mathbf{w}$. Using (5), it follows that conditioned on $\tilde{\mathbf{d}}$ and $\hat{\mathbf{H}}$, the elements of $\mathbf{v}$ are i.i.d. $\mathcal{CN}(0, \nu)$ where $\nu \triangleq \eta\|\hat{\mathbf{H}}^\dagger\tilde{\mathbf{d}}\|^2 + \gamma/\rho$. Since the base station guarantees that $\|\hat{\mathbf{H}}^\dagger\tilde{\mathbf{d}}\|^2 \le \gamma$ holds for the event $\bar{\mathcal{O}}_\gamma$, the variance $\nu$ can be bounded according by

$$\frac{\gamma}{\rho} \le \nu \le \eta\gamma + \frac{\gamma}{\rho}. \tag{13}$$

This leads to an effective signal-to-interference and noise ratio (SINR) lying below $\frac{\rho}{\gamma}$ (in the case of perfect CSI) and above $\frac{\rho}{\gamma}\frac{1}{1+\beta\rho^{1-\alpha}}$ (cf. (5)). Hence, for reciprocal systems ($\alpha = 1$), CSI imperfections reduce the worst-case SINR only by a constant factor of $1/(1+\beta)$. In contrast, for systems with feedback ($\alpha = 0$), the worst-case SINR saturates at $1/(\gamma\beta)$ for increasing SNR.

The symbol error probability at the receiver conditioned on no outage is

$$\Pr\{\mathcal{E}_s|\bar{\mathcal{O}}_\gamma\} = \Pr\{\mathbf{d} \ne \hat{\mathbf{d}}|\bar{\mathcal{O}}_\gamma\},$$

where $\hat{\mathbf{d}} = \mathcal{Q}\{\mathcal{M}_\tau\{\tilde{\mathbf{r}}\}\}$. Closed-form expressions for $\Pr\{\mathcal{E}_s|\bar{\mathcal{O}}_\gamma\}$ (and consequently for $\Pr\{\mathcal{E}_b|\bar{\mathcal{O}}_\gamma\}$) are difficult to obtain since this requires to take into account the extended symbol alphabet and the corresponding extended decision regions. Nevertheless, the bit error probability conditioned on no outage can be written as

$$\Pr\{\mathcal{E}_b|\bar{\mathcal{O}}_\gamma\} = \mathrm{E}\{h(\nu)|\bar{\mathcal{O}}_\gamma\} = \int_0^\infty h(\nu)\, f(\nu|\bar{\mathcal{O}}_\gamma)\, d\nu, \tag{14}$$

where $h(\nu) \triangleq \Pr\{\mathcal{E}_b|\bar{\mathcal{O}}_\gamma, \nu\}$ and $f(\nu|\bar{\mathcal{O}}_\gamma)$ is the probability density function of $\nu$ conditioned on no outage.

It follows from (13) that $f(\nu|\bar{\mathcal{O}}_\gamma)$ has finite support, $f(\nu|\bar{\mathcal{O}}_\gamma) = 0$ for $\nu \notin [\gamma/\rho, \gamma\eta + \gamma/\rho]$. Using this fact and the relation $\Pr\{\mathcal{E}_b|\bar{\mathcal{O}}_\gamma\} = \mathrm{E}\{h(\nu)|\bar{\mathcal{O}}_\gamma\}$, it follows that there must be a $\nu_0 \in [\gamma/\rho, \gamma\eta + \gamma/\rho]$ such that $\mathrm{E}\{h(\nu)|\bar{\mathcal{O}}_\gamma\} = h(\nu_0)$. Writing $\nu_0 = \theta_0\gamma\eta + \frac{\gamma}{\rho}$ with appropriately chosen $0 \le \theta_0 \le 1$, this entails

$$\Pr\{\mathcal{E}_b|\bar{\mathcal{O}}_\gamma\} = h\left(\theta_0\gamma\eta + \frac{\gamma}{\rho}\right), \tag{15}$$

Conjecturing that $h(\nu)$ is a strictly increasing function (i.e., an increase in the interference plus noise variance $\nu$ causes a larger error probability), upper and lower bounds on $\Pr\{\mathcal{E}_b|\bar{\mathcal{O}}_\gamma\}$ are hence obtained by setting $\theta_0 = 1$ and $\theta_0 = 0$, respectively:

$$h\left(\frac{\gamma}{\rho}\right) \le \Pr\{\mathcal{E}_b|\bar{\mathcal{O}}_\gamma\} \le h\left(\gamma\eta + \frac{\gamma}{\rho}\right). \tag{16}$$

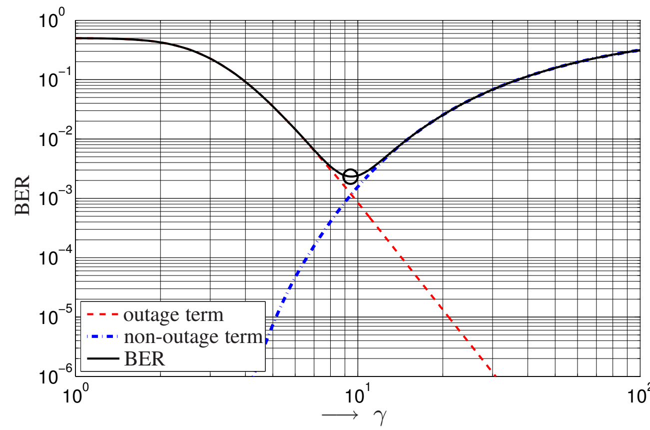



**Fig. 2**. Bit error probability $P_\mathcal{E}(\gamma)$, outage and non-outage probability terms versus power scaling factor $\gamma$ for a $6 \times 6$ system with 4-QAM symbol alphabet, perfect CSI, and $\rho = 20$ dB. The minimum of $P_\mathcal{E}(\gamma)$ at $\gamma^\star = 9.4$ is marked with a circle.

### 4.2. Performance Optimization

The choice of $\gamma$ optimizing bit error probability is given by (cf. (11))

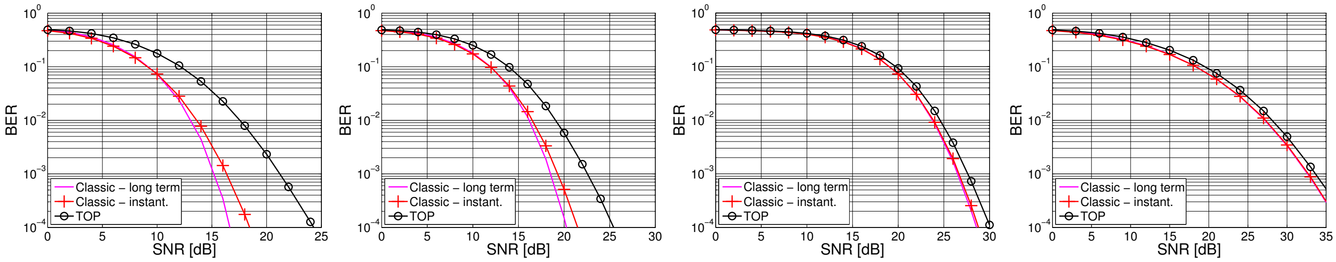
$$\gamma^\star = \arg\min_{\gamma\in\mathbb{R}^+} P_\mathcal{E}(\gamma). \tag{17}$$

This amounts to a trade-off between the outage probability and the BER at the receiver conditioned on no outage. By inserting (12) and (15) into (11), we obtain

$$P_\mathcal{E}(\gamma) = \frac{1}{2}\, g\big((1+\eta)\gamma\big) + \big[1 - g\big((1+\eta)\gamma\big)\big]\, h(\theta_0\gamma\eta+\gamma/\rho), \tag{18}$$

which is the basis for numerically solving (17). The first term of the sum in (18) is referred to as the outage term and the second as the non-outage term. Unfortunately, $\theta_0$ is unknown a priori. Upper and lower bounds for $P_\mathcal{E}(\gamma)$ can be obtained by using (16) instead of (15). For practical purposes, the functions $g(\xi)$ and $h(\nu)$ are determined using Monte-Carlo simulations and an approximation of $P_\mathcal{E}(\gamma)$ is then obtained by replacing $\theta_0$ with a specifically chosen $0 \le \theta \le 1$. The resulting approximation for $P_\mathcal{E}(\gamma)$ can then be used to determine the optimal scaling factor $\gamma^\star$ in (17). This can be done offline for any desired SNR.

As a numerical illustration for the case of perfect CSI ($\eta = 0$), Fig. 2 depicts the overall bit error probability $P_\mathcal{E}(\gamma) = \frac{1}{2}\, g(\gamma) + [1 - g(\gamma)]\, h(\gamma/\rho)$, and its two constituents $\Pr\{\mathcal{O}_\gamma\} = g(\gamma)$ and $\Pr\{\mathcal{E}_b|\bar{\mathcal{O}}_\gamma\} = h(\gamma/\rho)$ versus the scaling parameter $\gamma$, again for a $6\times 6$ system with 4-QAM symbol alphabet at an SNR of $\rho = 20$ dB. In this example $\gamma^\star = 9.4$. It can be seen from this simulation example that the minimization of $P_\mathcal{E}(\gamma)$ with respect to $\gamma$ amounts to balancing the outage and the non-outage term. Furthermore, we observe that increasing the SNR only affects the non-outage term which will be shifted to higher $\gamma$ values while the outage probability will remain unaffected. Correspondingly, $\gamma^\star$ will become larger for higher SNRs.




(a) (b) (c) (d)

**Fig. 3**. BER versus SNR for transmit outage precoding (labeled 'TOP') and classic precoding: (a) perfect CSI, (b) $\alpha = 1$, $\beta = 1$, (c) $\alpha = 1$, $\beta = 10$, and (d) $\alpha = 0.5$, $\beta = 1$.

## 5. SIMULATION RESULTS

In this section, we present numerical results to assess the performance of the proposed transmit outage precoding scheme for an uncoded $6\times6$ system with 4-QAM symbol alphabet. While the optimal scaling factor $\gamma^\star$ was computed numerically (using $\theta = 0.3 \ldots 0.5$) as outlined in Section 4.2, the BER results in this section have been obtained by simulating the overall transceiver chain. Classical precoding (see (6)) under the instantaneous and the long-term power constraint serves as performance benchmarks. Fig. 3 shows BER versus *nominal* SNR $\rho$ for four scenarios that differ in CSI accuracy.

In the case of perfect CSI (cf. Fig. 3(a)), we observe that transmit outage precoding (labeled 'TOP') performs about 5 dB poorer than classical precoding under the instantaneous power constraint. However, we emphasize that our scheme uses much less average transmit than classical VP precoding. In fact, if the BER is plotted versus average receive SNR instead of nominal SNR, our scheme actually has a slight performance advantage (in addition to not requiring communication of the scaling factor). Also note that classical precoding with long-term power constraint performs best, however, at the cost of strongly fluctuating instantaneous transmit power.

Fig. 3(b) shows results for imperfect CSI characterized by $\alpha = \beta = 1$, i.e., $\eta = 1/\rho$. Here, it is seen that all curves are shifted to higher SNRs since interference kicks in as a result of non-ideal channel inversion at the transmitter. Furthermore, the gaps between the different precoding schemes are noticeably smaller for the same reason. When CSI accuracy becomes worse ($\alpha = 1$, $\beta = 10$), the just mentioned behavior is even more pronounced (cf. Fig. 3(b)). Here, the gap between classical precoding and our scheme is reduced to less than 1 dB since all schemes are mostly interference limited. Note that this is a scenario where the uplink pilot signaling is 10 dB weaker than the downlink power, which is representative e.g. for battery-operated user terminals. Finally, Fig. 3(d) presents the BER for imperfect CSI with $\alpha = 0.5$, $\beta = 1$. This is the worst CSI accuracy considered here and all schemes are seen to be interference limited and to perform almost equal. It is seen that in this case even the diversity is reduced.

We re-iterate that the average transmit power of our scheme is much smaller than that of classical precoding and all plots would change in favor of our scheme if BER were plotted versus average receive SNR.

## 6. CONCLUSIONS

We proposed a novel vector perturbation (VP) based precoding scheme that satisfies an instantaneous power constraint by using a fixed power scaling factor and avoiding transmission in the case of outage events that correspond to poorly conditioned channel realizations. Compared to classical VP precoding this has the advantage that no channel- and data-dependent scaling factor needs to be communicated continuously to the receivers. Based on a theoretical analysis of the bit error probability of transmit outage precoding we presented a performance optimization and corroborated the performance of the new scheme via numerical simulations. In particular, it was seen that under realistic assumptions regarding the accuracy of transmit-side channel state information, transmit outage precoding performs as well as classical precoding while saving a significant portion of the available transmit power.

The effect of the multi-user interference on the high-SNR performance of VP precoding is investigated in detail in a companion paper [4]. One of the results therein establishes a diversity order of $\alpha M$ regardless the power constraint for $\alpha \leq 1$, which is confirmed by our simulation results.

## 7. REFERENCES

[1] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication - Part I: Channel inversion and regularization," *IEEE Trans. Comm.*, 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 multiantenna multiuser communication - Part II: Perturbation," *IEEE Trans. Comm.*, vol. 53, no. 3, pp. 537–544, March 2005.

[3] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Trans. Wireless Comm.*, vol. 3, no. 4, pp. 1305–1316, July 2004.

[4] J. Jaldén, J. Maurer, and G. Matz, "On the diversity order of vector perturbation precoding with imperfect channel state information," in *Proc. IEEE SPAWC'08*, July 2008.

[5] A. Goldsmith, *Wireless Communications*, Cambridge Univ. Press, Cambridge (UK), 2005.

[6] 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.

[7] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468–1589, July 1999.

[8] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2201–2214, Aug. 2002.

[9] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Processing*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.