ON THE DIVERSITY-COMPLEXITY TRADEOFF IN MIMO SPATIAL MULTIPLEXING SYSTEMS

Johannes Maurer, Gerald Matz, and Dominik Seethaler

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Vienna, Austria
phone: +43 1 58801 38968; fax: +43 1 58801 38999; email: jmaurer@nt.tuwien.ac.at

ABSTRACT

In MIMO spatial multiplexing systems, maximum likelihood detection achieves maximum receive diversity at the cost of high computational complexity. In contrast, low cost receivers like zero-forcing equalization result in minimum receive diversity. In this paper, we present a novel detection scheme that allows for a continuous tradeoff between diversity and complexity. The proposed receiver consists of a partial equalization stage and of a mismatched ML detector (implemented via a sphere decoder variant). Simulation results confirm that our scheme allows to trade diversity for complexity.

1. INTRODUCTION

Two of the most important benefits of multiple-input multiple-output (MIMO) wireless systems are boosted capacity (multiplexing gain) and improved reliability (diversity gain) [1]. In this paper, we consider $M_{\mathrm{R}} \times M_{\mathrm{T}}$ spatial multiplexing systems that offer multiplexing gain M_{T} and allow to achieve receive diversity gain M_{R}, provided that appropriate detection algorithms like maximum-likelihood (ML) detection are used. Unfortunately, ML detection is computationally very expensive, even when implemented using the sphere-decoding algorithm [2, 3]. On the other extreme, suboptimum detection schemes like zero-forcing (ZF) and MMSE equalization, nulling and canceling (NC), and decision-feedback techniques [4, 5] are computationally much less expensive but achieve only diversity $M_{\mathrm{R}} - M_{\mathrm{T}} + 1$.

In this paper, we propose a novel receiver that allows to trade diversity and complexity in a continuous manner. Our receiver consists of two stages. The first stage partially equalizes the channel to improve the condition number and the second stage performs mismatched ML detection via a sphere decoder variant suited for low signal to noise ratios (SNR). Diversity and complexity essentially depend on the amount of equalization. The resulting diversity-complexity tradeoff is studied using analytical methods and numerical simulations.

System Model. For simplicity, we restrict to MIMO systems with equal number of transmit and receive antennas, $M_{\mathrm{R}} = M_{\mathrm{T}} = M$. The transmit vector $\mathbf{x} \in \mathcal{A}^M$ (with $\mathcal{A}$ denoting the symbol alphabet) is i.i.d. with $\mathcal{E}\{\mathbf{x}\mathbf{x}^H\} = \mathbf{I}$. Assuming a flat-fading channel[1], the length-M receive vector equals

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{n}\,. \tag{1}$$

Here, $\mathbf{H}$ is the $M \times M$ channel matrix and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2\mathbf{I})$ denotes spatially white complex Gaussian noise. The channel $\mathbf{H}$ is assumed perfectly known at the receiver. For further reference, we next briefly review ML and ZF detection.

ML Detection. ML detection [1, 2] is optimal in the sense of minimizing error probability. It amounts to solving

$$\hat{\mathbf{x}}_{\mathrm{ML}} = \underset{\mathbf{x}\in\mathcal{A}^M}{\arg\min}\, \|\mathbf{r} - \mathbf{H}\mathbf{x}\|^2.$$

The computational complexity of ML detection in general grows exponentially with M, even if it is implemented using the sphere-decoding algorithm [3].

Conditioned on the channel $\mathbf{H}$, the pair-wise error probability (PEP) of ML detection equals [1]

$$P_{\mathrm{ML}}(\mathbf{x}_1 \to \mathbf{x}_2|\mathbf{H}) = \mathrm{Q}\left(\frac{\|\mathbf{H}\boldsymbol{\delta}\|}{\sqrt{2}\sigma_n}\right) \tag{2}$$

where $\boldsymbol{\delta} = \mathbf{x}_1 - \mathbf{x}_2$ is the error vector and $\mathrm{Q}(\cdot)$ denotes the Q-function. For i.i.d. Rayleigh fading the mean PEP is upper bounded as [1]

$$P_{\mathrm{ML}}(\mathbf{x}_1 \to \mathbf{x}_2) \;\le\; C_{\mathrm{ML}}\left(4 + \frac{\|\boldsymbol{\delta}\|^2}{\sigma_n^2}\right)^{-M}, \tag{3}$$

where $C_{\mathrm{ML}} > 0$ is a constant. Defining the diversity order as

$$d = \lim_{1/\sigma_n^2\to\infty} \frac{\log P(\mathbf{x}_1 \to \mathbf{x}_2)}{\log \sigma_n^2}, \tag{4}$$

it follows from (3) that $d_{\mathrm{ML}} = M$.

ZF Detection. ZF detection is based on quantizing the ZF equalized receive vector, i.e.,

$$\hat{\mathbf{x}}_{\mathrm{ZF}} = \mathcal{Q}_{\mathcal{A}}\{\mathbf{y}\}\,, \qquad \mathbf{y} = \mathbf{H}^{\#}\mathbf{r} = \mathbf{x} + \mathbf{H}^{\#}\mathbf{n}\,, \tag{5}$$

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

[1] The assumption of flat fading is no serious restriction since frequency-selective channels can be converted into parallel flat-fading channels using OFDM.

where $\mathcal{Q}_{\mathcal{A}}\{\cdot\}$ denotes component-wise quantization with respect to $\mathcal{A}$ and $\mathbf{H}^{\#} = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H$ denotes the pseudo-inverse of $\mathbf{H}$.

For given $\mathbf{H}$, the PEP of the ZF detector is given by [6]

$$P_{\text{ZF}}(\mathbf{x}_1 \to \mathbf{x}_2|\mathbf{H}) = \text{Q}\left(\frac{1}{\sqrt{2}\sigma_n}\frac{\|\boldsymbol{\delta}\|^2}{\sqrt{\boldsymbol{\delta}^H(\mathbf{H}^H\mathbf{H})^{-1}\boldsymbol{\delta}}}\right),$$

and for i.i.d. Rayleigh fading, the mean PEP is bounded as [7]

$$P_{\text{ZF}}(\mathbf{x}_1 \to \mathbf{x}_2) \;\le\; C_{\text{ZF}}\left(1+\frac{\|\boldsymbol{\delta}\|^2}{4\sigma_n^2}\right)^{-1}.$$

The ZF detector thus yields diversity gain $d_{\text{ZF}} = 1$.

2. PROPOSED RECEIVER

ML and ZF are opposite extremes in terms of diversity and complexity. Our goal in this paper is to devise a scheme that realizes a continuous trade-off between these extremes. To this end, we first provide a unifying reformulation of ML and ZF detection.

2.1. Unifying Framework for ML and ZF

In order to place ML and ZF detection on a common basis, we start out with the ZF domain relation $\mathbf{y} = \mathbf{x} + \mathbf{H}^{\#}\mathbf{n}$ (cf. (5)). Note that $\mathbf{y}$ is a sufficient statistic for $\mathbf{x}$. The noise in the ZF domain is distributed as $\mathbf{H}^{\#}\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2\mathbf{G}^{-1})$, where $\mathbf{G} = \mathbf{H}^H\mathbf{H}$ denotes the Gram matrix of $\mathbf{H}$. Consequently, the ZF domain receive vector $\mathbf{y}$ conditioned on $\mathbf{x}$ has distribution $\mathcal{CN}(\mathbf{x}, \sigma_n^2\mathbf{G}^{-1})$. This allows to rewrite the ML detector (2) as

$$\hat{\mathbf{x}}_{\text{ML}} = \underset{\mathbf{x}\in\mathcal{A}^M}{\arg\min}\,(\mathbf{y}-\mathbf{x})^H\mathbf{G}(\mathbf{y}-\mathbf{x}).$$

Similarly, the component-wise quantization (5) performed by the ZF detector can be reformulated as

$$\hat{\mathbf{x}}_{\text{ZF}} = \underset{\mathbf{x}\in\mathcal{A}^M}{\arg\min}\,(\mathbf{y}-\mathbf{x})^H(\mathbf{y}-\mathbf{x}).$$

Both ML and ZF detection can thus be written in terms of a positive definite quadratic form in the ZF domain, the difference being that ML takes the full correlation matrix into account by using $\mathbf{G} = \mathbf{G}^1$ in the quadratic form, whereas ZF detection completely ignores the noise correlation by using a quadratic form induced by $\mathbf{I} = \mathbf{G}^0$.

2.2. Proposed Scheme

The observations in the previous subsection motivate us to propose an intermediate receiver between the ML and ZF extremes by taking part of the noise correlation matrix into account, i.e.,

$$\hat{\mathbf{x}} = \underset{\mathbf{x}\in\mathcal{A}^M}{\arg\min}\,(\mathbf{y}-\mathbf{x})^H\mathbf{G}^{1-\alpha}(\mathbf{y}-\mathbf{x}), \tag{6}$$

where $0 \le \alpha \le 1$. Choosing $\alpha = 0$ or $\alpha = 1$ yields the ML and ZF detector, respectively. For $0 < \alpha < 1$, we expect that (6) leads to some tradeoff between ML and ZF. In particular, using the eigenvalue decomposition [8] $\mathbf{G} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^H$ it can easily be shown that $\mathbf{G}^{1-\alpha} = \mathbf{V}\boldsymbol{\Lambda}^{1-\alpha}\mathbf{V}^H$ has condition number $\kappa(\mathbf{G}^{1-\alpha}) = \kappa^{1-\alpha}(\mathbf{G})$ that decays exponentially with α. Small condition numbers are desirable since detection complexity increases with growing condition number [9]. On the other hand, the part $\mathbf{G}^{\alpha}$ of the noise correlation is ignored in (6), which is expected to have a negative impact on the detection performance.

2.3. Practical Implementation

Rather than solving (6) directly, we next derive a formulation that is more convenient for implementation and interpretation. To this end, we take the square-root of $\mathbf{G}^{1-\alpha}$, and rewrite the quadratic form in (6) as

$$(\mathbf{y}-\mathbf{x})^H\mathbf{G}^{\frac{1-\alpha}{2}}\mathbf{G}^{\frac{1-\alpha}{2}}(\mathbf{y}-\mathbf{x}) = \left\|\mathbf{G}^{\frac{1-\alpha}{2}}\mathbf{y} - \mathbf{G}^{\frac{1-\alpha}{2}}\mathbf{x}\right\|^2. \tag{7}$$

Using the singular value decomposition [8] $\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^H$ of the channel matrix, the first term on the right-hand side in (7) can be developed as

$$\begin{aligned}\tilde{\mathbf{r}} \;&=\; \mathbf{G}^{\frac{1-\alpha}{2}}\mathbf{y} = \mathbf{G}^{\frac{1-\alpha}{2}}\mathbf{H}^{\#}\mathbf{r} = \mathbf{V}\boldsymbol{\Lambda}^{\frac{1-\alpha}{2}}\mathbf{V}^H\,\mathbf{V}\mathbf{D}^{-1}\mathbf{U}^H\mathbf{r}\\ &=\; \mathbf{V}\mathbf{D}^{-\alpha}\mathbf{U}^H\mathbf{r} = \mathbf{W}_{\alpha}^{-1}\mathbf{r}\,.\end{aligned} \tag{8}$$

Here, we used the unitarity of $\mathbf{V}$ and the fact that $\boldsymbol{\Lambda} = \mathbf{D}^2$. The expression $\mathbf{G}^{\frac{1-\alpha}{2}}\mathbf{y}$ is thus equivalent to applying a *partial equalizer* $\mathbf{W}_{\alpha}^{-1} = \mathbf{V}\mathbf{D}^{-\alpha}\mathbf{U}^H$ to the receive vector $\mathbf{r}$ (indeed, $\mathbf{W}_1^{-1} = \mathbf{H}^{\#}$). To further understand the effect of this partial equalizer, we insert (1) into (8), which yields

$$\tilde{\mathbf{r}} = \mathbf{W}_{\alpha}^{-1}(\mathbf{H}\mathbf{x}+\mathbf{n}) = \mathbf{C}_{\alpha}\mathbf{x} + \tilde{\mathbf{n}}\,. \tag{9}$$

Here, $\tilde{\mathbf{n}} = \mathbf{W}_{\alpha}^{-1}\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2\mathbf{G}^{-\alpha})$ denotes spatially correlated noise and $\mathbf{C}_{\alpha} = \mathbf{W}_{\alpha}^{-1}\mathbf{H}$ is what remains of the channel after applying the partial equalizer. It turns out that

$$\mathbf{C}_{\alpha} = \mathbf{V}\mathbf{D}^{-\alpha}\mathbf{U}^H\,\mathbf{U}\mathbf{D}\mathbf{V}^H\,\mathbf{x} = \mathbf{V}\mathbf{D}^{1-\alpha}\mathbf{V}^H = \mathbf{G}^{\frac{1-\alpha}{2}}\,.$$

Hence, (6) can finally be written as

$$\hat{\mathbf{x}} = \underset{\mathbf{x}\in\mathcal{A}^M}{\arg\min}\,\left\|\tilde{\mathbf{r}} - \mathbf{C}_{\alpha}\mathbf{x}\right\|^2. \tag{10}$$

This looks like the ML detector for the linear model (9) apart of the fact that the correlation $\sigma_n^2\mathbf{G}^{-\alpha}$ of the noise $\tilde{\mathbf{n}}$ is being ignored. Hence, we refer to (10) as *mismatched ML detection*.

To summarize, we have shown that our proposed scheme (6) can be implemented in terms of two stages, illustrated in Fig. 1: the first stage (see (8)) partially equalizes the channel and the second stages performs mismatched ML detection using the equalizer output according to (10), ignoring noise correlations. For small α (close to 0), there is little equalization

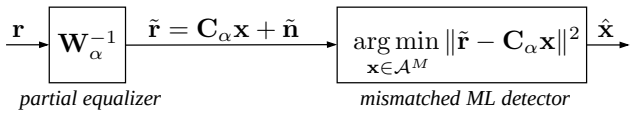


Fig. 1. *Proposed receiver consisting of a partial equalizer and mismatched ML detection.*

and correspondingly low noise correlation being ignored. For large α (close to 1), the mismatched ML detector performs poorly since it ignores strong noise correlations; however, in this case the channel is almost fully equalized (i.e., $\mathbf{C}_\alpha$ almost equals the identity matrix), which results in strongly reduced complexity. In the next section, we analyze this tradeoff in more detail.

3. DIVERSITY-COMPLEXITY TRADEOFF

In this section, we show that by varying α the proposed detector realizes a continuos tradeoff between diversity and complexity.

3.1. Diversity

We have performed extensive simulations (see Section 4) to assess the performance of our proposed scheme. In all our results we observed that the parameter α has an impact on the receive diversity. More specifically, diversity measurements with different number of antennas, different symbol constellations, and varying α suggest the following conjecture.

Diversity Conjecture. *For i.i.d. Rayleigh fading channels with $M_R = M_T = M$, the receive diversity achieved by the proposed receiver equals $d = \alpha + (1-\alpha)\,M$.*

This conjecture says that the diversity of our scheme is an α-dependent convex combination of the minimum diversity $d_{\text{ZF}} = 1$ and the maximum diversity $d_{\text{ML}} = M$. While all our numerical simulations were in perfect agreement with this conjecture, up to now we were not able to find a general proof. While the conditional PEP of our proposed receiver can be obtained via standard techniques,

$$P(\mathbf{x}_1 \to \mathbf{x}_2|\mathbf{H}) = \mathrm{Q}\bigg(\frac{1}{\sqrt{2}\sigma_n}\frac{\|\mathbf{D}^{1-\alpha}\mathbf{V}^H\boldsymbol{\delta}\|^2}{\|\mathbf{D}^{1-2\alpha}\mathbf{V}^H\boldsymbol{\delta}\|}\bigg)\,, \qquad (11)$$

averaging with respect to the channel for arbitrary M and α is difficult. Below, we sketch a proof for the case $\alpha = 1/2$.

Special Case $\boldsymbol{\alpha = 1/2}$. For $\alpha = 1/2$, (11) simplifies to

$$P(\mathbf{x}_1 \to \mathbf{x}_2|\mathbf{H}) = \mathrm{Q}\bigg(\frac{1}{\sqrt{2}\sigma_n}\frac{\boldsymbol{\delta}^H\mathbf{S}\boldsymbol{\delta}}{\|\boldsymbol{\delta}\|}\bigg)\,,$$

where $\mathbf{S} = \mathbf{G}^{1/2} = \mathbf{V}\mathbf{D}\mathbf{V}^H$ denotes the positive definite square root of the Gram matrix $\mathbf{G}$. Since $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, both $\mathbf{G}$ and $\mathbf{S}$ are isotropically distributed (in fact, $\mathbf{G}$ has Wishart distribution). This means that $\mathbf{S}$ and $\mathbf{T}\mathbf{S}\mathbf{T}^H$ have the same distribution if $\mathbf{T}$ is a fixed unitary matrix. Choosing $\mathbf{T}$ as a Householder reflection [10] such that $\mathbf{T}\boldsymbol{\delta} = \|\boldsymbol{\delta}\|\mathbf{e}_1$ with $\mathbf{e}_1 = [1, 0, \ldots, 0]^T$, it follows that[2]

$$\begin{aligned}\boldsymbol{\delta}^H\mathbf{S}\boldsymbol{\delta} &= \boldsymbol{\delta}^H\mathbf{T}^H\mathbf{T}\mathbf{S}\mathbf{T}^H\mathbf{T}\boldsymbol{\delta} = \|\boldsymbol{\delta}\|^2\mathbf{e}_1^H\mathbf{T}\mathbf{S}\mathbf{T}^H\mathbf{e}_1 \\ &= \|\boldsymbol{\delta}\|^2[\mathbf{T}\mathbf{S}\mathbf{T}^H]_{11} \sim \|\boldsymbol{\delta}\|^2 s_{11}\,,\end{aligned}$$

where $s_{11} = [\mathbf{S}]_{11}$ is the element of $\mathbf{S}$ in the first row and first column. Hence, the quadratic form $\boldsymbol{\delta}^H\mathbf{S}\boldsymbol{\delta}$ has the same distribution as (a scaled version of) s_{11}. It follows that

$$\begin{aligned}P(\mathbf{x}_1 \to \mathbf{x}_2) &= \mathcal{E}\bigg\{\mathrm{Q}\bigg(\frac{\boldsymbol{\delta}^H\mathbf{S}\boldsymbol{\delta}}{\sqrt{2}\sigma_n\|\boldsymbol{\delta}\|}\bigg)\bigg\} = \mathcal{E}\bigg\{\mathrm{Q}\bigg(\frac{\|\boldsymbol{\delta}\|s_{11}}{\sqrt{2}\sigma_n}\bigg)\bigg\} \\ &\le \mathcal{E}\bigg\{\exp\bigg(-\frac{\|\boldsymbol{\delta}\|^2 s_{11}^2}{4\sigma_n^2}\bigg)\bigg\}. \qquad (12)\end{aligned}$$

To evaluate (12), we need the distribution of s_{11} and hence also of $\mathbf{S}$. We start out from the Wishart distribution of the Gram matrix [11], $f_{\mathbf{G}}(\mathbf{G}) = C\,\exp\big(-\mathrm{tr}\{\mathbf{G}\}\big)$ for $\mathbf{G} > \mathbf{0}$, where C is a normalization constant. The distribution of $\mathbf{S}$ is then obtained via the random matrix transformation [12]

$$f_{\mathbf{S}}(\mathbf{S}) = \begin{cases}|\det\mathbf{J}|\,f_{\mathbf{G}}(\mathbf{S}^2)\,, & \mathbf{S} > \mathbf{0}, \\ 0\,, & \text{else},\end{cases} \qquad (13)$$

where $\mathbf{J} = \frac{\partial\mathbf{S}^2}{\partial\mathbf{S}}$. According to [13], $|\det\mathbf{J}| = \prod_{k\le l}(d_k + d_l)$, where d_k denotes the singular values of the channel matrix $\mathbf{H}$. Unfortunately, there seems to be no general formula expressing this Jacobian in terms of the elements of $\mathbf{S}$ for any M, i.e., $|\det\mathbf{J}|$ must be calculated for any M of interest separately. Once this has been done, the expectation in (12) can be calculated to obtain the average PEP.

We illustrate the procedure for $M = 2$. Here, the matrix $\mathbf{S}$ has four degrees of freedem, i.e.,

$$\mathbf{S} = \begin{pmatrix} s_{11} & |s_{12}|e^{j\phi_{12}} \\ |s_{12}|e^{-j\phi_{12}} & s_{22}\end{pmatrix}.$$

The partial derivative $\partial\mathbf{S}$ reduces to the derivatives with respect to s_{11}, $|s_{12}|$, ϕ_{12}, and s_{22}. Straightforward calculation yields the Jacobian $|\det\mathbf{J}| = 8(s_{11}+s_{22})^2\big(|s_{12}|^2 - s_{11}s_{22}\big)$. The pdf of s_{11} is obtained by calculating the marginal of (13) with respect to s_{11}. The side-constraint $\mathbf{S} > \mathbf{0}$ is taken into account by integrating within the range $s_{22} > 0$, $|s_{12}| \le \sqrt{s_{11}s_{22}}$, and $-\pi < \phi_{12} \le \pi$. This yields

$$\begin{aligned}f(s_{11}) = \frac{1}{4}&\exp\big(-s_{11}^2\big)\big(2s_{11} + 4s_{11}^3 + \sqrt{\pi}(2s_{11}^2 - 1)\big) \\ &+ \frac{\sqrt{\pi}}{2}\mathrm{Q}\Big(\sqrt{2}s_{11}\Big)\,.\end{aligned}$$

[2]The relation $x \sim y$ means that x and y have the same distribution.

Plugging this expression into (12) yields

$$P(\mathbf{x}_1 \to \mathbf{x}_2) \leq \int_{-\infty}^{\infty} \exp\!\left(\frac{-s_{11}^2\|\boldsymbol{\delta}\|^2}{4\sigma_n^2}\right) f(s_{11})\, ds_{11}$$
$$\leq \frac{1}{4}\Bigg[\frac{32}{(4+\gamma)^2} + \frac{4\pi}{(4+\gamma)^{\frac{3}{2}}} + \frac{4}{4+\gamma} - \frac{\pi}{\sqrt{4+\gamma}} + \frac{1}{\sqrt{\gamma}}\left(\pi + j\mathrm{B}\!\left(-\frac{4}{\gamma}, \frac{1}{2}, 0\right)\right)\Bigg],$$

where $\mathrm{B}(z,a,b) = \int_0^z u^{a-1}(1-u)^{b-1}\, du$ is the incomplete Beta function and $\gamma = \|\boldsymbol{\delta}\|^2/\sigma_n^2$. Computing the limit in (4) finally yields a diversity of $d = \frac{3}{2}$ which is exactly halfway between the full diversity of 2 and the minimum diversity of 1, in accordance with our conjecture.

3.2. Complexity

We next investigate the impact of α on the complexity of the mismatched ML detector when the latter is implemented via the sphere decoding (SD) algorithm [2] (the complexity of the partial equalizer can be neglected). We expect that complexity decreases with α since the partial equalization stage improves the channel condition number, i.e., $\kappa(\mathbf{C}_\alpha) = \kappa^{1-\alpha}(\mathbf{H})$, and small channel condition numbers were previously observed to result in smaller SD complexity [9]. However, using the standard version of the SD it turned out that this is true only for small α. For larger α, the partial equalizer causes strong noise enhancement. In this case, the SD typically descends directly to the optimal solution but continues traversing the tree back and forth in the search for an even better solution.

To exploit the improved condition number for large α, in spite of the strong noise enhancement, we propose to use the SD variant that has been proposed in [14] and is suited even for low-SNR situations. We next review the main idea of this SD variant using a slightly simpler derivation.

With SD, the QR decomposition $\mathbf{C}_\alpha = \mathbf{QR}$ of the equivalent channel $\mathbf{C}_\alpha$ is used to rewrite the metric in (10) as

$$\left\|\tilde{\mathbf{r}} - \mathbf{C}_\alpha \mathbf{x}\right\|^2 = \left\|\mathbf{q} - \mathbf{R}\mathbf{x}\right\|^2 \tag{14}$$

where $\mathbf{q} = \mathbf{Q}^H\tilde{\mathbf{r}}$. The upper triangular matrix $\mathbf{R}$ corresponds to a tree that is traversed to find the ML solution. In the tree traversal, only nodes corresponding to solutions that lie within a sphere of radius R about the receive vector are visited. To further reduce the number of nodes visited, [14] proposed a tree pruning as follows. First split (14) into two terms including, respectively, the first $k-1$ and the last $M-k+1$ layers,

$$\left\|\mathbf{q} - \mathbf{R}\mathbf{x}\right\|^2 = \left\|\mathbf{q}_k - [\mathbf{R}_k^1\, \mathbf{R}_k^2]\mathbf{x}_k\right\|^2 + \left\|\mathbf{q}^k - \mathbf{R}^k\mathbf{x}^k\right\|^2$$

where $\mathbf{q}_k = [q_1, \cdots, q_{k-1}]$, $\mathbf{q}^k = [q_k, \cdots, q_M]$, $\mathbf{x}_k = [x_1, \cdots, x_{k-1}]$, $\mathbf{x}^k = [x_k, \cdots, x_M]$, and $\mathbf{R}_k^1 = [\mathbf{R}]_{ij}$ for $i,j \in \{1,\ldots,k-1\}$, $\mathbf{R}_k^2 = [\mathbf{R}]_{ij}$ for $i \in \{1,\ldots,k-1\}$ and $j \in \{k,\ldots,M\}$, $\mathbf{R}^k = [\mathbf{R}]_{ij}$ for $i,j \in \{k,\ldots,M\}$, denote the north-west, north-east, and south-east corners of $\mathbf{R}$, respectively. At the kth layer, the sphere bound $\left\|\mathbf{q}-\mathbf{R}\mathbf{x}\right\|^2 \leq R^2$ now amounts to

$$\left\|\mathbf{q}^k - \mathbf{R}^k\mathbf{x}^k\right\|^2 \leq R^2 - \left\|\mathbf{q}_k - [\mathbf{R}_k^1\, \mathbf{R}_k^2]\mathbf{x}_k\right\|^2 \leq R^2 - \rho_k^2. \tag{15}$$

The lower bound ρ_k^2 for $\left\|\mathbf{q}_k - [\mathbf{R}_k^1\, \mathbf{R}_k^2]\mathbf{x}_k\right\|^2$ is obtained via the Rayleigh-Ritz theorem [8] as

$$\rho_k^2 = \sigma_{\min}^2(\mathbf{R}_k^1)\, \|\mathcal{Q}_{\mathcal{A}}\{\mathbf{b}_k\} - \mathbf{b}_k\|^2,$$

with $\mathbf{b}_k = (\mathbf{R}_k^1)^{-1}(\mathbf{q}_k - \mathbf{R}_k^2\mathbf{x}^k)$ and $\sigma_{\min}(\mathbf{R}_k^1)$ denoting the minimum singular value of $\mathbf{R}_k^1$. According to (15), at each layer a reduced sphere radius is used that allows for a significant tree pruning and hence smaller number of visited nodes.

We analyze the behavior of this modified SD for $\alpha = 1$ (ZF detection), where the complexity (in terms of nodes visited) of the standard SD is very high due to strong noise enhancement. In contrast, with the tree-pruning variant only the nodes associated to the ZF solution are visited. For $\alpha = 1$, $\mathbf{C}_\alpha = \mathbf{Q} = \mathbf{R} = \mathbf{I}$, $\mathbf{b}_k = \mathbf{q}_k$, and thus (15) simplifies to

$$\|\mathbf{q}^k - \mathbf{x}^k\|^2 \leq R^2 - \|\mathcal{Q}_{\mathcal{A}}\{\mathbf{q}_k\} - \mathbf{q}_k\|^2, \tag{16}$$

where the lower bound is $\rho_k^2 = \|\mathcal{Q}_{\mathcal{A}}\{\mathbf{q}_k\} - \mathbf{q}_k\|^2$. When the initial sphere radius R is found via the ZF solution, i.e. $R = \|\mathcal{Q}_{\mathcal{A}}\{\mathbf{q}\} - \mathbf{q}\|$, the upper bound in (16) reduces to $\|\mathcal{Q}_{\mathcal{A}}\{\mathbf{q}^k\} - \mathbf{q}^k\|^2$ which equals the (partial) distance of the ZF solution from the observation. Since for $\alpha = 1$ the ZF solution is optimal, there is no other solution within the sphere and hence no other nodes than those associated to the ZF solution are visited.

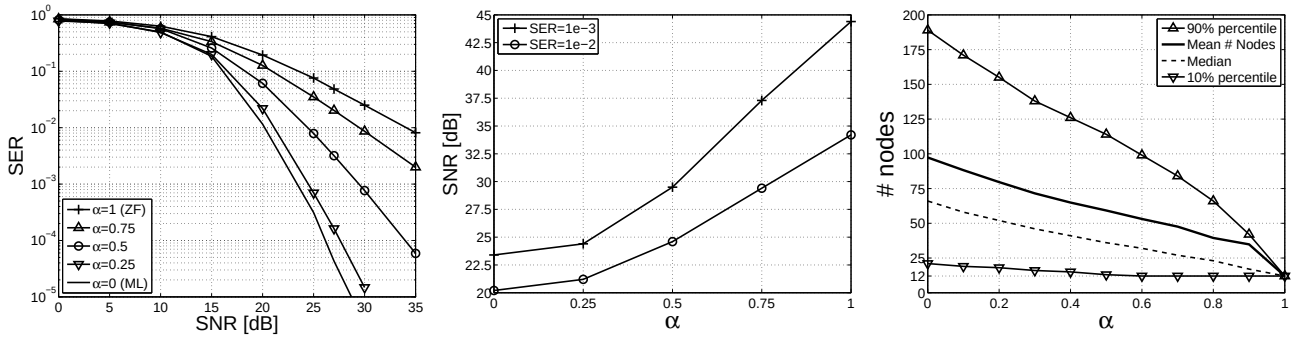
We assessed the complexity of our proposed receiver via extensive numerical simulations (see Section 4 for examples). To this end, we computed empirical distributions of the number of nodes visited by the SD from a huge number of realizations of an i.i.d. channel. We denote by n either an arbitrary percentile of this distribution or its mean. All our results suggested that the relation between the number n of visited nodes and the parameter α can be reasonably approximated as

$$n \approx \alpha\, n_{\mathrm{ZF}} + (1-\alpha)\, n_{\mathrm{ML}}. \tag{17}$$

Hence, the number of nodes visited for a certain α is essentially a convex combination of the number of nodes n_{ML} visited in the case of ML detection and the number of nodes n_{ZF} visited with ZF detection, respectively. Defining the complexity savings over ML detection as $\Delta n = n_{\mathrm{ML}} - n$, we obtain $\Delta n \approx \alpha(n_{\mathrm{ML}} - n_{\mathrm{ZF}})$, i.e., the savings increase roughly linearly with α.

4. SIMULATION RESULTS

In this section, we provide exemplary simulation results illustrating the performance and complexity of the proposed receiver. We considered uncoded 4×4 and 6×6 MIMO spatial



(a) (b) (c)

Fig. 2. *Performance and complexity of proposed receiver: (a) SER versus SNR for various α for $M=6$, (b) required SNR versus α for various target SERs ($M=4$), (c) number of visited nodes (10%, 50%, and 90% percentile and mean) versus α for $M=6$.*

multiplexing system with 16-QAM symbol alphabet. All results were obtained using between 500 and $5 \cdot 10^6$ realizations of an i.i.d. Rayleigh fading channel.

Fig. 2(a) shows symbol error rate (SER) versus signal-to-noise ratio $\text{SNR} = \frac{M}{\sigma_n^2}$ for $\alpha \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ and $M = 4$. Close inspection reveals that the diversity achieved with different α is in agreement with our conjecture in Section 3.

In Fig. 2(b), the SNR required for a SER of 10^{-3} and 10^{-2} is depicted versus α (again, $M = 4$). It is seen that particularly for small α (up to $\alpha = 0.25$) the SNR penalty with respect to ML detection is rather small (i.e., only 1 dB for $\alpha = 0.25$). Simulations with more antennas showed that this effect is even more pronounced for larger MIMO systems.

Finally, Fig. 2(c) shows the number of visited nodes (mean as well as 10%, 50%, and 90% percentile) versus α for $M = 6$. The number of visited nodes decreases essentially linearly with increasing α in all cases. All curves are in good agreement with our rule of thumb (17).

5. CONCLUSION

Motivated by a unifying formulation of ML and ZF detection, we proposed a novel receiver that consists of a partial equalizer that improves the channel condition number, followed by a mismatched ML detector that ignores the noise correlation resulting from the first stage. The actual implementation uses a sphere decoder variant specifically suited for low SNR situations. This is necessary since the partial equalizer can cause severe noise enhancement. Theoretical investigations and numerical simulations showed that our receiver allows a continuous tradeoff between diversity and complexity, and can thus be adjusted to any desired level of performance or complexity. For fixed target symbol error rate, we observed that significant complexity savings can be achieved at the cost of a very small SNR penalty.

ACKNOWLEDGMENT

The authors thank Arnold Neumaier for helpful discussions.

REFERENCES

[1] A. Paulraj, R. U. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge Univ. Press, Cambridge (UK), 2003.

[2] U. Fincke and M. Phost, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Math. Comp.*, vol. 44, pp. 463–471, April 1985.

[3] 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.

[4] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communications architecture," *Elect. Lett.*, vol. 35, pp. 14–16, Jan. 1999.

[5] W.-J. Choi, R. Negi, and J. M. Cioffi, "Combined ML and DFE decoding for the V-BLAST system," in *Proc. IEEE ICC-00*, New Orleans, LA, June 2000, pp. 18–22.

[6] E. Biglieri, G. Taricco, and A. Tulino, "Performance of space-time codes for a large number of antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 1794–1803, July 2002.

[7] Dhananjay A. Gore, Jr. Robert W. Heath, and Arogyaswami J. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Comm. Letters*, vol. 6, no. 11, pp. 491–493, November 2002.

[8] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge (UK), 1999.

[9] H. Artés, D. Seethaler, and F. Hlawatsch, "Efficient detection algorithms for MIMO channels: A geometrical approach to approximate ML detection," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2808–2820, Nov. 2003.

[10] L. L. Scharf, *Statistical Signal Processing*, Addison Wesley, Reading (MA), 1991.

[11] T. Ratnarajah, R. Vaillancourt, and M. Alvo, "Complex random matrices and rayleigh channel capacity," *Communication in Information and Systems*, vol. 3, no. 2, pp. 119–138, October 2003.

[12] Robb J. Muirhead, Ed., *Aspects of Multivariate Statistical Theory*, Wiley, Hoboken, New Jersey, 2005.

[13] Ingram Olkin and Herman Rubin, "Multivariate beta distributions and independence properties of the wishart distribution," *Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 261–269, March 1964.

[14] M. Stojnic, H. Vikalo, and B. Hassibi, "A branch and bound approach to speed up the sphere decoder," in *Proc. IEEE ICASSP 2005*, Philadelphia, USA, May 2005, pp. 429–433.