Optimum Receivers and Low-Dimensional Spreaded Modulation for Multiuser Space–Time Communications

Matthias Brehler and Mahesh K. Varanasi

Abstract—**The jointly optimum receiver is obtained for multiuser communications in a frequency non-selective Rayleigh fading channel with N_{T} transmit antennas per user and N_{R} receive antennas. Based on a general analysis of quadratic receivers in zero-mean complex Gaussian vectors, asymptotically tight expressions (for high SNR) for the pairwise error probabilities are derived. Subsequently, it is shown that N_{T}-dimensional single-user signaling suffices to provide full diversity order $N = N_{\mathrm{T}}N_{\mathrm{R}}$ for all the users. In other words, the presence of other users does not increase the minimum dimension required beyond what is needed for the single-user space–time channel.**

For the special case of low-rank "CDMA" signaling with $N_{\mathrm{T}} = 1$ and provided the signatures of any two users are linearly independent, it is shown that the error probability of a K-user system asymptotically approaches single-user like performance for every user. Remarkably therefore, an increase in the number of users, and hence an increase in the aggregate spectral efficiency, does not require the users to transmit with more power to achieve single-user like performance asymptotically. A signal design algorithm is proposed to illustrate this point and examples are given. These results are then generalized to the multiple transmit antenna case. A new $(N_{\mathrm{T}} + 1)$-dimensional signaling strategy is proposed for the multiuser channel that leverages existing single-user space–time signal designs while ensuring full diversity order and single-user like performance asymptotically for every user.

Index Terms—**Asymptotic efficiency, asymptotics, CDMA, error analysis, fading channels, maximum likelihood receiver, multiuser communications, space–time modulation.**

I. Introduction

MULTIPLE antenna communication has received considerable attention in recent years due in large part to the information theoretic work in [1, 2], which showed that the use of multiple transmit and receive antennas could achieve considerable gains on the Rayleigh fading channel when the receiver has perfect side information about the channel state. Motivated by these promises, several researchers have recently proposed multi-antenna coding and modulation schemes for coherent single-user channels (cf. [3–7]) to show that diversity communication systems, when designed intelligently, can yield significant improvements over single antenna channels. The information theory of the single-user space–time channel easily extends to the multiuser multi-antenna channel [1] and it can be inferred that the gains in the capacity region are every bit as dramatic for the multiuser channel as well.

In this paper, we present a theory of modulation and detection for *multiuser* space–time communication. Each user can employ N_{T} transmit antennas and is detected by a base-station with N_{R} receive antennas. We develop multiuser modulation schemes that guarantee each user full order of diversity $N = N_{\mathrm{T}}N_{\mathrm{R}}$.

When only receive antenna diversity is employed, it is well known that in the multiuser narrowband channel (in which all users modulate an identical waveform) the jointly optimum detector achieves the same order of diversity for every user as in a single user channel [8]. This is illustrated in Figure 1, which shows the bit error rate (BER) for one, five, and ten users, each employing BPSK modulation, and that are detected with $N_{\mathrm{R}} = 2$ receive antennas. The optimum multiuser detector achieves the same diversity order for every user as is achievable in a single user channel, but it incurs a penalty in terms of signal-to-noise ratio (SNR), when compared to single-user performance. In the example, for ten users the SNR gap to the single user is close to 8 dB at a BER of 10^{-2}.

For the special case of one transmit antenna per user we present a method for modulation that spreads the symbol of each user with a unit-norm spreading sequence of very low dimension ($D \geq 2$) and wherein these spreading sequences are designed to guarantee asymptotic (high SNR) single-user like performance for every user. By an "asymptotic single-user like" performance we mean that for high SNR the upper bound on the multiuser BER converges to the upper bound on the single-user BER. In particular, the lowest admissible (complex) signal space dimension $D = 2$ appears to be the most desirable (in terms of energy vs spectral efficiency tradeoff), independently of the number of users. The signaling scheme that uses the corresponding signature signals, which are "optimized" according to a signal design criterion we develop in this paper, may be thought of as *enlightened* or low-rank CDMA signaling in that the spread factor is just two.

In the case of binary modulation, when the upper bound on the single-user BER coincides with its lower bound, the optimum multiuser detector achieves exactly single-user performance asymptotically, i.e., the asymptotic efficiency of all users is unity. In other words, the signal design enables an increase of aggregate spectral efficiency of the multiuser system due to an increase in the numbers of users with no loss in energy efficiency for sufficiently high SNR, when compared to a single user. Using these designs, we next extend the above example and see that ten users can be detected with a BER that is within 2 dB of the single-user performance at a BER of 10^{-2}; at SNRs above 12 dB, the performance of the multiuser system is virtually indistinguishable from that of the single user (Fig. 4).

To obtain these results and their extensions to multiple transmit antennas per user, we use the asymptotic analysis in [9],

Sumitted July 31, 2001; revised August 8, 2002. This work was supported in part by NSF grants CCK-0112977 and ANIR-9725778. The material in this paper was presented in part at the Conference on Information Sciences and Systems, Baltimore, MD, March 2001, the IEEE Information Theory Workshop, Cairns, Australia, Sept. 2001, and the IEEE International Symposium on Information Theory, Lausanne, Switzerland, July 2002. The authors are with the Department of Electrical and Computer Engineering, University of Colorado, Boulder, CO 80309-0425 (e-mail: brehler@dsp.colorado.edu, varanasi@colorado.edu).

where we develop a general theory for the analysis of quadratic receivers in Rayleigh fading channels and apply it to single-user coherent and noncoherent space–time receivers. Independently from that work, [10] analyzes the optimum coherent receiver for the single-user space–time channel and also obtains asymptotically tight pairwise error probabilities for the single-user case. In this paper, we start-out with a general synchronous multiuser modulation model and apply the methodology of [9] to arrive at asymptotic tight expressions for pairwise error probabilities of modulation schemes and channels not previously considered and/or analyzed. For example, for the narrowband multiuser channel and optimum detection, the new asymptotic pairwise error probabilities generalize the corresponding result in [8], in that more general modulation schemes and possibly correlated diversity branches are included. Even the single-user space–time case profits from our general expressions, because an asymptotically tight approximation of the pairwise error probability for correlated fading and non-full rank codes are obtained (recall that [4, 6] consider the Chernoff bound in i.i.d. fading, [9] only full-rank codes). For the multiuser channel, the asymptotic pairwise error probabilities are subsequently used to derive signal design criteria which in turn enable us to design signal sets for both, single and multiple transmit antenna systems (per user), that guarantee asymptotic single-user like performance for every user. As pointed out above, a remarkable feature of these signal sets is that the minimum required signal space dimension is independent of the number of users for sufficiently high SNR.

Interestingly, in the multiple transmit antenna per user problem, if every user transmits symbols from a full diversity order $N = N_{\mathrm{T}} N_{\mathrm{R}}$, single-user space–time constellation[1] in a multiuser channel, we show that each user achieves the same full diversity order N in spite of the presence of other users. Recall, that to achieve full order of diversity in the single-user channel $D \geq N_{\mathrm{T}}$ signal space dimensions are necessary [4, 6, 9], so that $D \geq N_{\mathrm{T}}$ is also the condition required for every user in a multiuser channel to achieve the full diversity order N. However, with an increase of the number of users, an increasing SNR penalty is incurred, just like in the narrowband case with one transmit antenna.

The modulation method we propose for $N_{\mathrm{T}} > 1$ leverages existing single-user space–time constellations and alleviates the SNR penalty asymptotically and requires at least (only) $N_{\mathrm{T}} + 1$ dimensions. Before transmitting, the single-user space–time symbols of each user are "spread" by a low-dimensional "spreading matrix." This can be thought of as a generalization of the low-rank CDMA signaling method described above to multiple antennas per user. In particular, we present a criterion to design each user's low-dimensional spreading matrix that depends only on the dimensionality of the chosen single-user space–time constellation. Consequently, different spectral efficiencies can be attained without changing or redesigning the spreading matrices.

In all our examples we keep the number of dimensions as close as possible to the minimum necessary, so that our modulation/detection perspective suffices, since such "short" signal matrices are more amenable to a (super-) symbol interpretation. Our results on performance analysis, however, are also valid for large number of dimensions, as in systems where each user may employ single-user (or other, yet to be designed) space–time trellis codes [4, 6]. All results are derived for (jointly) optimum detection, which can be efficiently implemented by closest point search algorithms for (general) lattices [12]. In particular, the generalized sphere decoder of [13] was used to obtain some of the numerical results.

In Section II, we describe a general K user, N_{T} transmit, N_{R} receive antenna symbol-synchronous system model. Based on our general results on the asymptotic analysis of quadratic receivers in complex zero-mean Gaussian vectors [9], we analyze, in Section III, the optimum multiuser receiver and obtain asymptotically tight expressions for the pairwise error probabilities. In Section IV we interpret these probabilities for the single transmit (IV-A) and multiple transmit antennas (IV-B) cases in terms of the minimum dimension needed to achieve full order of diversity for all users. We also propose new optimized multiuser signal design strategies that leverage single-user space–time designs in order to deliver single-user like performance in the high SNR regime. In Section IV-C, we investigate conditions to re-write the system model to make it amenable to (generalized) sphere-decoding. While most of the paper focuses on uplink communications (users to a central base-station), the system model and the general analysis of the pairwise error probability can be easily adapted to downlink (base-station to users) communications as well, and we do this in Section V. We conclude in Section VI. *Notation:* Throughout the paper $^{\mathsf{T}}$ denotes transpose and † complex conjugate transpose. The multi-variate circularly symmetric, complex Gaussian distribution with mean-vector $\mathbf{m}$ (and covariance matrix $\mathbf{K}$) is denoted by $\mathcal{CN}(\mathbf{m})$ ($\mathcal{CN}(\mathbf{m}, \mathbf{K})$). $E[\cdot]$ denotes the expected value of the expression in brackets. For any matrix $\mathbf{A}$ we write its determinant as $|\mathbf{A}|$ and its trace as $\operatorname{tr}(\mathbf{A})$. The product of the non-zero eigenvalues of a matrix $\mathbf{A}$ is denoted by $|\mathbf{A}|_{\mathsf{NZ}}$. For any vector $\mathbf{a}$, we write its ℓ_2 norm as $\sqrt{\mathbf{a}^{\dagger}\mathbf{a}} = \|\mathbf{a}\|$. The logarithm to the base b is denoted by $\log_b$, the natural logarithm by $\ln$. The Kronecker (or tensor) product of two matrices is denoted by $\otimes$.

II. Multi-Antenna, Multiuser Discrete Time System Model

We describe a system model for K users communicating simultaneously in a common D–dimensional signal space.[2] Each of the K users employs N_{T} transmit antennas to send information symbol-synchronously to an N_{R} receive antenna array of the base-station (if perfect synchronization proves difficult and/or there are multiple fading paths, [14] offers a way to construct a robust basis). Since there are N_{T} transmit antennas and D dimensions, each user transmits one out of M possible $D \times N_{\mathrm{T}}$ complex-valued signal matrices, drawn from the set $\mathcal{S}_k = \left\{\mathbf{S}_{k1}, \mathbf{S}_{k2}, \ldots, \mathbf{S}_{kM}\right\}$ with $\mathbf{S}_{km} \in \mathbb{C}^{D \times N_{\mathrm{T}}}$. The signal

[1] In this paper, we take a perspective of modulation and detection and thus consider the transmitted signal matrices, originating for example from the Alamouti scheme [11], as space–time (super-) symbols originating from a constellation rather than a codeword of a (block-) code.

[2] We suggest the term "space-dimension" communication rather than "space–time" communications. The latter implies a basis of time-translates of a single waveform (so that D corresponds to the length of the coherence interval in symbol durations), which is restrictive as pointed out in [9].

matrices $\mathbf{S}_{km}$ may be thought of as space–time (block-) codewords where each element of the matrix is drawn from a finite, QAM-like constellation with $\mathcal{S}_k$ being the kth user's codebook, or they may be thought of as super-symbols of some arbitrary constellation $\mathcal{S}_k$. Hence, the receiver is a decoder or a detector in the two cases, respectively. We will refer to the matrices $\mathbf{S}_{km}$ as (super-) symbols or signals or codewords as is appropriate and use the general term receiver when the terms detector or decoder are both applicable. To succinctly write the discrete-time model for this system we need more definitions.

Let H_i denote the ith hypothesis with $1 \leq i \leq M^K$. Without loss of generality let i determine uniquely the K-tuple $(i_1, i_2, \ldots, i_K)$ according to

$$i = \sum_{k=1}^{K} (i_k - 1) M^{k-1} + 1.$$

We let hypothesis H_i denote that user k transmits the signal $\mathbf{S}_{k i_k}$ for each k. Define the $D \times K N_{\mathrm{T}}$ matrix of signals corresponding to hypothesis H_i as $\mathbf{F}_i = [\mathbf{S}_{1 i_1}, \mathbf{S}_{2 i_2}, \ldots, \mathbf{S}_{K i_K}]$. Thus the discrete-time model for the nth receive antenna can be written as

$$\mathbf{y}_n = \mathbf{F}_i \mathbf{W}^{1/2} \widehat{\mathbf{h}}_n + \boldsymbol{\eta}_n, \tag{1}$$

where $\mathbf{y}_n$ is the D-dimensional vector of observations, $\mathbf{W} = \mathrm{diag}\left\{ w_1, w_2, \ldots, w_K \right\} \otimes \mathbf{I}_{N_{\mathrm{T}}}$ with w_k being the kth user's average transmitted symbol energy, $\widehat{\mathbf{h}}_n^{\mathsf{T}} = \left[\mathbf{h}_{1n}^{\mathsf{T}},\, \mathbf{h}_{2n}^{\mathsf{T}}, \ldots,\, \mathbf{h}_{Kn}^{\mathsf{T}} \right]$ is a $K N_{\mathrm{T}}$-dimensional vector of $\mathcal{CN}(\mathbf{0})$ distributed fading coefficients with $\mathbf{h}_{kn}$ containing the N_{T} fading coefficients from the kth user's transmit antennas to receive antenna n, and $\boldsymbol{\eta}_n$ is the D–dimensional $\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_D)$ distributed additive noise vector. To obtain the sufficient statistics for all N_{R} receive antennas, we simply stack the $\mathbf{y}_n$ to obtain

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{N_{\mathrm{R}}} \end{bmatrix} = \left(\mathbf{I}_{N_{\mathrm{R}}} \otimes \mathbf{F}_i \mathbf{W}^{1/2} \right) \widehat{\mathbf{h}} + \boldsymbol{\eta}, \tag{2}$$

where $\widehat{\mathbf{h}}^{\mathsf{T}} = \left[\widehat{\mathbf{h}}_1^{\mathsf{T}}, \ldots, \widehat{\mathbf{h}}_{N_{\mathrm{R}}}^{\mathsf{T}} \right]$ contains the fading coefficients in "receiver-antenna" order. For the analysis to come, it will be more convenient to organize the fading coefficients user-wise in the vector $\mathbf{h}^{\mathsf{T}} = \left[\mathbf{h}_1^{\mathsf{T}}, \ldots, \mathbf{h}_K^{\mathsf{T}} \right]$, where $\mathbf{h}_k^{\mathsf{T}} = \left[\mathbf{h}_{k1}^{\mathsf{T}}, \ldots, \mathbf{h}_{kN_{\mathrm{R}}}^{\mathsf{T}} \right]$, which also requires the introduction of $\boldsymbol{\mathcal{S}}_{km} = \mathbf{I}_{N_{\mathrm{R}}} \otimes \mathbf{S}_{km}$, $\boldsymbol{\mathcal{F}}_i = \left[\boldsymbol{\mathcal{S}}_{1 i_1}, \boldsymbol{\mathcal{S}}_{2 i_2}, \ldots, \boldsymbol{\mathcal{S}}_{K i_K} \right]$, and $\boldsymbol{\mathcal{W}} = \mathbf{W} \otimes \mathbf{I}_{N_{\mathrm{R}}}$. With these definitions the $D N_{\mathrm{R}}$ sufficient statistics can be written as

$$\mathbf{y} = \boldsymbol{\mathcal{F}}_i \boldsymbol{\mathcal{W}}^{1/2} \mathbf{h} + \boldsymbol{\eta}. \tag{3}$$

We denote the correlation matrix of the fading coefficients as $\boldsymbol{\Sigma} = E\left[\mathbf{h}\mathbf{h}^{\dagger} \right]$. $\boldsymbol{\Sigma}_{kk} = E\left[\mathbf{h}_k \mathbf{h}_k^{\dagger} \right]$ denotes the kth diagonal $N \times N$ block of $\boldsymbol{\Sigma}$ and thus is the kth user's fading correlation matrix. The signals and fading processes are normalized so that $\overline{\gamma}_k = w_k / \sigma^2$ represents the kth user's average transmitted symbol energy to noise variance ratio, which in single-user channels is often denoted as E_S/N_0 (we set $\sigma^2 = N_0$ in plots). We refer to $\overline{\gamma}_k$ as the signal-to-noise ratio (SNR) in the text, but caution the reader that the often-used average received SNR of user k equals $\overline{\gamma}_k / D$.

Specifically, each user's signals are normalized such that their average energy is unity, i.e.,

$$E\left[\mathrm{tr}\left(\mathbf{S}^{\dagger}_{km} \mathbf{S}_{km} \right) \right] = 1 \;\forall\, k, \tag{4}$$

where the expected value is taken over m. The fading coefficients are normalized such that

$$E\left[\mathbf{h}_k^{\dagger} \boldsymbol{\mathcal{S}}^{\dagger}_{km} \boldsymbol{\mathcal{S}}_{km} \mathbf{h}_k \right] = N_{\mathrm{R}} \;\forall\, k. \tag{5}$$

For equi-probable symbols this condition can be written as

$$\sum_{m=1}^{M} \mathrm{tr}\left(\boldsymbol{\Sigma}_{kk} \boldsymbol{\mathcal{S}}^{\dagger}_{km} \boldsymbol{\mathcal{S}}_{km} \right) = M N_{\mathrm{R}}. \tag{6}$$

In i.i.d. fading the stated conditions lead to $\boldsymbol{\Sigma}_{kk} = \mathbf{I}_N$.

III. Optimum Receiver and Analysis

We consider coherent detection and assume that the fading coefficients are perfectly known at the receiver. To estimate the fading coefficients with sufficient accuracy in practice, the channel coherence time must be long relative to the symbol-duration T. In this paper we focus on the ideal coherent case and express the optimum receiver in terms of a quadratic form in the observations and the fading coefficients. This formulation allows us to make use of the general results of [9] for the asymptotically tight analysis of the pairwise error probabilities.

A. Maximum Likelihood Receiver

The likelihood function of the sufficient statistics $\mathbf{y}$ given the fading coefficients $\mathbf{h}$ and the true hypothesis H_i (i.e. $\mathbf{F}_i$, is transmitted) is

$$p\left(\mathbf{y} \middle| \mathbf{h}, H_i \right) = \frac{\exp\left(-\sigma^{-2} \left\| \mathbf{y} - \boldsymbol{\mathcal{F}}_i \boldsymbol{\mathcal{W}}^{1/2} \mathbf{h} \right\|^2 \right)}{\pi^{D N_{\mathrm{R}}} \sigma^{2 D N_{\mathrm{R}}}}. \tag{7}$$

Defining the new $(K N_{\mathrm{T}} + D) N_{\mathrm{R}}$–dimensional sufficient statistic $\mathbf{z}^{\mathsf{T}} = \sigma^{-1} \begin{bmatrix} \mathbf{h}^{\mathsf{T}} & \mathbf{y}^{\mathsf{T}} \end{bmatrix}$ and the matrix

$$\mathbf{Q}_i = \begin{bmatrix} \boldsymbol{\mathcal{W}}^{1/2} \boldsymbol{\mathcal{F}}_i^{\dagger} \boldsymbol{\mathcal{F}}_i \boldsymbol{\mathcal{W}}^{1/2} & -\boldsymbol{\mathcal{W}}^{1/2} \boldsymbol{\mathcal{F}}_i^{\dagger} \\ -\boldsymbol{\mathcal{F}}_i \boldsymbol{\mathcal{W}}^{1/2} & \mathbf{0}_{D N_{\mathrm{R}}} \end{bmatrix}, \tag{8}$$

the jointly optimum coherent receiver Φ can be expressed as

$$\Phi : \hat{i} = \arg \min_{1 \leq i \leq M^K} \mathbf{z}^{\dagger} \mathbf{Q}_i \mathbf{z} = \arg \min_{1 \leq i \leq M^K} \delta_i, \tag{9}$$

where δ_i is defined implicitly. Note that the sufficient statistics $\mathbf{z}$ are $\mathcal{CN}(\mathbf{0},\, \mathbf{K}_{\mathbf{zz}|H_i})$ distributed, where

$$\begin{aligned} \mathbf{K}_{\mathbf{zz}|H_i} &= E\left[\mathbf{z}\mathbf{z}^{\dagger} \right] \\ &= \begin{bmatrix} \sigma^{-2} \boldsymbol{\Sigma} & \sigma^{-2} \boldsymbol{\Sigma} \boldsymbol{\mathcal{W}}^{1/2} \boldsymbol{\mathcal{F}}_i^{\dagger} \\ \sigma^{-2} \boldsymbol{\mathcal{F}}_i \boldsymbol{\mathcal{W}}^{1/2} \boldsymbol{\Sigma} & \sigma^{-2} \boldsymbol{\mathcal{F}}_i \boldsymbol{\mathcal{W}}^{1/2} \boldsymbol{\Sigma} \boldsymbol{\mathcal{W}}^{1/2} \boldsymbol{\mathcal{F}}_i^{\dagger} + \mathbf{I} \end{bmatrix}. \end{aligned} \tag{10}$$

B. Bounds on Symbol and Bit Error Rate

Let $\mathcal{E}_k(\Phi)$ denote the event that the receiver Φ detects user k erroneously. Then $\Pr\left\{\mathcal{E}_k(\Phi)\middle| H_i\right\}$ is the symbol error probability of the kth user detected by receiver Φ conditioned on the hypothesis H_i. It is the probability of the union of the corresponding $(M-1)M^{K-1}$ possible events of the form $\left\{\delta_j < \delta_i\right\}$. Since the probability of the union is usually not computable, consider the union upper bound, which is the sum of the pairwise error probabilities $\Pr\left\{\delta_j < \delta_i\right\}$.[3] A lower bound is obtained by considering the pairwise probability $\Pr\left\{\delta_{\hat{j}} < \delta_i\right\}$, where $H_{\hat{j}}$ corresponds to one of the $M-1$ hypotheses H_j that result in an error only for user k when compared to H_i. The lower bound can be tightened by choosing $H_{\hat{j}}$ such that $\Pr\left\{\delta_{\hat{j}} < \delta_i\right\}$ is maximized.

The kth user's symbol error rate (SER) $\Pr\left\{\mathcal{E}_k(\Phi)\right\}$ for equiprobable symbols is then bounded as

$$\begin{aligned}\Pr\left\{\mathcal{E}_k(\Phi)\right\} &= M^{-K}\sum_{i=1}^{M^K}\Pr\left\{\mathcal{E}_k(\Phi)\middle| H_i\right\} && (11)\\ &\le M^{-K}\sum_{i=1}^{M^K}\sum_{\forall j\in\Lambda_i(k)}\Pr\left\{\delta_j<\delta_i\right\}, && (12)\\ \Pr\left\{\mathcal{E}_k(\Phi)\right\} &\ge M^{-K}\sum_{i=1}^{M^K}\Pr\left\{\delta_{\hat{j}}<\delta_i\right\}, && (13)\end{aligned}$$

where $\Lambda_i(k)$ is the set of the $(M-1)M^{K-1}$ indices of hypotheses in which the kth user's symbol differs from its symbol corresponding to the true hypothesis H_i.

To obtain bounds on the average bit error rate (BER) P_k^{b} of the kth user, we introduce the event $H_i \to H_j$ that hypothesis H_i is detected as H_j (in the presence of all other hypotheses). Since the events $H_i \to H_j$ are mutually exclusive, the average bit error rate can be written as

$$P_k^{\mathrm{b}} = M^{-K}\sum_{i=1}^{M^K}\sum_{\forall j\in\Lambda_i(k)}\frac{b_{ij}(k)}{\log_2 M}\Pr\left\{H_i \to H_j\right\}, \qquad (14)$$

where $b_{ij}(k)$ is the number of erroneously detected bits of user k, when hypothesis H_i is detected as H_j. An upper bound on P_k^{b} is obtained by upper-bounding the probabilities $\Pr\left\{H_i \to H_j\right\}$ by $\Pr\left\{\delta_j < \delta_i\right\}$. A lower bound on P_k^{b} is obtained by lower bounding $b_{ij}(k)$ by one and using the fact that the inner sum of probabilities is equal to $\Pr\left\{\mathcal{E}_k(\Phi)\middle| H_i\right\}$, which in turn can be lower bounded by $\Pr\left\{\delta_{\hat{j}} < \delta_i\right\}$, as in (13).

C. Pairwise Error Probabilities

The pairwise error probabilities $\Pr\left\{\delta_j < \delta_i\right\}$ are crucial for the bounds on the symbol as well as the bit error rate. They can be obtained via the calculation of residues (cf. [9, 10, 15–17]). However, the residues depend on the eigenvalues of $\mathbf{C}_{ij} = \mathbf{K}_{\mathbf{zz}|H_i}(\mathbf{Q}_j - \mathbf{Q}_i)$ and do not in general give any insight into the dependencies on the system parameters of interest, such as the signal and fading correlations. A remedy for this is offered by the asymptotic (high SNR) analysis of the pairwise error probabilities in [9], where we examined the asymptotic analysis of quadratic receivers in Rayleigh fading channels and found formulas for the asymptotic error-probabilities that require "only" the evaluation of the asymptotic eigenvalues of $\mathbf{C}_{ij}$. The structure of these asymptotic eigenvalues follows the structure observed in [9]: half of the non-zero eigenvalues are positive and linear in σ^{-2}, and the other half converge to minus unity.

We state next the pairwise error probabilities for finite SNR in the following proposition, which can be easily obtained from, for example, [9].

Proposition 1 (Expression for $\Pr\left\{\delta_j < \delta_i\right\}$) Let $\left\{\lambda_l\right\}_{l=1}^{L}$ be the distinct non-zero eigenvalues of $\mathbf{C}_{ij} = \mathbf{K}_{\mathbf{zz}|H_i}(\mathbf{Q}_j - \mathbf{Q}_i)$ with multiplicities $\left\{\mu_l\right\}_{l=1}^{L}$, and let $\left\{\lambda_l\right\}_{l=1}^{L_n}$ be negative and $\left\{\lambda_l\right\}_{l=L_n+1}^{L}$ positive, respectively. Then

$$\Pr\left\{\delta_j < \delta_i\right\} = -\sum_{k=1}^{L_n}\mathrm{Res}\left(\frac{1}{s\,\prod_{l=1}^{L}\lambda_l^{\mu_l}\left(s+\frac{1}{\lambda_l}\right)^{\mu_l}},\ s_k = \frac{-1}{\lambda_k}\right).$$

The residue of a function $\mathrm{f}(s)$ in a pole a of multiplicity m can be calculated as

$$\mathrm{Res}\Big(\mathrm{f}(s),\ a\Big) = \frac{1}{(m-1)!}\lim_{s\to a}\frac{d^{m-1}}{ds^{m-1}}\Big[(s-a)^m\mathrm{f}(s)\Big].$$

For rational functions the limit is trivial, because the poles cancel with the $(s-a)^m$ terms.

Note that the calculation of the residues is numerically unstable for high-multiplicities of eigenvalues, so that for these cases one must use, for example, a saddle point integration technique [18].

In [9] we find the asymptotic eigenvalues of $\mathbf{C}_{ij}$ to in turn obtain the asymptotic expression for $\Pr\left\{\delta_j < \delta_i\right\}$ for the single-user case. Here, we consider the general multiuser setting and moreover also find the finite-SNR eigenvalues analytically, simplifying numerical calculations ([10] presents them for the single-user case). To this end, we introduce some assumptions and notation. We assume that the users are ordered such that users $1, 2, \ldots, e_{ij}$ suffer from an error, if the receiver would erroneously decide for hypothesis H_j when hypothesis H_i is true. To avoid a complication in notation, we do not denote this user-ordering with any special symbols, but assume it implicitly. Another notational convenience is to split up the transmitted signal into two parts, the first containing the signals of the e_{ij} users that suffer from an error relative to H_i, and the second part containing the $\bar{e}_{ij} = K - e_{ij}$ signals corresponding to the correctly

[3] $\Pr\left\{\delta_j < \delta_i\right\}$ is only an *error* probability in a binary hypothesis test. However, the term "pairwise error probability" is customarily used in the literature.

detected users, i.e.,

$$\mathbf{F}_i = \left[\mathbf{F}_i^{\text{e}}\, \mathbf{F}^{\text{c}}\right],\ \mathbf{F}_j = \left[\mathbf{F}_j^{\text{e}}\, \mathbf{F}^{\text{c}}\right], \tag{15}$$

where c signifies the common part in the two signals $\mathbf{F}_i$ and $\mathbf{F}_j$. The matrices $\mathbf{F}_i^{\text{e}}$ and $\mathbf{F}_j^{\text{e}}$ are $D \times e_{ij} N_{\text{T}}$ and $\mathbf{F}^{\text{c}}$ is $D \times \bar{e}_{ij} N_{\text{T}}$. Similarly, we define $\boldsymbol{\mathcal{F}}_i^{\text{e}}$, $\boldsymbol{\mathcal{F}}_j^{\text{e}}$, and $\boldsymbol{\mathcal{F}}^{\text{c}}$ (whose sizes are multiplied by N_{R} when compared to $\mathbf{F}_i^{\text{e}}$, $\mathbf{F}_j^{\text{e}}$, and $\mathbf{F}^{\text{c}}$, respectively). Furthermore, we define $\boldsymbol{\Sigma}_{\text{ee}}$ and $\boldsymbol{\mathcal{W}}_{\text{ee}}$ as the $e_{ij}N \times e_{ij}N$ upper-left block of $\boldsymbol{\Sigma}$ and $\boldsymbol{\mathcal{W}}$, respectively (recall $N = N_{\text{T}}N_{\text{R}}$). $\boldsymbol{\Sigma}_{\text{ec}}$ and $\boldsymbol{\Sigma}_{\text{cc}}$ ($\boldsymbol{\mathcal{W}}_{\text{cc}}$) are the corresponding upper- and lower-right blocks of $\boldsymbol{\Sigma}$ ($\boldsymbol{\mathcal{W}}$). With these definitions we state the eigenvalues of $\mathbf{C}_{ij}$ in the following proposition.

Proposition 2 (Eigenvalues of $\mathbf{C}_{ij}$)

Let $\boldsymbol{\mathcal{F}}_{ji}^{\text{e}} = \boldsymbol{\mathcal{F}}_j^{\text{e}} - \boldsymbol{\mathcal{F}}_i^{\text{e}}$ and let $\left\{\rho_l\right\}_{l=1}^{\tilde{r}}$ be the non-zero eigenvalues of $\mathbf{M}_{ij} = \boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\Sigma}_{\text{ee}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}\,\dagger}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}}$. Then $\mathbf{C}_{ij} = \mathbf{K}_{\mathbf{zz}|H_i}\left(\mathbf{Q}_j - \mathbf{Q}_i\right)$ has $2\tilde{r}$ non-zero eigenvalues $\left\{\lambda_l\right\}_{l=1}^{2\tilde{r}}$ that can be calculated as

$$\lambda_{l,\, l+\tilde{r}} = \frac{\rho_l}{2\sigma^2}\left(1 \mp \sqrt{1 + \frac{4\sigma^2}{\rho_l}}\right),$$

where the negative sign corresponds to the lth and the positive sign to the $(l+\tilde{r})$th eigenvalue. Furthermore we have that

- for small σ, the non-zero eigenvalues of $\mathbf{C}_{ij} = \mathbf{K}_{\mathbf{zz}|H_i}(\mathbf{Q}_j - \mathbf{Q}_i)$ are arbitrarily close to the $\tilde{r}$ non-zero eigenvalues of $\sigma^{-2}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\Sigma}_{\text{ee}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}\,\dagger}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}}$ and minus unity with multiplicity $\tilde{r}$.
- $\tilde{r} = rN_{\text{R}}$, if the fading correlation matrix $\boldsymbol{\Sigma}$ has full rank KN (as is assumed in this paper except in the section on the downlink model), and we define r as the rank of $\mathbf{F}_j^{\text{e}} - \mathbf{F}_i^{\text{e}}$.

Proof: In contrast to our presentation in [19], we do not require the invertability of the fading covariance matrix $\boldsymbol{\Sigma}$, thereby simplifying the treatment of the downlink model considered later. Note also, that [9, Appendix A] makes use of the classic result of [20] (repeated in [21, Appendix B]), where the characteristic function of a Hermitian quadratic form in *linearly independent* complex Gaussian random variables is derived. For zero-mean and linearly dependent random variables, the characteristic function of the quadratic form can be easily shown to be the same as derived in [20, 21].

The calculation of the eigenvalues mainly involves basic, but at times tedious, algebra with block-partitioned matrices, a similarity transformation, and the application of a determinantal equality. We block-partition the matrices $\mathbf{Q}_i$, $\mathbf{Q}_j$, and $\mathbf{K}_{\mathbf{zz}|H_i}$ as given in (8) and (10) in 3×3 blocks such that the diagonal blocks have sizes $e_{ij}N \times e_{ij}N$, $\bar{e}_{ij}N \times \bar{e}_{ij}N$, and $DN_{\text{R}} \times DN_{\text{R}}$ (recall $N = N_{\text{T}}N_{\text{R}}$ and $\bar{e}_{ij} = K - e_{ij}$). Then it is easy to express $(\mathbf{Q}_j - \mathbf{Q}_i)$ in terms of this partitioning and after some tedious algebra one can obtain $\mathbf{C}_{ij}$. Making use of

$$\begin{aligned}\boldsymbol{\mathcal{F}}_i\boldsymbol{\mathcal{W}}^{1/2}\boldsymbol{\Sigma}\boldsymbol{\mathcal{W}}^{1/2}\boldsymbol{\mathcal{F}}_i^{\dagger} = {}&\\ &\boldsymbol{\mathcal{F}}_i^{\text{e}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\Sigma}_{\text{ee}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\mathcal{F}}_i^{\text{e}\dagger} + \boldsymbol{\mathcal{F}}_i^{\text{e}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\Sigma}_{\text{ec}}\boldsymbol{\mathcal{W}}_{\text{cc}}^{1/2}\boldsymbol{\mathcal{F}}_i^{\text{c}\dagger} +\\ &\boldsymbol{\mathcal{F}}_i^{\text{c}}\boldsymbol{\mathcal{W}}_{\text{cc}}^{1/2}\boldsymbol{\Sigma}_{\text{ce}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\mathcal{F}}_i^{\text{e}\dagger} + \boldsymbol{\mathcal{F}}_i^{\text{c}}\boldsymbol{\mathcal{W}}_{\text{cc}}^{1/2}\boldsymbol{\Sigma}_{\text{cc}}\boldsymbol{\mathcal{W}}_{\text{cc}}^{1/2}\boldsymbol{\mathcal{F}}_i^{\text{c}\dagger},\end{aligned}$$

to simplify the $(3,1)$–block of $\mathbf{C}_{ij}$, we can write $\mathbf{C}_{ij}$ as

$$\mathbf{C}_{ij} = \sigma^{-2}\mathbf{ABC} + \mathbf{D}, \tag{16}$$

where

$$\mathbf{A} = \begin{bmatrix} \boldsymbol{\Sigma}_{\text{ee}} \\ \boldsymbol{\Sigma}_{\text{ce}} \\ \boldsymbol{\mathcal{F}}_i^{\text{e}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\Sigma}_{\text{ee}} + \boldsymbol{\mathcal{F}}^{\text{c}}\boldsymbol{\mathcal{W}}_{\text{cc}}^{1/2}\boldsymbol{\Sigma}_{\text{ce}} \end{bmatrix}, \tag{17}$$

$$\mathbf{B} = \boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\left(\boldsymbol{\mathcal{F}}_j^{\text{e}} - \boldsymbol{\mathcal{F}}_i^{\text{e}}\right)^{\dagger}, \tag{18}$$

$$\mathbf{C} = \begin{bmatrix} \boldsymbol{\mathcal{F}}_j^{\text{e}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2} & \boldsymbol{\mathcal{F}}^{\text{c}}\boldsymbol{\mathcal{W}}_{\text{cc}}^{1/2} & -\mathbf{I}_{DN_{\text{R}}} \end{bmatrix}, \tag{19}$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{0}_{e_{ij}N} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{\bar{e}_{ij}N} & \mathbf{0} \\ -\left(\boldsymbol{\mathcal{F}}_j^{\text{e}} - \boldsymbol{\mathcal{F}}_i^{\text{e}}\right)\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2} & \mathbf{0} & \mathbf{0}_{DN_{\text{R}}} \end{bmatrix}. \tag{20}$$

Since the eigenvalues of $\mathbf{C}_{ij}$ are the same as the eigenvalues of $\widehat{\mathbf{C}}_{ij} = \mathbf{T}\mathbf{C}_{ij}\mathbf{T}^{-1}$ for any invertible matrix $\mathbf{T}$, we are free to choose

$$\mathbf{T} = \begin{bmatrix} \mathbf{I}_{e_{ij}N} & \mathbf{0} & \mathbf{0} \\ -\boldsymbol{\Sigma}_{\text{ce}} & \mathbf{I}_{\bar{e}_{ij}N} & \mathbf{0} \\ -\boldsymbol{\mathcal{F}}_i^{\text{e}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2} & -\boldsymbol{\mathcal{F}}^{\text{c}}\boldsymbol{\mathcal{W}}_{\text{cc}}^{1/2} & \mathbf{I}_{DN_{\text{R}}} \end{bmatrix}, \tag{21}$$

$$= \begin{bmatrix} \mathbf{I}_{e_{ij}N} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{\Sigma}_{\text{ce}} & \mathbf{I}_{\bar{e}_{ij}N} & \mathbf{0} \\ \left[\mathbf{T}^{-1}\right]_{31} & \boldsymbol{\mathcal{F}}^{\text{c}}\boldsymbol{\mathcal{W}}_{\text{cc}}^{1/2} & \mathbf{I}_{DN_{\text{R}}} \end{bmatrix}^{-1}, \tag{22}$$

where $\left[\mathbf{T}^{-1}\right]_{31} = \boldsymbol{\mathcal{F}}_i^{\text{e}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2} + \boldsymbol{\mathcal{F}}^{\text{c}}\boldsymbol{\mathcal{W}}_{\text{cc}}^{1/2}\boldsymbol{\Sigma}_{\text{ce}}$. Defining $\boldsymbol{\mathcal{F}}_{ji}^{\text{e}} = \boldsymbol{\mathcal{F}}_j^{\text{e}} - \boldsymbol{\mathcal{F}}_i^{\text{e}}$, we can calculate $\widehat{\mathbf{C}}_{ij}$ as given on top of the following page.

Since the second (block-) column of $\widehat{\mathbf{C}}_{ij}$ is identical to zero and we are only interested in the non-zero eigenvalues of $\mathbf{C}_{ij}$, we can expunge this column and the corresponding row and find the eigenvalues of

$$\widetilde{\mathbf{C}}_{ij} = \begin{bmatrix} \sigma^{-2}\boldsymbol{\Sigma}_{\text{ee}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}\,\dagger}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2} & -\sigma^{-2}\boldsymbol{\Sigma}_{\text{ee}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}\,\dagger} \\ -\boldsymbol{\mathcal{F}}_{ji}^{\text{e}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2} & \mathbf{0}_{DN_{\text{R}}} \end{bmatrix}.$$

The eigenvalues of $\widetilde{\mathbf{C}}$ are easily found by applying a determinantal equality [22, Section 0.8.5], i.e.,

$$\begin{aligned}\left|\widetilde{\mathbf{C}}_{ij} - \lambda\mathbf{I}\right| = \left|-\lambda\mathbf{I}\right| \Big|\sigma^{-2}\boldsymbol{\Sigma}_{\text{ee}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}\,\dagger}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2} - {}& \\ \lambda\mathbf{I}_{e_{ij}N} - \sigma^{-2}\boldsymbol{\Sigma}_{\text{ee}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}\,\dagger}\left(-\lambda\mathbf{I}\right)^{-1}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\Big|.&\end{aligned} \tag{24}$$

Let $\left\{\rho_l\right\}_{l=1}^{\tilde{r}}$ be the $\tilde{r}$ non-zero eigenvalues of $\mathbf{M}_{ij} = \boldsymbol{\Sigma}_{\text{ee}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}\,\dagger}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}$ (since the non-zero eigenvalues of any matrix product $\mathbf{XY}$ equal those of $\mathbf{YX}$, we may write $\left\{\rho_l\right\}_{l=1}^{\tilde{r}}$ in terms of several different matrix products [22, Theorem 1.3.20]). Then the $2\tilde{r}$ non-zero eigenvalues $\left\{\lambda_l\right\}_{l=1}^{2\tilde{r}}$ of $\widetilde{\mathbf{C}}_{ij}$ (and thus $\mathbf{C}_{ij}$) are found as the solutions of a standard quadratic equation and are given in the proposition. The asymptotic eigenvalues of $\mathbf{C}_{ij}$ are easily found by inserting the series expansion for $\sqrt{1 + \frac{4\sigma^2}{\rho_l}}$ for small σ.

Note that the eigenvalues $\left\{\rho_l\right\}_{l=1}^{\tilde{r}}$ are not necessarily distinct, so that the eigenvalues $\left\{\lambda_l\right\}_{l=1}^{2\tilde{r}}$ are not necessarily distinct. Furthermore, if the fading covariance matrix $\boldsymbol{\Sigma}$ has full

$$\widehat{\mathbf{C}}_{ij} = \begin{bmatrix} \sigma^{-2}\boldsymbol{\Sigma}_{\text{ee}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}\,\dagger}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2} & \mathbf{0} & -\sigma^{-2}\boldsymbol{\Sigma}_{\text{ee}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}\,\dagger} \\ \sigma^{-2}\boldsymbol{\Sigma}_{\text{ce}}\left(\mathbf{I}_{e_{ij}N} - \boldsymbol{\Sigma}_{\text{ee}}\right)\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}\,\dagger}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2} & \mathbf{0}_{\bar{e}_{ij}N} & -\sigma^{-2}\boldsymbol{\Sigma}_{\text{ce}}\left(\mathbf{I}_{e_{ij}N} - \boldsymbol{\Sigma}_{\text{ee}}\right)\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}\,\dagger} \\ -\boldsymbol{\mathcal{F}}_{ji}^{\text{e}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2} & \mathbf{0} & \mathbf{0}_{DN_{\text{R}}} \end{bmatrix}. \tag{23}$$

rank KN (thus $\boldsymbol{\Sigma}_{\text{ee}}$ has full rank $e_{ij}N$), the rank of $\mathbf{M}_{ij}$ is equal to the rank of $\boldsymbol{\mathcal{F}}_{ji}^{\text{e}} = \boldsymbol{\mathcal{F}}_{j}^{\text{e}} - \boldsymbol{\mathcal{F}}_{i}^{\text{e}}$, which in turn equals N_{R} times the rank of $\mathbf{F}_{ji}^{\text{e}} = \mathbf{F}_{j}^{\text{e}} - \mathbf{F}_{i}^{\text{e}}$, because there exists a permutation matrix $\mathbf{P}$ such that $\boldsymbol{\mathcal{F}}_{ji}^{\text{e}}\mathbf{P} = \mathbf{I}_{N_{\text{R}}} \otimes \mathbf{F}_{ji}^{\text{e}}$ and the latter matrix has rank rN_{R}, where r is the rank of $\mathbf{F}_{ji}^{\text{e}}$ [23, Theorem 4.2.15]. ■

It easily follows from the proposition that the eigenvalues of $\mathbf{C}_{ij}$ are the same as the eigenvalues of $\mathbf{C}_{ji}$ and thus $\Pr\left\{\delta_j < \delta_i\right\} = \Pr\left\{\delta_i < \delta_j\right\}$, which can be exploited to simplify numeric evaluations. Even for the single-user case the above proposition allows us to slightly generalize the results of [9], in which the asymptotic eigenvalues (and pairwise error probability) were only obtained under the assumption of full rank $\boldsymbol{\mathcal{F}}_{ji}$.

With the asymptotic eigenvalues of Proposition 2 and the results of [9], one easily finds the asymptotic pairwise error-probability given in the next proposition. For ease of notation, we introduce $|\mathbf{X}|_{\mathsf{NZ}}$ as the product of the non-zero eigenvalues of $\mathbf{X}$.

Proposition 3 (Asymptotic Pairwise Error Probability) For coherent detection the pairwise error probability $\Pr\left\{\delta_j < \delta_i\right\}$ of the optimum receiver Φ approaches arbitrarily closely

$$\Pr{}^{\text{a}}\left\{\delta_j < \delta_i\right\} = \frac{\sigma^{2rN_{\text{R}}}\binom{2rN_{\text{R}}-1}{rN_{\text{R}}}}{\left|\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\Sigma}_{\text{ee}}\boldsymbol{\mathcal{W}}_{\text{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}\,\dagger}\boldsymbol{\mathcal{F}}_{ji}^{\text{e}}\right|_{\mathsf{NZ}}}$$

as σ goes to zero.

D. Asymptotic Single-User Like Performance

The asymptotic multiuser efficiency of the kth user is a well known asymptotic measure of performance of a multiuser receiver [24]. It is the fraction of the energy required by a single-user (without interference) to attain asymptotically the same error rate as the kth user in a multiuser channel. For its calculation, one usually requires the exact asymptotic error rate of the kth user and the single user, which are often hard to obtain in fading channels for M-ary modulation with $M > 2$. Particularly for space–time single-user constellations, exact expressions for the symbol or bit error rate are often intractable and only bounds on the error rate are available. On the other hand, as we have seen above, the union bound on the kth user's symbol error rate is easily obtained, as is of course the union bound on the single-user error rate. If these converge, the multiuser receiver achieves asymptotic single-user like performance, a formal definition of which follows.

Definition 1 (Asymptotic Single-User Like Performance) Let

$$P_{\text{SU}}^{\text{UB}}\left(w, \boldsymbol{\Sigma}_{\text{SU}}\right) = M^{-1}\sum_{m=1}^{M}\sum_{\substack{n=1\\n\neq m}}^{M}\Pr\left\{\delta_n^{\text{SU}} < \delta_m^{\text{SU}}\right\}$$

be the union bound on the symbol error rate of a single user with average energy w and fading covariance $\boldsymbol{\Sigma}_{\text{SU}}$, employing the same M-ary modulation as the kth user in a K-user channel. The multiuser receiver Φ achieves asymptotic single-user like performance, if the union bound on $\Pr\left\{\mathcal{E}_k(\Phi)\right\}$ converges to $P_{\text{SU}}^{\text{UB}}$ as $\sigma \to 0$ for each k, i.e., if

$$\lim_{\sigma\to 0}\frac{M^{-K}\sum_{i=1}^{M^K}\sum_{\forall j\in\Lambda_i(k)}\Pr\left\{\delta_j < \delta_i\right\}}{P_{\text{SU}}^{\text{UB}}\left(w_k, \boldsymbol{\Sigma}_{kk}\right)} = 1\ \forall k,$$

where $\Lambda_i(k)$ is the set of the $(M-1)M^{K-1}$ indices of hypotheses in which the kth user's symbol differs from its symbol corresponding to the true hypothesis H_i.

When binary modulation is employed, asymptotic single-user like performance corresponds to each user achieving an asymptotic efficiency of 1. For $M > 2$ it amounts to saying that asymptotically the SER of the kth user in a K-user channel is bounded by the same union bound as a single-user with the same energy and fading covariance.

IV. Interpretations and Signal Designs for Uplink Communications

The asymptotic result of the previous section encompasses many special cases of interest, some of which we explore in this section. For example, we specialize to $N_{\text{T}} = 1$ in Section IV-A and gain some insights into this case, which help understand the multiple transmit antenna case. In Section IV-B we consider multiple transmit antennas but specialize to $K = 1$ first, before we discuss the general K-user, N_{T}-antenna problem. We focus on giving specific results and interpretations for the asymptotic pairwise error probabilities and assume that it is understood that the corresponding optimum receiver can be obtained by applying the specifics to Φ as defined above.

A. One Transmit-Antenna per User

We distinguish between linear versus general M-ary/block-coded modulation. In linear modulation, each user modulates its signature sequence by a symbol drawn from a fixed alphabet, like a QAM or PSK constellation. In M-ary or block-coded modulation the kth user's mth signal vector may be a block-code over a finite alphabet or a super-symbol drawn from an arbitrary constellation.

A.1 Linear Modulation – Multiuser Detection and CDMA Signature Sequence Design

If we specialize $\mathbf{F}_i = \mathbf{F}\mathbf{B}_i$, where $\mathbf{B}_i$ is a $K \times K$ diagonal matrix containing the users' constellation symbols, the system model of (1) corresponds to a synchronous code division multiple access (CDMA) model. Note that we do not make any assumptions on the number of dimensions D, so that the important case of overloaded systems with $D < K$ are included in the analysis. In the multiuser/CDMA detection literature usually only $D \geq K$ is considered [24, 25].

For the corollary, we define $\mathbf{F}^{\mathrm{e}}$ to be made up from the columns (signatures sequences) of the e_{ij} users that suffer from an error. $\mathbf{B}_i^{\mathrm{e}}, \mathbf{B}_j^{\mathrm{e}}$ are diagonal matrices that contain the information symbols of these users.

Corollary 1: (Asymptotic Pairwise Error Probability for Coherent CDMA Detection)
Assuming $\mathbf{\Sigma} = \mathbf{I}_{KN_{\mathrm{R}}}$[4] and that any subset of D columns of $\mathbf{F}$ span the D-dimensional signal space, we have for $e_{ij} \leq D$ that the pairwise error probability of the optimum detector Φ approaches arbitrarily closely

$$\Pr^{\mathrm{a}}\left\{\delta_j < \delta_i\right\} = \frac{\sigma^{2e_{ij}N_{\mathrm{R}}}\binom{2e_{ij}N_{\mathrm{R}}-1}{e_{ij}N_{\mathrm{R}}}}{\left|\mathbf{W}_{\mathrm{ee}}\left(\mathbf{B}_j^{\mathrm{e}}-\mathbf{B}_i^{\mathrm{e}}\right)^{\dagger}\mathbf{F}^{\mathrm{e}\dagger}\mathbf{F}^{\mathrm{e}}\left(\mathbf{B}_j^{\mathrm{e}}-\mathbf{B}_i^{\mathrm{e}}\right)\right|^{N_{\mathrm{R}}}}$$

as σ goes to zero.

For $D \leq e_{ij}$ we have

$$\Pr^{\mathrm{a}}\left\{\delta_j < \delta_i\right\} = \frac{\sigma^{2DN_{\mathrm{R}}}\binom{2DN_{\mathrm{R}}-1}{DN_{\mathrm{R}}}}{\left|\mathbf{F}^{\mathrm{e}}\left(\mathbf{B}_j^{\mathrm{e}}-\mathbf{B}_i^{\mathrm{e}}\right)\mathbf{W}_{\mathrm{ee}}\left(\mathbf{B}_j^{\mathrm{e}}-\mathbf{B}_i^{\mathrm{e}}\right)^{\dagger}\mathbf{F}^{\mathrm{e}\dagger}\right|^{N_{\mathrm{R}}}}.$$

Consider the case $D = 1$ and note that as the number of users increases, the aggregate spectral efficiency $K\log_2 M$ increases linearly since all the users employ the same "signature signal" which, without being wasteful of bandwidth, can be taken to be the minimum bandwidth sinc pulse or a raised cosine pulse with sufficient roll-off to ensure robustness to timing jitter and user quasi-synchronism. The second bound of Corollary 1 implies that there is no loss of diversity order compared to a single-user channel for any of the users. As an aside, we note here that for BPSK modulation this bound specializes to the one given in [8]. Without any bandwidth expansion compared to a single-user channel, multiple users can be accommodated with no loss of diversity order. There would be, however, a loss of energy efficiency in that each user would have to transmit at a somewhat higher power to achieve the performance it would have in the absence of other users, and this loss would increase with the number of users (see Figure 1).

Consider the case $D = 2$ with aggregate spectral efficiency of $\frac{K}{2}\log_2 M$ bits/dimension that also linearly increases with an increase in the number of users albeit at half the rate of the narrowband channel. In this case, it is easy to design two-dimensional signature sequences that ensure that any two users are assigned linearly independent signals so that the probability of error events involving more than one user decays with a diversity order of two. Consequently, each user not only achieves full order of diversity but even the above-mentioned loss of asymptotic effective energy relative to single-user performance is eliminated. In summary, with a bandwidth expansion by a factor of two relative to a single-user channel, an increasing number of users can be accommodated and received with a reliability that is asymptotically equivalent to "single-user like" performance for every user, in the sense that the upper bound on the multiuser BER converges to the single-user upper bound. We generalize this interpretation to any $D > 1$ in the following corollary.

Corollary 2: (Asymptotic single-user like performance for one transmit antenna) For $D > 1$ and $\mathbf{F}$ such that any subset of D columns span the D-dimensional signal space, the optimum receiver Φ achieves asymptotic single-user like performance.

Proof: For $e_{ij} = 1$ (only one user is detected erroneously) we see from Corollary 1 that the corresponding pairwise error probabilities are independent of any of the interfering $K-1$ users' quantities. In other words, the pairwise error probabilities with $e_{ij} = 1$ of the kth user in a K-user channel coincide with the pairwise error probabilities of the single-user channel. Moreover, the pairwise error probabilities with $e_{ij} > 1$ decay at least like $\sigma^{4N_{\mathrm{R}}}$ (have diversity order $2N_{\mathrm{R}}$) and consequently can be asymptotically neglected in the upper bound on the kth user's SER (11). Thus the requirements of Definition 1 are easily seen to be fulfilled. ■

Note that for BPSK signaling ($M = 2$), $D \geq K$ (and consequently the signal correlation matrix $\mathbf{F}^{\dagger}\mathbf{F}$ has full rank K), this result specializes to the well-known finding of [26] that the asymptotic efficiency of optimally detected CDMA in Rayleigh fading is unity (see also [24, pg. 206]). The much stronger result of Corollary 2 however is that that one can achieve an asymptotic efficiency of one with only two spreading dimensions, independent of the number of users. We next address the signal design question.

Signal Design To obtain spreading signals for $D > 1$ dimensions we suggest a signal design algorithm that minimizes the maximum of a per-user asymptotic performance criterion over all users. This criterion is derived from the upper bound on the kth user's asymptotic bit error rate, which in turn results from (14) by upper-bounding $\Pr\left\{H_i \rightarrow H_j\right\}$ by $\Pr^{\mathrm{a}}\left\{\delta_j < \delta_i\right\}$ for small σ, i.e.,

$$P_k^{\mathrm{b}} \leq M^{-K}\sum_{i=1}^{M^K}\sum_{\forall j\in\Lambda_i(k)}\frac{b_{ij}(k)}{\log_2 M}\Pr^{\mathrm{a}}\left\{\delta_j < \delta_i\right\} \quad (25)$$

$$= \sum_{d=1}^{D} c_d(k)\sigma^{2dN_{\mathrm{R}}}, \quad (26)$$

where $c_d(k)$ contains all the coefficients with diversity order dN_{R} in the upper bound on the kth user's BER. The signal design algorithm must minimize $\max_{1\leq k\leq K} c_2(k)$. While the terms $c_d(k)$ with $d > 1$ asymptotically do not influence the BER, we conjecture that by minimizing $c = \max_{1\leq k\leq K} c_2(k)$, the convergence of the upper-bound to the lower-bound is improved, so that the BER of a system employing optimized signals is improved at finite $\frac{w_k}{\sigma^2}$. Note that asymptotically the BER does not depend on the signals (nor the dimensionality D) provided that at least any two signature sequences are linearly independent (and all signature sequences are normalized to have energy one).

Evaluating $c_2(k)$ is cumbersome (particularly for large number of users) and depends on the user's symbol alphabet

[4] In this and some of the following corollaries, the assumption of i.i.d. fading is only made to enable a compact presentation of the results.

(through M in (25)). In other words, if the user's constellation size (M) is increased to increase spectral efficiency, the design criterion changes. This motivates the introduction of a simplified design criterion, which can be efficiently computed.

For binary modulation the criterion can be simplified by realizing that $\left(\mathbf{B}_j^{\mathrm{e}} - \mathbf{B}_i^{\mathrm{e}}\right) \mathbf{W}_{\mathrm{ee}} \left(\mathbf{B}_j^{\mathrm{e}} - \mathbf{B}_i^{\mathrm{e}}\right)^\dagger = 4\mathbf{W}_{\mathrm{ee}}$ independent of (i, j), which reduces the complexity of evaluating $c_2(k)$. For M-ary modulation, the signals designed for the binary case still fulfill the criterion that any D-signals span the D-dimensional signal space, so that these still perform well.

For $D = 2$ dimensions, we further propose to simplify the algorithm by only considering error events that affect two users (with BPSK) so that we minimize

$$\widetilde{c} = \max_{1 \le k \le K} \sum_{\substack{l=1 \\ l \ne k}}^{K} \left(w_k w_l \left(1 - |\rho_{kl}|^2 \right) \right)^{-N_{\mathrm{R}}}, \tag{27}$$

where $\rho_{kl} = \mathbf{f}_k^\dagger \mathbf{f}_l$ and $\mathbf{f}_k$ is the kth column of $\mathbf{F}$, i.e. the kth user's signature sequence. In our examples, we could not tell a difference in performance between signature sequences that were designed to minimize this simplified criterion compared to signature sequences that minimize $\max_{1 \le k \le K} c_2(k)$. This is not too surprising, because for the $e_{ij} > 2$ error events we need to add terms of the form $\left| \mathbf{F}^{\mathrm{e}} \mathbf{W}_{\mathrm{ee}} \mathbf{F}^{\mathrm{e}\dagger} \right|^{-N_{\mathrm{R}}}$ to (27), and we have

$$\begin{aligned} \left| \mathbf{F}^{\mathrm{e}} \mathbf{W}_{\mathrm{ee}} \mathbf{F}^{\mathrm{e}\dagger} \right|^{1/2} &= \left| \sum_{l=1}^{e_{ij}} w_l \mathbf{f}_l \mathbf{f}_l^\dagger \right|^{1/2} \\ &= \left| (e_{ij} - 1)^{-1} \sum_{l=1}^{e_{ij}-1} \sum_{k=l+1}^{e_{ij}} \left(w_l \mathbf{f}_l \mathbf{f}_l^\dagger + w_k \mathbf{f}_k \mathbf{f}_k^\dagger \right) \right|^{1/2} \\ &\ge \sum_{l=1}^{e_{ij}-1} \sum_{k=l+1}^{e_{ij}} \left| (e_{ij} - 1)^{-1} \left(w_l \mathbf{f}_l \mathbf{f}_l^\dagger + w_k \mathbf{f}_k \mathbf{f}_k^\dagger \right) \right|^{1/2} \end{aligned}$$

by Minkowski's inequality [22, Theorem 7.8.8]. Since $\left| w_l \mathbf{f}_l \mathbf{f}_l^\dagger + w_k \mathbf{f}_k \mathbf{f}_k^\dagger \right| = w_k w_l \left(1 - |\rho_{kl}|^2 \right)$ and minimizing (27) tends to maximize these terms, the simplified criterion gives essentially the same results as the original one. Note that any two columns of a signature sequence matrix $\mathbf{F}$ that is optimized with respect to the simplified criterion will be linearly independent, and thus single-user like performance is guaranteed asymptotically, independent of the specific modulation scheme. Consequently, the same spreading sequences can be used for various constellations and thus spectral efficiencies.

The simplified criterion also gives some insight into the question, as to what kind of signal sets are optimal: For equal-energy users, $c_2(k)$ is equalized for all users. For small number of users, this leads to the absolute values of the cross-correlations ρ_{kl} being equalized. For example, the signature sequences designed for $K = 4$ equal-energy users have identical absolute values of cross-correlations of 0.58. For unequal-energies, the numerical optimization returns signature sequences whose absolute values of cross-correlations among the high-energy users are higher than among the low-energy users. However, in our numerical evaluations, the performance of signature sequences designed for unequal energies and those designed for equal energies is indistinguishable from each other over the complete SNR range, whether employed in equal or unequal energy situations.[5] Henceforth, we consider equal-energy situations only.

For $D > 2$, minimizing $c = \max_{1 \le k \le K} c_2(k)$ corresponds to minimizing $\widetilde{c}$, cf. (27), provided that at least any three columns of $\mathbf{F}$ are linearly independent. However, minimizing (27) does not automatically ensure that any three or even D columns of $\mathbf{F}$ are linearly independent. Thus the performance of the resulting signal sets can be (slightly) improved, if one adds a cost function to $c_2(k)$ that ensures that any D columns of $\mathbf{F}$ that include $\mathbf{f}_k$ span the signal space. The obvious choice for the cost function is $\lambda \sum \left| \mathbf{F}^{\mathrm{e}} \mathbf{W}_{\mathrm{ee}} \mathbf{F}^{\mathrm{e}\dagger} \right|^{-N_{\mathrm{R}}}$, where $e_{ij} = D$ and the sum is over all possible $\mathbf{F}^{\mathrm{e}}$ that include $\mathbf{f}_k$. λ is a weighting factor, that we chose in the order of 10^{-5}. However, we will argue using some examples below, that using more than $D = 2$ dimensions is not advantageous, when the total spectral efficiency remains fixed. Thus we mainly concentrate on $D = 2$, which is already sufficient to ensure asymptotic single-user like performance, cf. Corollary 2.

For the numerical optimization involved we relied on the Matlab Optimization Toolbox, whose algorithm for this minimax problem is based on standard Sequential Quadratic Programming (SQP). The optimization was repeated several times with a different random initialization for each run to decrease the chance of obtaining only a local minimum (in practice, the solutions returned almost always identical minima). We state some of the spreading matrices used in the examples in Appendix A.

Figure 2 shows the performance of multiuser systems employing optimized signature sequences in $D = 2$ dimensions (in all plots plain line styles give simulation results and lines with a triangle upper bounds). The equal-energy users ($\mathbf{W} = \mathbf{I}$) transmit spreaded BPSK symbols from one transmit antenna to one receive antenna in i.i.d. fading ($\mathbf{\Sigma} = \mathbf{I}$). The design algorithm yielded signal sets for which the users' performances are identical for finite SNR, so that we plot the upper bound and simulated BER of one user only. We see that asymptotically single-user performance is achieved. However, for increasing number of users the asymptote is reached for increasing SNR only.

Figure 3 shows the performance of the signature sequences for $K = 10$ users of the previous plot in comparison with narrowband communications ($D = 1$) and a single user in one dimension with a spectral efficiency of 5 bits/dimension (thus the single user employs 32-QAM). As before we choose $\mathbf{W} = \mathbf{I}_K$ and $\mathbf{\Sigma} = \mathbf{I}_K$ for one transmit and receive antenna. The advantage of the CDMA system over narrowband signaling at a BER of 10^{-2} is over 15 dB when compared to the $K = 10$ narrowband system and about 4 dB when compared to the $K = 5$ narrowband system, which has the same spectral efficiency as the $K = 10$, $D = 2$ CDMA system. While for a BER 10^{-2} the gap to the single user employing 32-QAM is about 1 dB, the single user is asymptotically out-performed by roughly 6 dB. Note that the single-user 32-QAM performance is what one would get for orthogonal signaling among multiple users, leading also to a

[5] We considered up to $K = 8$ user with a energy disparity of up to 10 dB.

total spectral efficiency of 5 bits/dimension.

Figure 4 displays the performance of the same systems as Figure 3 (in terms of K, D) but the receiver uses $N_{\mathrm{R}} = 2$ receive antennas. Since for $K > 4$ the signal sets that the design procedure returns do not have equal absolute values of cross-correlations (although $c_2(k)$ is equalized among the users), the CDMA signature sequences depend on the number of receive antennas and are specifically designed for $N_{\mathrm{R}} = 2$. The performance of all systems improves dramatically by the additional antenna. However, relatively to one receive antenna, the single-user employing 32-QAM modulation cannot profit as much and is now out-performed by both, the narrowband $K = 5$ user system and the $K = 10$ user CDMA system. Thus, high spectral efficiencies seem to be more (energy-) efficiently reached by a CDMA system than by orthogonal users employing large constellations.

Figure 5 explores the effect of increasing D, the number of dimensions, for a fixed number of users. As one might expect, doubling the number of dimensions from $D = 2$ to $D = 4$ for $k = 10$ users improves the performance for low SNR considerably. However, the spectral efficiency drops from 5 bits/dimension to 2.5 bits/dimension. The same spectral efficiency is achieved by $K = 5$ users in $D = 2$ dimensions. The number of users is of course not necessarily a design parameter. However, the desired spectral efficiency usually is, and for say 10 users using BPSK a spectral efficiency of 2.5 bits/dimension can be achieved either by all 10 users signaling in a common 4-dimensional signal space or two sets of five users signaling in two orthogonal 2-dimensional spaces. From the figure, we see that the latter choice gives a performance that is indistinguishable from the 10 user system in 4 dimensions. Moreover, since the complexity of joint detection for 5 users is by a factor of $2^5 = 32$ lower than detecting $K = 10$ users, there seems to be no point in increasing the number of dimensions beyond $D = 2$, the minimum to achieve single-user performance. Note again that increasing D does not benefit SER performance asymptotically, because $D = 2$ is sufficient to ensure single-user like performance, cf. Corollary 2.

Figure 6 gives an example of asymptotic "single-user like" performance. For the plot, each of the three users employs the most energy efficient 8-QAM constellation. For comparison, the BER of a single user also using 8-QAM is given in terms of bounds and simulations. Note that the upper bound on the $D = 2$, $K = 3$ multiuser system converges to the upper bound on the single-user's BER, as predicted by the analysis above. Moreover, the simulated BERs are also close. On the other hand, at a BER of 10^{-2}, the gap to narrowband signaling is roughly 5 dB.

A.2 M-ary or Block Coded Modulation

In this section, we interpret the kth $D \times 1$ column $\mathbf{s}_{ki_k}$ of $\mathbf{F}_i$ as a super-symbol of user k which can be thought of as belonging to some dense lattice (or more generally to an arbitrary non-lattice constellation), whose individual scalar elements may be drawn from a regular QAM-like alphabet or may be arbitrary complex numbers, not necessarily restricted to be part of a finite alphabet. When the "codeword" interpretation is appropriate, the receiver may be thought of as a decoder.

We rewrite the general expression of the asymptotic pairwise error probability from Proposition 3.

Corollary 3: (Asymptotic Pairwise Error Probability for Coherent Decoding)
Assuming $\mathbf{\Sigma} = \mathbf{I}_{KN_{\mathrm{R}}}$ and that $\left(\mathbf{F}_j^{\mathrm{e}} - \mathbf{F}_i^{\mathrm{e}}\right)$ has rank $r \leq \min(D, e)$, the pairwise error probability of the optimum decoder Φ approaches arbitrarily closely

$$\Pr{}^{\mathrm{a}}\left\{\delta_j < \delta_i\right\} = \frac{\sigma^{2rN_{\mathrm{R}}}\binom{2rN_{\mathrm{R}}-1}{rN_{\mathrm{R}}}}{\left|\mathbf{W}_{\mathrm{ee}}\left(\mathbf{F}_j^{\mathrm{e}} - \mathbf{F}_i^{\mathrm{e}}\right)^{\dagger}\left(\mathbf{F}_j^{\mathrm{e}} - \mathbf{F}_i^{\mathrm{e}}\right)\right|_{\mathsf{NZ}}^{N_{\mathrm{R}}}}$$

as σ goes to zero.

Let us reconcile this result for the fictitious case where all K users cooperate so that we have an equivalent single-user, K-transmit, N_{R}-receive antenna channel. In this case, the maximum diversity order for a given K could be achieved if $r = K = e$ and the proposition corresponds to the well-known rank criterion [4, 6], for which of course we need $D \geq K$. By the use of space–time codes such as the orthogonal designs of [5] or the algebraic codes of [27] which satisfy this rank criterion, one can achieve full diversity order (namely KN_{R}).

However, the multiuser rank criterion is very different from the single-user criterion because while in the single-user channel with K transmit antennas, signals transmitted over the different transmit antennas can be dependent (i.e., a super-information symbol is encoded into a $D \times N_{\mathrm{T}}$ matrix), the columns of this matrix in the multiuser channel arise from the independent transmission of vectors of length D each from the K different users. This makes the problems of modulation and coding in some sense more challenging for multiuser channels than they are for single-user space–time communications, because at this point it is not even clear that multiuser codes exist that guarantee a diversity order of $N_{\mathrm{R}}e_{ij}$, when e_{ij} users' are detected erroneously (for this the difference of the code matrices $\left(\mathbf{F}_j^{\mathrm{e}} - \mathbf{F}_i^{\mathrm{e}}\right)$ must have full rank).

B. *Multiple Transmit-Antennas per User*

The classical single-user multiple transmit antenna space–time coding analysis also profits from our general analysis: in contrast to the earlier, Chernoff bound based approaches, our analysis provides asymptotically tight expressions for the pairwise error probability and considers possibly correlated fading (as does [9, 10]). Finally, for the multi-transmit antenna, multiuser space time channel, we propose a signal design algorithm that ensures single-user like performance asymptotically.

B.1 One User

In this case one user transmits a $D \times N_{\mathrm{T}}$ signal matrix $\left\{\mathbf{S}_m\right\}_{m=1}^{M}$ with average energy w. $\mathbf{\Sigma}$ simplifies to the $N \times N$ fading correlation matrix associated with all the antennas. Recalling that we defined $\boldsymbol{\mathcal{S}}_m = \mathbf{I}_{N_{\mathrm{R}}} \otimes \mathbf{S}_m$, Proposition 3 easily simplifies to the following corollary.

Corollary 4: (Asymptotic Pairwise Error Probability for Single User Reception)

Assuming that $(\mathbf{S}_j - \mathbf{S}_i)$ has rank $r \leq \min(D, N_{\mathrm{T}})$, the pairwise error probability of the optimum receiver Φ approaches arbitrarily closely

$$\Pr{}^{\mathrm{a}}\left\{\delta_j < \delta_i\right\} = \frac{\left(\frac{w}{\sigma^2}\right)^{-rN_{\mathrm{R}}} \binom{2rN_{\mathrm{R}}-1}{rN_{\mathrm{R}}}}{\left|\boldsymbol{\Sigma}\left(\boldsymbol{\mathcal{S}}_j - \boldsymbol{\mathcal{S}}_i\right)^\dagger \left(\boldsymbol{\mathcal{S}}_j - \boldsymbol{\mathcal{S}}_i\right)\right|_{\mathsf{NZ}}}$$

as σ goes to zero.

Note that in addition to revealing the rank and determinant criterion of [4, 6] for i.i.d. fading, this formula is is also asymptotically tight and considers the more general case of correlated fading (cf. [10]). As a consequence of the asymptotic tightness, asymptotic lower bounds on symbol and bit error rates can be obtained. For correlated fading and full-diversity space–time codes, the fading correlation does not affect the determinant criterion, because $\left|\boldsymbol{\Sigma}\left(\boldsymbol{\mathcal{S}}_j - \boldsymbol{\mathcal{S}}_i\right)^\dagger \left(\boldsymbol{\mathcal{S}}_j - \boldsymbol{\mathcal{S}}_i\right)\right| = |\boldsymbol{\Sigma}| \left|\left(\mathbf{S}_j - \mathbf{S}_i\right)^\dagger \left(\mathbf{S}_j - \mathbf{S}_i\right)\right|^{N_{\mathrm{R}}}$. Consequently, full-diversity space–time codes that were optimized for i.i.d. are also asymptotically optimal for correlated fading. The analysis presented here also strengthens our result in [9] by providing exact expressions for the asymptotic pairwise error probabilities in case $\mathbf{S}_i - \mathbf{S}_j$ is low rank.

B.2 Multiple Users

The observations we made for the various special cases allow us to finally draw some conclusions about Proposition 3 for multiuser communication when each user employs N_{T} transmit antennas. Interestingly, if every user employs a full-diversity space–time code/constellation (requiring $D \geq N_{\mathrm{T}}$) in the same D dimensional signal space, the optimum receiver still achieves asymptotically a diversity order of $N = N_{\mathrm{R}} N_{\mathrm{T}}$,[6] i.e., no loss in diversity order occurs when compared to the single-user case, without any bandwidth expansion (this was also realized independently in [27] by using the weaker Chernoff analysis that does not yield asymptotically tight bounds on pairwise error rates). However, a loss in energy-efficiency occurs when more users are added. This behavior mirrors exactly the $N_{\mathrm{T}} = 1$, $D = 1$ narrowband case discussed above. We saw that we could improve on this behavior by expanding the signal space to $D = 2$ dimensions to design signals such that the optimum receiver achieves single-user like performance asymptotically. We propose to generalize this idea to the multiple transmit antenna case by signaling according to

$$\mathbf{F}_i = \mathbf{F} \begin{bmatrix} \mathbf{B}_{1i_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{2i_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{B}_{Ki_K} \end{bmatrix}, \qquad (28)$$

where $\mathbf{F}$ is a $D \times KD_{\mathrm{SU}}$ fixed "signature" matrix and $\mathbf{B}_{ki_k}$ are $D_{\mathrm{SU}} \times N_{\mathrm{T}}$ single-user full-diversity symbol matrices. The latter can originate from, for example, the Alamouti scheme [11], orthogonal designs [5], or the algebraic codes of [28, 29], the universal space–time codes [30], or any other single-user space–time constellation that guarantees full transmit antenna diversity. If the matrix $\mathbf{F}$ is suitably chosen, every user can achieve asymptotically single-user like performance. Since the necessary increase in signal space dimensions D will turn out to be independent of the number of users (as for $N_{\mathrm{T}} = 1$ above), the aggregate spectral efficiency can be increased by adding users, with no loss in energy-efficiency for sufficiently high SNR (increasing with the number of users). For example, if each user employs N_{T} transmit antennas and a single-user space–time code with the minimum required dimensions N_{T} (cf. [4, 6, 9]), our proposed modulation scheme requires just one more dimension to achieve asymptotic single-user like performance. Denoting the D_{SU} columns of $\mathbf{F}$ that correspond to user k as $\mathbf{F}_k$ (so that $\mathbf{F} = [\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_K]$), we state our claim precisely in the following Corollary.

Corollary 5: (Asymptotic single-user like performance for $N_{\mathrm{T}} \geq 1$) If every user employs a full-diversity constellation $\left\{\mathbf{B}_{km}\right\}_{m=1}^{M}$ and $\left\{\mathbf{F}_k\right\}_{k=1}^{K}$ is such that

- $\mathbf{F}_k^\dagger \mathbf{F}_k = \mathbf{I}_{D_{\mathrm{SU}}}\ \forall k$,
- any compound matrix $[\mathbf{F}_m, \mathbf{F}_n]$, $m \neq n$ has rank greater than or equal to $2D_{\mathrm{SU}} - N_{\mathrm{T}} + 1$,

then the optimum receiver Φ achieves asymptotic single-user like performance.

Proof: To simplify notation, we introduce $\boldsymbol{\mathcal{F}}_k = \mathbf{I}_{N_{\mathrm{R}}} \otimes \mathbf{F}_k$ and $\boldsymbol{\mathcal{B}}_{km} = \mathbf{I}_{N_{\mathrm{R}}} \otimes \mathbf{B}_{km}$.

Consider the asymptotic pairwise error probabilities with $e_{ij} = 1$ first. Similarly to the proof of Corollary 2, we see that these probabilities are independent of interfering users (just insert $\mathbf{F}^{\mathrm{e}}_{ji} = \boldsymbol{\mathcal{F}}_1 \left(\boldsymbol{\mathcal{B}}_{1j_1} - \boldsymbol{\mathcal{B}}_{1i_1}\right)$ etc. into Proposition 3. Furthermore, they coincide with the pairwise error probabilities of a single user only transmitting $\mathbf{B}_{1i_1}$ (without spreading) because for $e_{ij} = 1$ we have

$$\begin{aligned} \boldsymbol{\mathcal{F}}^{\mathrm{e}\,\dagger}_{ji} \boldsymbol{\mathcal{F}}^{\mathrm{e}}_{ji} &= \left(\mathbf{I}_{N_{\mathrm{R}}} \otimes \left(\mathbf{B}_{1j_1} - \mathbf{B}_{1i_1}\right)^\dagger \mathbf{F}_k^\dagger\right) \\ &\quad \left(\mathbf{I}_{N_{\mathrm{R}}} \otimes \mathbf{F}_k \left(\mathbf{B}_{1j_1} - \mathbf{B}_{1i_1}\right)\right) \\ &= \mathbf{I}_{N_{\mathrm{R}}} \otimes \left(\left(\mathbf{B}_{1j_1} - \mathbf{B}_{1i_1}\right)^\dagger \mathbf{F}_1^\dagger \mathbf{F}_1 \left(\mathbf{B}_{1j_1} - \mathbf{B}_{1i_1}\right)\right), \end{aligned}$$

where we applied some basic properties of the Kronecker product. Obviously, the last expression coincides with the usual single-user channel expression for $\mathbf{F}_k^\dagger \mathbf{F}_k = \mathbf{I}$.

Now consider the pairwise error probabilities that correspond to $e_{ij} \geq 2$. By a rank inequality ([22, Section 0.4.5]), one can easily show that for $e_{ij} = 2$ the matrix

$$\boldsymbol{\mathcal{F}}^{\mathrm{e}}_{ji} = [\boldsymbol{\mathcal{F}}_1, \boldsymbol{\mathcal{F}}_2] \begin{bmatrix} \boldsymbol{\mathcal{B}}_{1j_1} - \boldsymbol{\mathcal{B}}_{1i_1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\mathcal{B}}_{2j_2} - \boldsymbol{\mathcal{B}}_{2i_2} \end{bmatrix} \qquad (29)$$

has got rank greater than or equal to $(N_{\mathrm{T}} + 1)N_{\mathrm{R}}$. Thus the diversity order of all pairwise error probabilities corresponding to error events with $e_{ij} \geq 2$ is at least $(N_{\mathrm{T}} + 1)N_{\mathrm{R}}$ and consequently these probabilities can be asymptotically neglected in the upper bound (11) when compared to the single-user pairwise error probabilities whose diversity order is $N_{\mathrm{T}} N_{\mathrm{R}}$. Consequently

[6] Recall that $\boldsymbol{\mathcal{F}}^{\mathrm{e}}_{ji} = \left[\boldsymbol{\mathcal{S}}_{1j_1} - \boldsymbol{\mathcal{S}}_{1i_1}, \ldots, \boldsymbol{\mathcal{S}}_{e_{ij}j_{e_{ij}}} - \boldsymbol{\mathcal{S}}_{e_{ij}i_{e_{ij}}}\right]$ and for full-diversity space time codes $\boldsymbol{\mathcal{S}}_{kj_k} - \boldsymbol{\mathcal{S}}_{ki_k}$ has full rank N for $k \leq e_{ij}$ so that $\boldsymbol{\mathcal{F}}^{\mathrm{e}}_{ji}$ is guaranteed to have at least rank N.

the requirements of Definition 1 are easily seen to be fulfilled. ■

The key-point of the Corollary is that $D = 2D_{\text{SU}} - N_{\text{T}} + 1$ dimensions are sufficient to achieve asymptotic single-user like performance, which is again independent of the number of users. For example, when the single-user constellation requires only the minimum dimensions $D_{\text{SU}} = N_{\text{T}}$, only one more dimension is required for the multiuser channel. In our signal design we concentrate on $D = 2D_{\text{SU}} - N_{\text{T}} + 1$, because by increasing the number of dimensions the performance is–at least asymptotically–*not* enhanced.

The design algorithm discussed in Section IV-A.1 can be adapted to design an optimized signature matrix $\mathbf{F}$. As before (25), we expand the kth user's BER into terms corresponding to the diversity order. Since the single-user space–time symbols $\mathbf{B}_{ki_k}$ are assumed to be drawn from a full transmit diversity achieving constellation, the minimum diversity order of any error event for any user is $N_{\text{T}}N_{\text{R}}$. We conjecture that by minimizing the maximum of the coefficients $c_{N_{\text{T}}+1}(k)$ $(1 \le k \le K)$, corresponding to diversity order $(N_{\text{T}}+1)N_{\text{R}}$, the convergence to the upper bound on the single-user's error rate is improved.[7] However, the computational complexity of this criterion is quite high, since on the order of M^K terms have to be evaluated to obtain the $c_{N_{\text{T}}+1}(k)$. Therefore, we suggest to generate the spreading matrices $\left\{\mathbf{F}_k\right\}_{k=1}^{K}$ by minimizing the simplified criterion

$$c = \max_{1\le k\le K} \sum_{\substack{l=1\\ l\neq k}}^{K} \left|\mathbf{F}_k\mathbf{F}_k^\dagger + \mathbf{F}_l\mathbf{F}_l^\dagger\right|^{-N_{\text{R}}}, \tag{30}$$

under the constraint that each matrix $\mathbf{F}_k$ has orthonormal columns. Note that the criterion (30) is easily evaluated, because it is independent of the specific single-user space–time constellation. It also ensures in a "fair" way that all pairs $[\mathbf{F}_l, \mathbf{F}_k]$, $l \neq k$, fulfill the rank criterion, which is necessary (and sufficient) to achieve single-user like performance.

Since we want to constrain our signals to have orthonormal columns, a numerical optimization in the usual Euclidian space would have to incorporate these constraints. The optimization can be performed unconstrained in the Grassmann manifold $G(D_{\text{SU}}, D)$, the space of all D_{SU}-dimensional subspaces of $\mathbb{C}^D$. A parameterization of $G(D_{\text{SU}}, D)$ which employs $2D_{\text{SU}}D - D_{\text{SU}}^2$ real parameters is explicitly detailed in [31] following, for example, [32]. However, with D^2 real parameters we can over-parameterize $G(D_{\text{SU}}, D)$ and obtain a somewhat simpler parameterization (cf. [31,32]).[8] The latter parameterization is built by expressing any rectangular matrix $\mathbf{U} \in \mathbb{C}^{D\times D_{\text{SU}}}$ as the product of a square $D \times D$ unitary matrix and a fixed rectangular $D \times D_{\text{SU}}$ matrix $\mathbf{W}$ (typically $\mathbf{W}$ is taken to be the first D_{SU} columns of $\mathbf{I}_D$). Any of the $D \times D$ unitary matrices can in turn be expressed as the product of a real diagonal matrix $\mathbf{\Phi}$ and "simple" unitary matrices $\mathbf{V}^{pq}(\phi_{pq}, \theta_{pq})$ (complex Givens rotation matrices, each parameterized by two angular parameters and two indices). In formulas, we have

$$\mathbf{U} = \mathbf{\Phi}\left(\prod_{p=1}^{D-1}\prod_{q=D}^{p+1}\mathbf{V}^{pq}(\phi_{pq}, \theta_{pq})\right)\mathbf{W}, \tag{31}$$

where we define the matrix products to be left-sided and the second product decreases its index by minus unity for each multiplication. The k, l element of $\mathbf{V}^{pq}(\phi_{pq}, \theta_{pq})$ is defined as

$$[\mathbf{V}^{pq}]_{k,l}(\phi_{pq}, \theta_{pq}) = \begin{cases} 1 & k = l,\ k \neq p, q \\ \cos(\phi_{pg}) & k = l, k = p, q \\ -\sin(\phi_{pq})\exp(-j\theta_{pq}) & k = p, l = q \\ \sin(\phi_{pq})\exp(j\theta_{pq}) & k = q, l = p \\ 0 & \text{else} \end{cases}$$

Each user's spreading matrix $\mathbf{F}_k$ is parameterized in this way and thus the following numerical optimization can be performed without constraints. As before we used standard methods from the Matlab Optimization Toolbox to perform the optimization. We give spreading matrices for $K = 6$ users with $N_{\text{T}} = 2$ transmit antennas each in Appendix A.

For our examples, we choose the constituent space–time symbols $\mathbf{B}_{i_k}$ according to Alamouti's scheme [11] for $N_{\text{T}} = 2$ transmit antennas. We designed spreading matrices for up to three equal-energy users in the minimum number of dimensions ($D = 3$) that guarantee single-user like performance asymptotically. From Figure 7 we see that the simulated BER for a single-user employing 4-PSK symbols in the Alamouti scheme is close to its upper bound, to which the upper bounds on the multiuser systems ($K = 2$, $K = 3$) converge. The $K = 3$ user system in which each user employs 4-PSK symbols has a total spectral efficiency of 4 bits/dimension (three users each transmitting 2 bits from each of 2 antennas in 3 dimensions). At a BER of 10^{-3}, the optimum detector can detect the three users within 1 dB of the single-user's BER (while achieving twice the spectral efficiency). If one would try to achieve a spectral efficiency of 4 bits/dimension with orthogonal users, each user would have to employ 16-QAM modulation. From the figure we see that at a BER of 10^{-3} there is roughly a 3 dB gap between the $K = 3$ spread-matrix design and the orthogonal system. Obviously, this comparison does not take into account any complexity or implementational issues (the optimum multiuser detector does not simplify for the Alamouti scheme, in contrast to the single-user detector [11]). On the other hand, for larger systems with more users and/or antennas, the gap between spread-matrix and orthogonal communications is expected to grow.

Figure 8 confirms this growing gap for ($K = 6$, $N_{\text{T}} = 2$, $D = 3$) and ($K = 8$, $N_{\text{T}} = 1$, $D = 2$). Using QPSK (in conjunction with the Alamouti scheme for $N_{\text{T}} = 2$) a total spectral efficiency of 8 bits/dimension is obtained with low-dimensional spreading. To obtain the same spectral efficiency with orthogonal users, each user would have to employ 256-QAM. Even for moderate E_b/N_0, when the BER of the spread systems have not yet converged to the corresponding single-user BER (of a single-user employing QPSK), the performance gain of low-dimensional spreading is substantial.

[7] Note that if the matrices $[\mathbf{F}_m, \mathbf{F}_n]$, $m \neq n$, fulfill the rank criterion and each $\mathbf{F}_k$ has orthonormal columns the terms $c_{N_{\text{T}}}(k)$ do not depend asymptotically on the specific spreading matrices.

[8] A parameterization in the usual Euclidian space would require $2DD_{\text{SU}}$ real parameters. For the "minimal" choices $D_{\text{SU}} = N_{\text{T}}$ and $D = D_{\text{SU}} + 1$ more than for the over-parameterization of $G(D_{\text{SU}}, D)$.

C. Sphere Decoding

The optimum multiuser receiver as given in (9) has a complexity of M^K, i.e., it is exponential in the number of users. If the information symbols are drawn from a lattice (like in PAM or QAM), the sphere-decoder [33] can be used to find the maximum-likelihood solution efficiently (cf. the recent semi-tutorial [12]). This holds if there are at least as many observations as unknown information symbols. If there are less observations than information symbols, the generalized sphere decoder [13] requires a complexity that is exponential in the difference of information symbols and observations.

We next show how the channel model (cf. (3)) can be adapted so that the sphere decoder is applicable when low-dimensional spreading and a suitable space-time code are used. To this end the model must be written such that $\mathbf{y} = \mathbf{M}\mathbf{b}_i + \boldsymbol{\eta}$, where $\mathbf{M}$ is a spreading/channel/code matrix, and $\mathbf{b}_i$ a vector of independent PAM/QAM information symbols. It is easily seen that this is possible whenever

$$\mathbf{B}_m \mathbf{h}_{kn} = \widetilde{\mathbf{H}}_{kn} \mathbf{u}_{i_k}, \tag{32}$$

where $\widetilde{\mathbf{H}}_{kn}$ is an equivalent channel/code matrix of the kth user and $\mathbf{u}_{i_k}$ is a vector of independent PAM/QAM symbols of the kth user. Obviously, for $N_{\mathrm{T}} = 1$ and low-dimensional spreading the condition is trivially met, and the sufficient statistics can be written as

$$\mathbf{y} = \boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{W}}^{1/2}\mathbf{H}\mathbf{b}_i + \boldsymbol{\eta}, \tag{33}$$

where $\mathbf{H}$ is a $KN_{\mathrm{R}} \times K$ block-diagonal matrix with $\left\{\mathbf{h}_k\right\}_{k=1}^{K}$ as diagonal elements.

For example, for $N_{\mathrm{T}} = 2$ and the Alamouti scheme the equivalent channel/code matrix to the kth user's nth receive antenna is

$$\widetilde{\mathbf{H}}_{kn} = \begin{bmatrix} h_{1kn} & jh_{1kn} & h_{2kn} & jh_{2kn} \\ h_{2kn} & -jh_{2kn} & -h_{1kn} & jh_{1kn} \end{bmatrix}. \tag{34}$$

The corresponding length-4 vector $\mathbf{u}_{i_k}$ contains the real and imaginary parts of the two complex QAM data symbols transmitted by user k. The model is unchanged to (33), if $\mathbf{H}$ is defined to be a $2KN_{\mathrm{R}} \times 4K$ block-diagonal matrix with the matrices $\left\{\mathbf{H}_k\right\}_{k=1}^{K}$ as diagonal elements, where $\mathbf{H}_k^{\mathsf{T}} = \left[\widetilde{\mathbf{H}}_{k1}^{\mathsf{T}}, \ldots, \widetilde{\mathbf{H}}_{kN_{\mathrm{R}}}^{\mathsf{T}}\right]$ and $\mathbf{b}_i^{\mathsf{T}} = \left[\mathbf{u}_{i_1}^{\mathsf{T}}, \ldots, \mathbf{u}_{i_K}^{\mathsf{T}}\right]$.

V. Adaptions for Downlink Communications

The model for downlink communications can be written exactly as in (3), with a change in structure of $\mathbf{h}$, the vector of fading coefficients. The observations $\mathbf{y}$ now depend on the receiving user k, as do the fading coefficients $\mathbf{h}$: The channel from the base-station to user k is the same for the signals of all users, so that

$$\mathbf{h} = \mathbf{h}(k) = \mathbf{1}_{K\times 1} \otimes \mathbf{h}_k \tag{35}$$

where $\mathbf{1}_{x\times y}$ is a $x \times y$ matrix of all ones and $\mathbf{h}_k$ is the $N = N_{\mathrm{T}}N_{\mathrm{R}}$-length vector whose $(t + (n-1)N_{\mathrm{T}})$th element is the fading coefficient from the base-stations tth transmit antenna to the kth user's nth receive antenna (as correspondingly defined before for the uplink). The fading correlation matrix is easily seen to be

$$\begin{aligned} \boldsymbol{\Sigma}(k) &= E\left[\mathbf{h}(k)\mathbf{h}(k)^{\dagger}\right] \\ &= \mathbf{1}_{K\times K} \otimes \boldsymbol{\Sigma}_{kk} \\ &= \left(\mathbf{1}_{K\times 1} \otimes \mathbf{I}_N\right)\boldsymbol{\Sigma}_{kk}\left(\mathbf{1}_{1\times K} \otimes \mathbf{I}_N\right), \end{aligned}$$

where we recall that $\boldsymbol{\Sigma}_{kk} = E\left[\mathbf{h}_k\mathbf{h}_k^{\dagger}\right]$ is the correlation matrix of the fading processes between the base-station's transmit antennas and the kth user's receive antennas. By a property of the Kronecker product ([23, Theorem 4.2.15]) $\boldsymbol{\Sigma}(k)$ has maximum rank N, if $\boldsymbol{\Sigma}_{kk}$ has full rank N, which we assume in the following. As a consequence of the rank deficiency of $\boldsymbol{\Sigma}(k)$, we can conclude immediately that for all error events the achievable diversity order is less than or equal to N. Thus, single-user like performance is not attainable in the downlink. We summarize the pairwise error probability for the downlink model in the following corollary, whose proof can be found in Appendix B.

Corollary 6 (Asymptotic Pairwise Error Probability for Downlink) For coherent detection in downlink communications at user's k site, the pairwise error probability $\Pr\left\{\delta_j < \delta_i\right\}$ of the optimum receiver Φ approaches arbitrarily closely

$$\Pr{}^{\mathrm{a}}\left\{\delta_j < \delta_i\right\} = \frac{\sigma^{2N}\binom{2N-1}{N}}{\left|\boldsymbol{\Sigma}_{kk}\right| \left|\sum_{l=1}^{K}\sum_{q=1}^{K}\sqrt{w_l w_q}\left(\mathbf{S}_{lj_l} - \mathbf{S}_{li_l}\right)^{\dagger}\left(\mathbf{S}_{qj_q} - \mathbf{S}_{qi_q}\right)\right|^{N_{\mathrm{R}}}}$$

as σ goes to zero and we assume that the sum of matrices in the denominator has full rank N_{T}. When this assumption does not hold and the rank of the sum is $1 \le r < N_{\mathrm{T}}$, then the pairwise error probability asymptotically declines like $\sigma^{2rN_{\mathrm{R}}}$. In the extreme case when the sum is zero (i.e. $r = 0$) the error probability floors.

In the uplink we found that even for narrowband communications ($D = N_{\mathrm{T}}$) every user can achieve full order of diversity N. From the above corollary we conclude that this result does in general not hold for the downlink. Consider $N_{\mathrm{T}} = 1$, binary modulation, and equal-energy users. Then it is easy to see that for many pairs of hypotheses (i, j) the sum in the denominator is indeed zero, leading to an error floor. For example for $K = 2$ users, if the base station transmits $+1$ for user one and -1 for user two, neither of the users can distinguish this hypothesis from the one corresponding to swapping the users' bits. This analytic result fits the intuition, that when all users experience the same random spreading (as opposed to different random spreading in the uplink narrowband channel) the receiver is not able to distinguish them.

For $N_{\mathrm{T}} = 1$, $D = 2$, and equal-energy users, we design spreading sequences by minimizing the maximum asymptotic upper bound on the users' BER, i.e., minimize $\max_{1\le k\le K} c_1(k)$ as given in (25).[9] The assumption of equal-energy user's, as received by user k, corresponds to the case that

[9]Note that for desirable downlink spreading sequences $c_1(k)$ (more generally $c_{N_{\mathrm{T}}}(k)$) is the only non-zero term in the expansion (25) and in contrast to the uplink *does* depend on the specific signals employed.

the base-station transmits the same power for all users. Clearly, this is not desired in a practical system, where the base-station typically adjusts the transmit powers according to the individual path losses and possible quality-of-service constraints. Designing spreading-sequences for the latter case is not an easy problem and beyond the scope of this paper. However, we try to gain some insight into the idealized case of equal-energy users, as to whether or not gains in energy-efficiency can be achieved when a spreaded system is compared to a system of orthogonal users with the same spectral-efficiency. Figure 9 shows simulated BER's for $K = 1$ user employing 16-QAM and QPSK modulation, and $K = 4$ equal-energy users in $D = 2$ dimensions, each employing QPSK modulation. From the figure, we see that the multiuser system has a roughly 3 dB *loss* in energy-efficiency when compared to the single-user performance with the same spectral efficiency. The gap to single-user performance QPSK performance is about 6 dB. Thus, in this example, at a given SNR per bit, orthogonal users achieve a higher spectral efficiency.

VI. Conclusions

The general analysis of [9] is applied to coherent multiuser space–time reception with a focus on the uplink channel. Consequently, asymptotically tight expressions for the pairwise error probabilities are obtained that are subsequently used as signal design criteria. We are able to arrive at several conclusions:

- If the common signal space of all users has at least dimensionality two, all users can be detected with asymptotic single-user like performance in a one transmit antenna per user CDMA system. An algorithm to design "optimum" spreading sequences is presented.
- For the "classical" single-user N_{T} transmit and N_{R} receive antenna space–time communications, we extend our previous approach by providing asymptotically tight bounds when the space–time symbols (or code matrices) do not fulfill the rank criterion or the fading correlation matrix has not full rank (cf. [10]).
- For the multiuser space–time problem it is established that every user achieves full order of diversity $N = N_{\mathrm{T}}N_{\mathrm{R}}$ when communicating with $D \geq N_{\mathrm{T}}$-dimensional single-user space–time codes in a common D-dimensional signal space. To achieve asymptotically single-user like performance for multiple users, at least $N_{\mathrm{T}} + 1$ dimensions are necessary, as opposed to N_{T} dimensions in the single-user channel.
- A modulation scheme is proposed that leverages single-user space–time constellations and guarantees both, a full diversity order, as well as an asymptotic (high SNR) single-user like performance for every user. The new scheme requires low-dimensional spreading matrices, which are obtained by numerical optimization of a simplified design criterion. A special feature of the simplified criterion, and thus of the resulting spreading matrices, is that they are independent of the particular single-user space–time constellations (of a given dimension), so that different spectral efficiencies can be attained without changing or redesigning the spreading matrices.

Acknowledgement

The authors would like to thank Pranav Dayal for programming the generalized sphere decoder.

Appendix A
Spreading Sequences and Matrices

- $K = 6$, $D = 2$, $N_{\mathrm{T}} = 1$, $N_{\mathrm{R}} = 1$

$$\mathbf{F}^{\mathsf{T}} = \begin{bmatrix} 0.87e^{j86^\circ} & 0.49e^{j179^\circ} \\ 0.86e^{j90^\circ} & 0.51e^{j74^\circ} \\ 0.93e^{-j107^\circ} & 0.38e^{j114^\circ} \\ 0.51e^{j31^\circ} & 0.86e^{-j165^\circ} \\ 0.49e^{j163^\circ} & 0.87e^{j76^\circ} \\ 0.38e^{-j113^\circ} & 0.93e^{-j72^\circ} \end{bmatrix} \tag{A.1}$$

- $K = 10$, $D = 2$, $N_{\mathrm{T}} = 1$, $N_{\mathrm{R}} = 2$

$$\mathbf{F}^{\mathsf{T}} = \begin{bmatrix} 0.11e^{j38^\circ} & 0.99e^{j164^\circ} \\ 0.87e^{j165^\circ} & 0.50e^{j137^\circ} \\ 0.30e^{-j145^\circ} & 0.96e^{j124^\circ} \\ 0.94e^{j112^\circ} & 0.33e^{-j129^\circ} \\ 0.90e^{-j60^\circ} & 0.43e^{j30^\circ} \\ 0.59e^{-j44^\circ} & 0.81e^{-j162^\circ} \\ 0.14e^{j146^\circ} & 0.99e^{j99^\circ} \\ 0.96e^{-j82^\circ} & 0.29e^{-j55^\circ} \\ 0.80e^{-j138^\circ} & 0.60e^{j93^\circ} \\ 0.72e^{-j20^\circ} & 0.69e^{j69^\circ} \end{bmatrix} \tag{A.2}$$

- $K = 6$, $D = 3$, $N_{\mathrm{T}} = 2$, $N_{\mathrm{R}} = 2$

$$\mathbf{F}^{\mathsf{T}} = \begin{bmatrix} 0.49e^{-j104^\circ} & 0.37e^{-j41^\circ} & 0.79e^{j90^\circ} \\ 0.55e^{-j106^\circ} & 0.78e^{-j170^\circ} & 0.32e^{-j160^\circ} \\ 0.20e^{j132^\circ} & 0.33e^{-j28^\circ} & 0.92e^{j119^\circ} \\ 0.87e^{-j59^\circ} & 0.48e^{j4^\circ} & 0.13e^{j47^\circ} \\ 0.45e^{j93^\circ} & 0.14e^{-j101^\circ} & 0.88e^{-j28^\circ} \\ 0.41e^{-j135^\circ} & 0.85e^{j20^\circ} & 0.34e^{-j80^\circ} \\ 0.21e^{j78^\circ} & 0.44e^{-j158^\circ} & 0.88e^{j165^\circ} \\ 0.96e^{j81^\circ} & 0.26e^{j5^\circ} & 0.12e^{j10^\circ} \\ 0.45e^{-j27^\circ} & 0.83e^{j160^\circ} & 0.32e^{-j14^\circ} \\ 0.83e^{j20^\circ} & 0.49e^{j14^\circ} & 0.28e^{-j42^\circ} \\ 0.42e^{-j142^\circ} & 0.77e^{j158^\circ} & 0.48e^{-j60^\circ} \\ 0.52e^{j139^\circ} & 0.63e^{-j88^\circ} & 0.57e^{-j115^\circ} \end{bmatrix} \tag{A.3}$$

Appendix B
Pairwise Error Probability for Downlink

To prove the corollary, we consider $\left|\boldsymbol{\mathcal{W}}_{\mathsf{ee}}^{1/2}\boldsymbol{\Sigma}_{\mathsf{ee}}\boldsymbol{\mathcal{W}}_{\mathsf{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\mathsf{e}\,\dagger}\boldsymbol{\mathcal{F}}_{ji}^{\mathsf{e}}\right|_{\mathsf{NZ}}$, insert $\boldsymbol{\Sigma}_{\mathsf{ee}} = \left(\mathbf{1}_{e_{ij}\times 1} \otimes \mathbf{I}_N\right)\boldsymbol{\Sigma}_{kk}\left(\mathbf{1}_{1\times e_{ij}} \otimes \mathbf{I}_N\right)$, and make use of $|\mathbf{AB}|_{\mathsf{NZ}} = |\mathbf{BA}|_{\mathsf{NZ}}$ ([22, Theorem 1.3.20]):

$$\begin{aligned} &\left|\boldsymbol{\mathcal{W}}_{\mathsf{ee}}^{1/2}\boldsymbol{\Sigma}_{\mathsf{ee}}\boldsymbol{\mathcal{W}}_{\mathsf{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\mathsf{e}\,\dagger}\boldsymbol{\mathcal{F}}_{ji}^{\mathsf{e}}\right|_{\mathsf{NZ}} \\ &\quad= \left|\left(\mathbf{1}_{e_{ij}\times 1} \otimes \mathbf{I}_N\right)\boldsymbol{\Sigma}_{kk}\left(\mathbf{1}_{1\times e_{ij}} \otimes \mathbf{I}_N\right)\boldsymbol{\mathcal{W}}_{\mathsf{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\mathsf{e}\,\dagger}\boldsymbol{\mathcal{F}}_{ji}^{\mathsf{e}}\boldsymbol{\mathcal{W}}_{\mathsf{ee}}^{1/2}\right|_{\mathsf{NZ}} \\ &\quad= \left|\boldsymbol{\Sigma}_{kk}\left(\mathbf{1}_{1\times e_{ij}} \otimes \mathbf{I}_N\right)\boldsymbol{\mathcal{W}}_{\mathsf{ee}}^{1/2}\boldsymbol{\mathcal{F}}_{ji}^{\mathsf{e}\,\dagger}\boldsymbol{\mathcal{F}}_{ji}^{\mathsf{e}}\boldsymbol{\mathcal{W}}_{\mathsf{ee}}^{1/2}\left(\mathbf{1}_{e_{ij}\times 1} \otimes \mathbf{I}_N\right)\right|_{\mathsf{NZ}} \\ &\quad= \left|\boldsymbol{\Sigma}_{kk}\left(\sum_{l=1}^{e_{ij}}\left(\boldsymbol{\mathcal{S}}_{lj_l} - \boldsymbol{\mathcal{S}}_{li_l}\right)^{\dagger}\right)\left(\sum_{q=1}^{e_{ij}}\left(\boldsymbol{\mathcal{S}}_{qj_q} - \boldsymbol{\mathcal{S}}_{qi_q}\right)^{\dagger}\right)\right|_{\mathsf{NZ}}, \end{aligned}$$

where we inserted the definition for $\boldsymbol{\mathcal{F}}_{ji}^{\mathsf{e}}$ to obtain the last equation. Since $\boldsymbol{\mathcal{S}}_{ki_k} = \mathbf{I}_{N_{\mathrm{R}}} \otimes \mathbf{S}_{ki_k}$ and by the mixed product property of the Kronecker product [23, Lemma 4.2.10] we can continue

$$= \left| \mathbf{\Sigma}_{kk} \left(\mathbf{I}_{N_{\mathrm{R}}} \otimes \left(\sum_{l=1}^{e_{ij}} \sum_{q=1}^{e_{ij}} \left(\mathbf{S}_{lj_l} - \mathbf{S}_{li_l} \right)^{\dagger} \left(\mathbf{S}_{qj_q} - \mathbf{S}_{qi_q} \right)^{\dagger} \right) \right) \right|_{\mathrm{NZ}}$$

$$= \left| \mathbf{\Sigma}_{kk} \right| \left| \sum_{l=1}^{K} \sum_{q=1}^{K} \left(\mathbf{S}_{lj_l} - \mathbf{S}_{li_l} \right)^{\dagger} \left(\mathbf{S}_{qj_q} - \mathbf{S}_{qi_q} \right)^{\dagger} \right|^{N_{\mathrm{R}}},$$

where we made use of $|\mathbf{I}_n \otimes \mathbf{A}| = |\mathbf{A}|^n$ and our assumptions on the rank of $\mathbf{\Sigma}_{kk}$ (full rank N) and the sum of matrices (full rank N_{T}, the index of the sum can be extended beyond e_{ij}, because all the extra terms are zero anyway).

References

[1] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. 1999, Originally a Bell Laboratories, Lucent Technologies, Technical Report, Oct. 1995.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[3] G. J. Foschini, "Layered space–time architecture for wireless communication in fading environments when using multiple antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[4] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, no. 4, pp. 527–537, Apr. 1999.

[5] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space–time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[6] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space–time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[7] A. R. J. Hammons and H. El Gamal, "On the theory of space–time codes for PSK modulation," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 524–542, Mar. 2000.

[8] E. A. Fain and M. K. Varanasi, "Diversity order gain for narrowband multiuser communications with pre-combining group detection," *IEEE Trans. Commun.*, vol. 48, no. 4, pp. 533–536, Apr. 2000.

[9] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receivers in Rayleigh fading channels with applications to a unified analysis of coherent and noncoherent space–time receivers," *IEEE Trans. Inform. Theory*, vol. 47, no. 5, pp. 2383–2399, Sept. 2001.

[10] S. Siwamogsatham, M. P. Fitz, and J. Grimm, "A new view of performance analysis of transmit diversity schemes in correlated Rayleigh fading," *IEEE Trans. Inform. Theory*, vol. 48, no. 4, pp. 950–956, Apr. 2002.

[11] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

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

[13] M. O. Damen, K. Abed-Meraim, and J.-C. Belfiore, "Generalised sphere decoder for asymmetrical space–time communication architecture," *IEE Electronics Letters*, vol. 36, no. 2, pp. 166–167, Jan. 2000.

[14] S. Zhou, G. B. Giannakis, and C. Le Martret, "Chip-interleaved block-spread code division multiple access," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 235–248, Feb. 2002.

[15] J. K. Cavers and P. Ho, "Analysis of the error performance of trellis-coded modulations in Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 40, no. 1, pp. 74–83, Jan. 1992.

[16] E. Biglieri, H. L. Owen, and E. W. Zegura, "Computing error probabilities over fading channels: A unified approach," *European Trans. on Telecommun.*, vol. 9, no. 1, Feb. 1998.

[17] M. J. Barrett, "Error probability for optimal and suboptimal quadratic detectors in rapid Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. 5, no. 2, pp. 302–304, Feb. 1987.

[18] C. W. Helstrom, *Elements of Signal Detection & Estimation*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[19] M. Brehler and M. K. Varanasi, "Coherent multiuser space–time communications: Optimum receivers and signal design," in *Proc. Conf. Inform. Sciences and Systems*, Baltimore, MD, Mar. 2001, Johns Hopkins University.

[20] G. L. Turin, "The characteristic function of hermitian quadratic forms in complex normal variables," *Biometrika*, vol. 47, pp. 199–201, June 1960.

[21] M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques*, An IEEE Press Classic Reissue, New York, 1996, Originally A McGraw-Hill Publication, 1966.

[22] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1993.

[23] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1994.

[24] S. Verdú, *Multiuser Detection*, Cambridge Univ. Press, New York, NY, 1998.

[25] Z. Zvonar, "Combined multiuser detection and diversity reception for wireless CDMA systems," *IEEE Trans. Veh. Technol.*, vol. 45, no. 1, pp. 205–211, Feb. 1996.

[26] Z. Zvonar and D. Brady, "Multiuser detection in single-path fading channels," *IEEE Trans. Commun.*, vol. 42, no. 2/3/4, pp. 1729–1739, Feb./Mar./Apr. 1994.

[27] M. O. Damen, *Joint Coding/Decoding in a Multiple Access System, Application to Mobile Communications*, Ph.D. thesis, École Nationale Supérieure des Télécommunications, Paris, France, 1999.

[28] M. O. Damen, K. Abed-Meraim, and J.-C. Belfiore, "Diagonal algebraic space–time block codes," *IEEE Trans. Inform. Theory*, vol. 4, no. 3, pp. 628–636, Mar. 2002.

[29] M. O. Damen, A. Tewfik, and J.-C. Belfiore, "A construction of a space–time code based on number theory," *IEEE Trans. Inform. Theory*, vol. 4, no. 3, pp. 753–760, Mar. 2002.

[30] H. El Gamal and M. O. Damen, "Universal space–time coding," submitted to *IEEE Trans. Inform. Theory*, Jan. 2002.

[31] D. Agrawal, T. J. Richardson, and R. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2618–2626, Sept. 2001.

[32] F. D. Murnaghan, *The Unitary and Rotation Groups*, vol. III of *Lectures on Applied Mathematics*, Spartan Books, Washington, DC, 1962.

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

Matthias Brehler (S'00–M'03) started his studies in electrical engineering at the Technische Universität München in 1992. From 1995 to 1996, he attended the University of Colorado, Boulder, where he received the M.S. degree in electrical engineering. In 1998 he received the Diplom–Ingenieur degree from the Technische Universität München. After completing his military service at the Universität der Bundeswehr München (Federal Armed Forces University, Munich), he returned to Boulder in 1999, where he obtained his Ph.D. in electrical engineering in Fall 2002 and is currently working as a Research Associate. His research interests include multiuser, multi-carrier, and space-time communications.

Mahesh K. Varanasi (S'87–M'89–SM'95) received the B.E. degree in electronics and communication engineering from Osmania University, Hyderabad, India, in 1984, and the M.S. and Ph.D. degrees in electrical engineering from Rice University, Houston, TX, in 1987 and 1989, respectively. In 1989, he joined the faculty of the University of Colorado at Boulder in the Electrical and Computer Engineering Department where he is now a Professor. His teaching interests include communication theory, information theory, and signal processing. His research interests include multiuser detection, space-time communications, equalization, signal design, diversity communications over fading channels, and power- and bandwidth-efficient multiuser communications.

Fig. 1. In the multiuser narrowband channel, every user achieves full order of diversity. However, for an increasing number of users an increasing SNR penalty is incurred. The lines marked with a triangle correspond to an analytic upper bound, the plain line styles to simulations (for $K > 1$).

Fig. 2. If the number of users K increases and D stays fixed at two, the asymptotic BER is reached for increasing SNR only.

Fig. 3. With only $D = 2$ dimensions the 10 user system can asymptotically achieve single-user performance and out-perform the $K = 10$ narrowband system by 15 dB at a BER of 10^{-2}.

Fig. 4. For $N_{\mathrm{T}} = 2$ receive antennas all systems improve when compared to Figure 3, but the single-user employing 32-QAM modulation does not benefit as much from the additional receive antenna as the other systems.

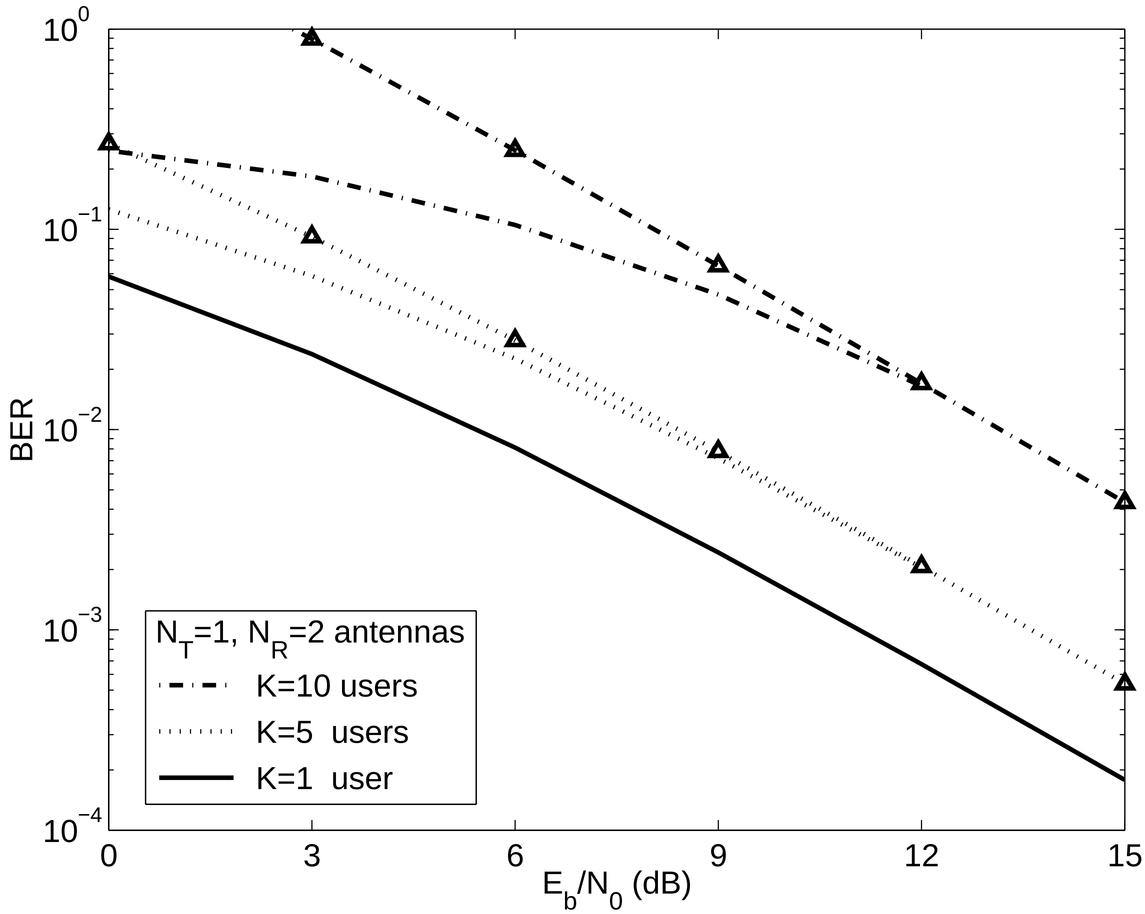
Fig. 5. Increasing the number of dimensions seems to gain nothing when compared to a system with $D = 2$ and the same spectral efficiency.

Fig. 6. For increasing constellation size, the multiuser system can still asymptotically achieve single-user like performance.

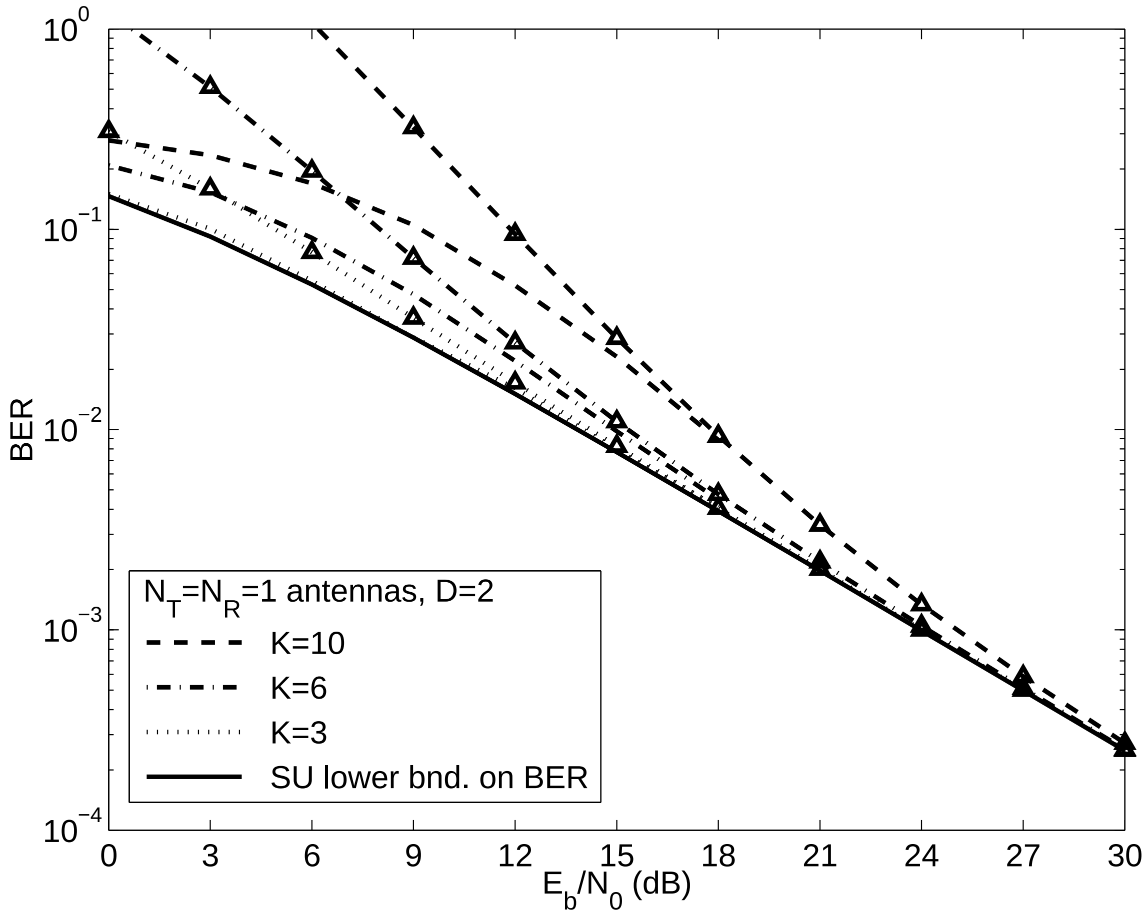
Fig. 7. Each user employs the Alamouti scheme and asymptotically achieves single-user like performance. The $K = 3$, 4-PSK system has the same spectral efficiency as the $K = 1$, 16-QAM system, but the latter has a 3 dB worse energy efficiency at BER 10^{-3}.

Fig. 8. For a spectral efficiency of 8 bits/dimension orthogonal users would have to employ 256-QAM, leading to a dismal performance when compared to low-dimensional spreading.

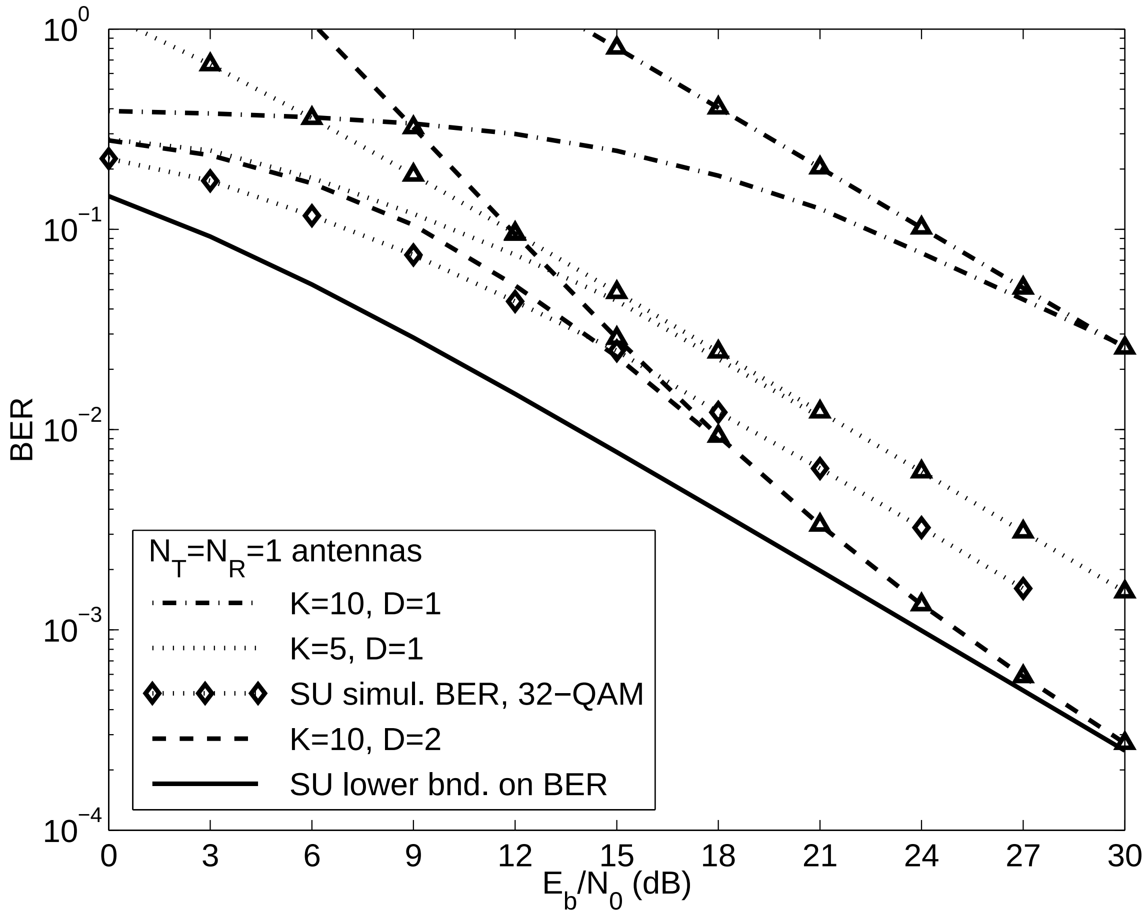
Fig. 9. For the downlink it seems to be more energy-efficient to increase spectral efficiency by increasing the constellation size of each (orthogonal) user rather than trying to communicate in a common signal space for several users.



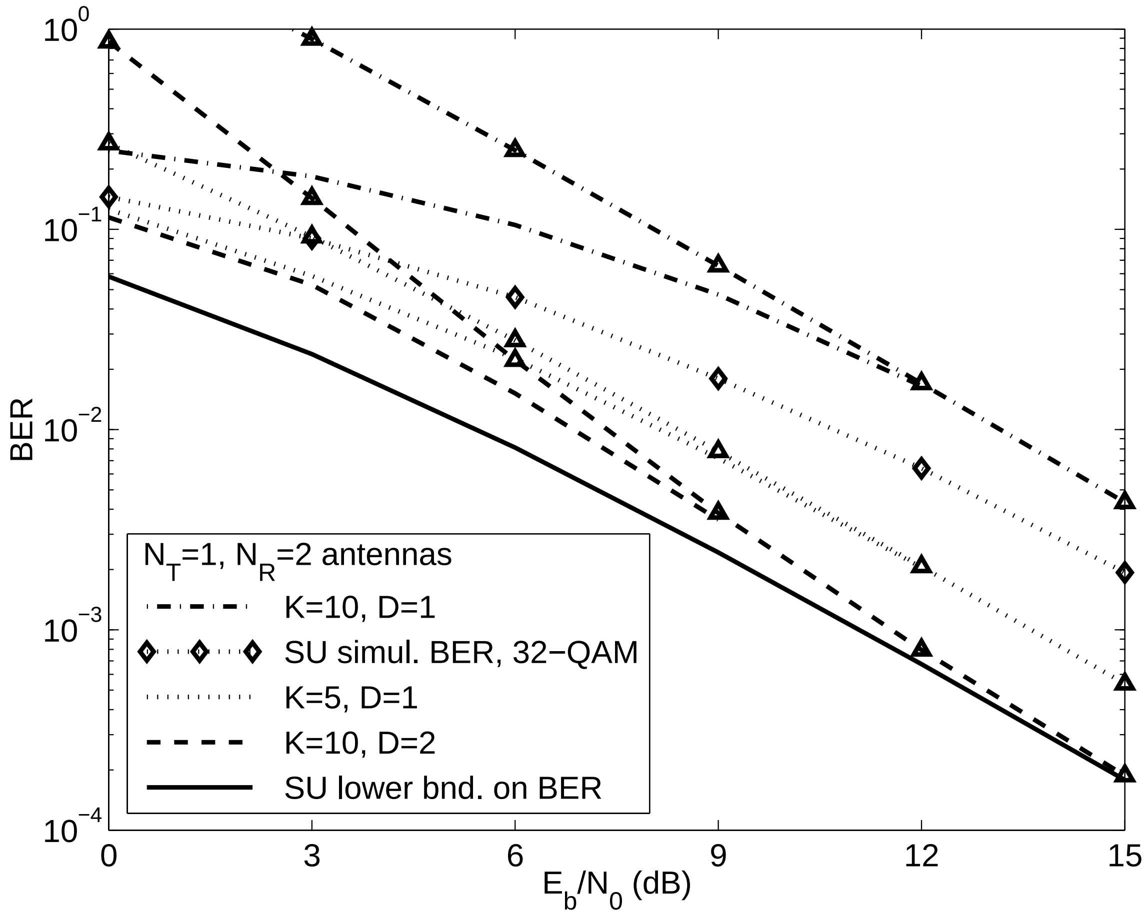
10^0
10^−1
10^−2
10^−3
10^−4
BER
0
3
6
9
12
15
E_b/N_0 (dB)
N_T=1, N_R=2 antennas
K=10 users
K=5 users
K=1 user



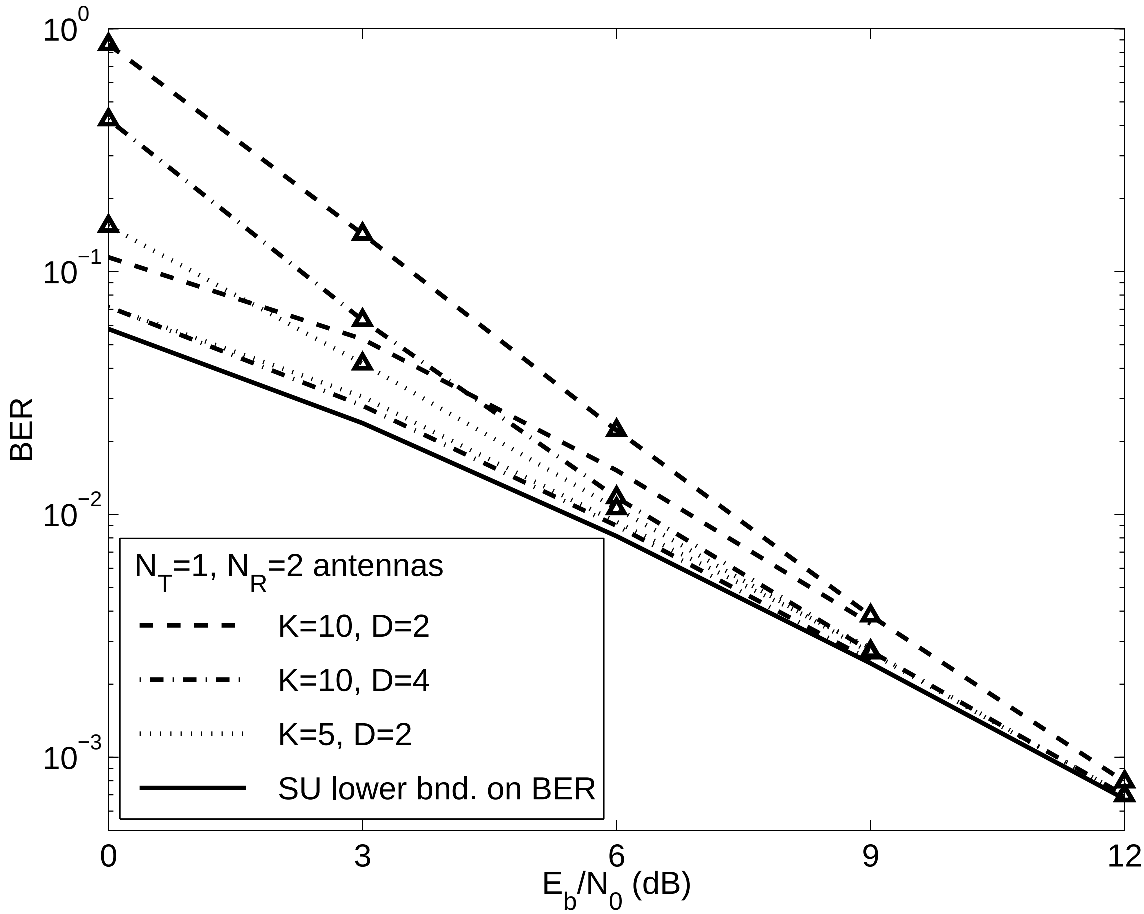
BER
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
0
3
6
9
12
15
18
21
24
27
30
E_b/N_0 (dB)
$N_T=N_R=1$ antennas, D=2
K=10
K=6
K=3
SU lower bnd. on BER



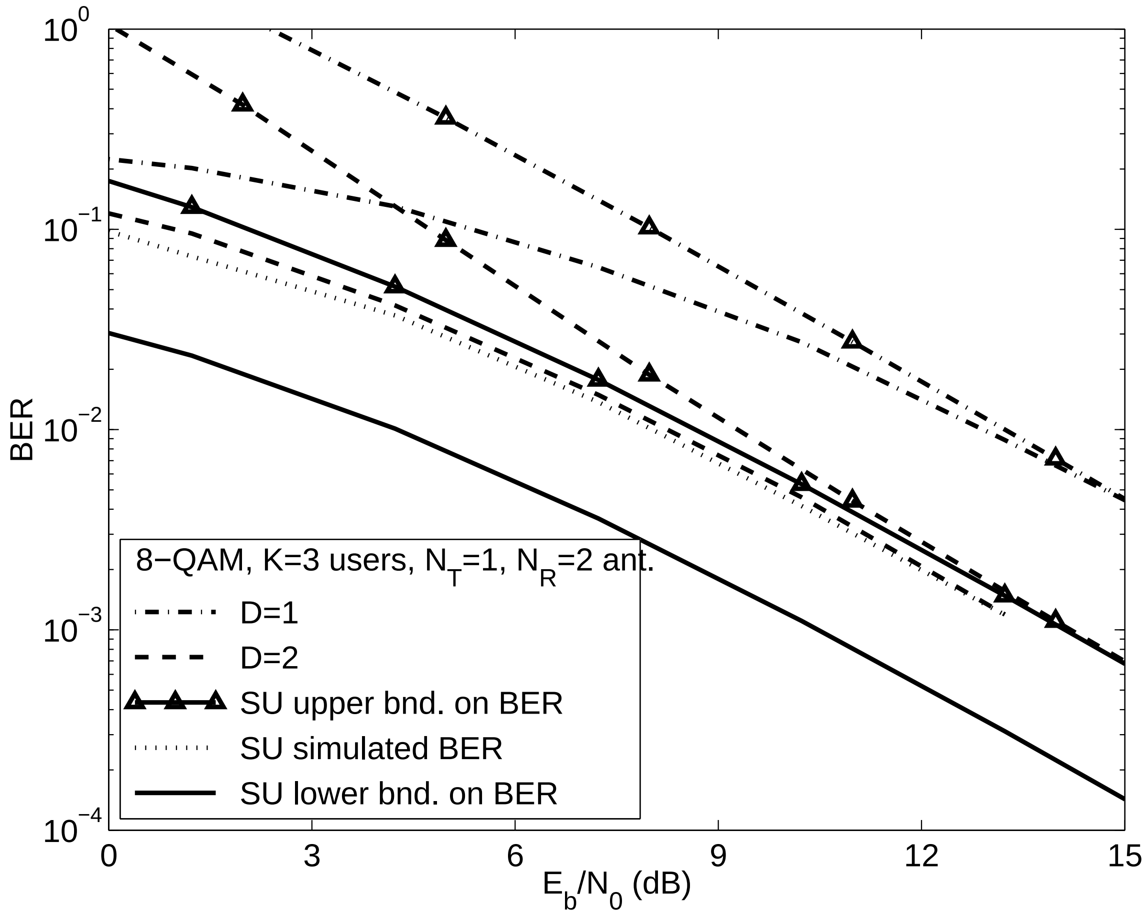
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
BER
0
3
6
9
12
15
18
21
24
27
30
E_b/N_0 (dB)
N_T=N_R=1 antennas
K=10, D=1
K=5, D=1
SU simul. BER, 32−QAM
K=10, D=2
SU lower bnd. on BER



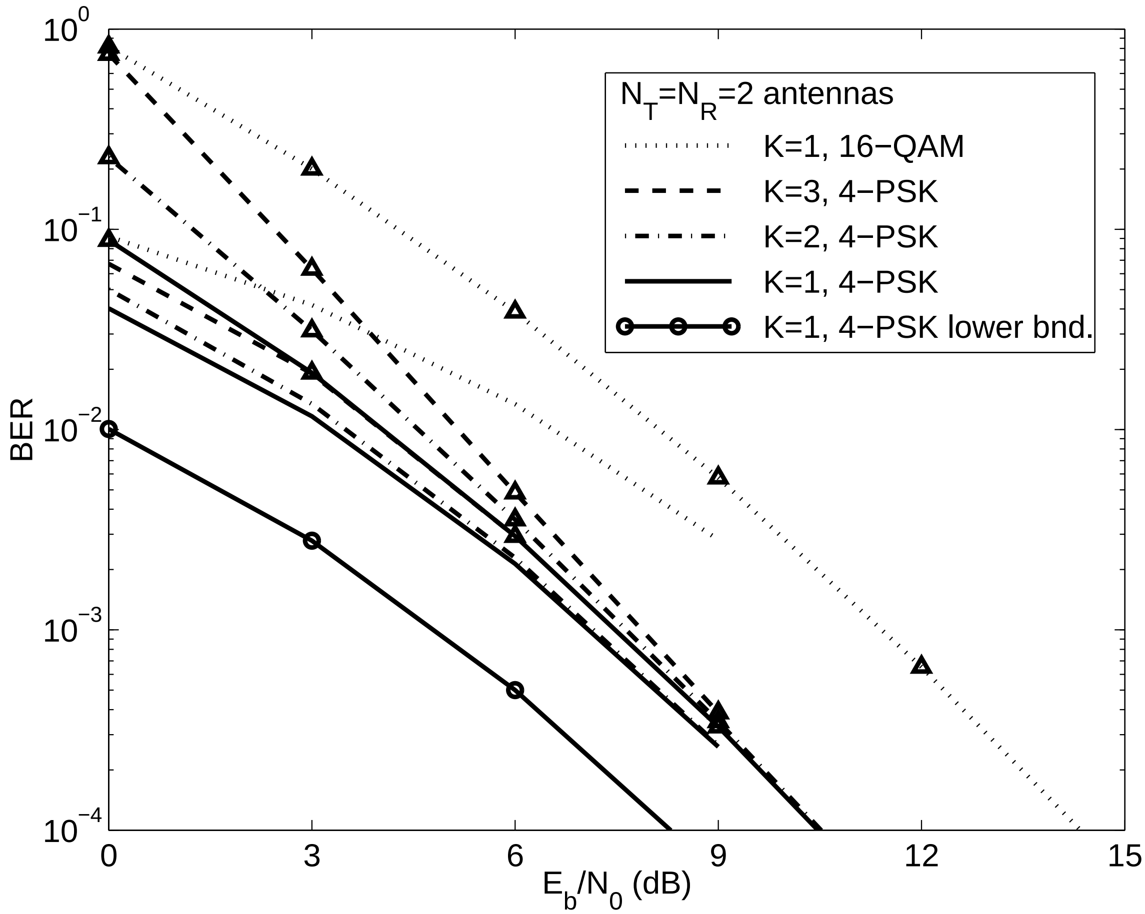
BER
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
0
3
6
9
12
15
E_b/N_0 (dB)
N_T=1, N_R=2 antennas
K=10, D=1
SU simul. BER, 32−QAM
K=5, D=1
K=10, D=2
SU lower bnd. on BER



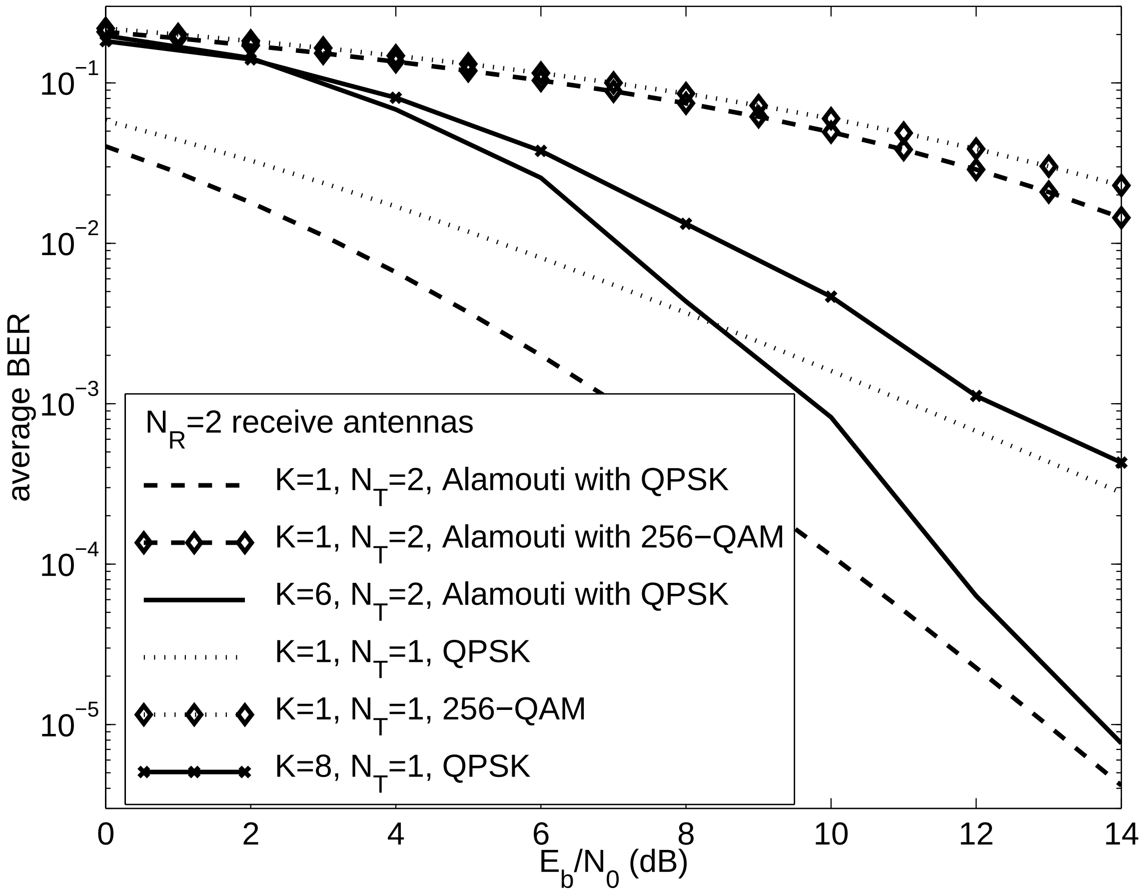
10^0
10^−1
10^−2
10^−3
BER
N_T=1, N_R=2 antennas
K=10, D=2
K=10, D=4
K=5, D=2
SU lower bnd. on BER
0
3
6
9
12
E_b/N_0 (dB)



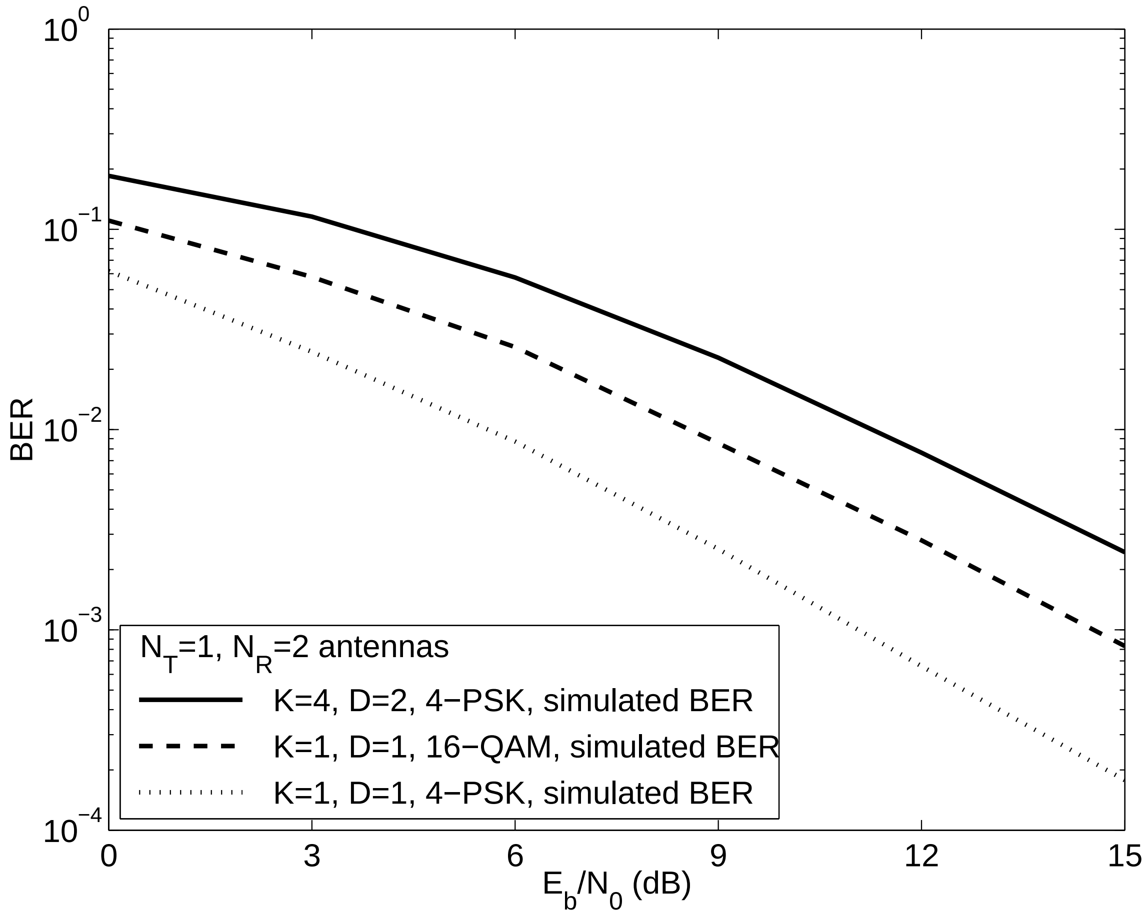
10^0
10^−1
10^−2
10^−3
10^−4
BER
0
3
6
9
12
15
E_b/N_0 (dB)
8−QAM, K=3 users, N_T=1, N_R=2 ant.
D=1
D=2
SU upper bnd. on BER
SU simulated BER
SU lower bnd. on BER



N_T=N_R=2 antennas
K=1, 16−QAM
K=3, 4−PSK
K=2, 4−PSK
K=1, 4−PSK
K=1, 4−PSK lower bnd.
BER
E_b/N_0 (dB)



average BER
E_b/N_0 (dB)
N_R=2 receive antennas
K=1, N_T=2, Alamouti with QPSK
K=1, N_T=2, Alamouti with 256−QAM
K=6, N_T=2, Alamouti with QPSK
K=1, N_T=1, QPSK
K=1, N_T=1, 256−QAM
K=8, N_T=1, QPSK
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
0
2
4
6
8
10
12
14



BER
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
0
3
6
9
12
15
E_b/N_0 (dB)
N_T=1, N_R=2 antennas
K=4, D=2, 4−PSK, simulated BER
K=1, D=1, 16−QAM, simulated BER
K=1, D=1, 4−PSK, simulated BER