Asymptotic Error Probability Analysis of Quadratic Receivers in Rayleigh-Fading Channels With Applications to a Unified Analysis of Coherent and Noncoherent Space–Time Receivers

Matthias Brehler, *Student Member, IEEE*, and Mahesh K. Varanasi, *Senior Member, IEEE*

***Abstract*—A general, asymptotic (high signal-to-noise (SNR)) error analysis is introduced for quadratic receivers in frequency-flat and multipath Rayleigh-fading channels with multiple transmit and receive antennas. Asymptotically tight expressions for the pairwise error probabilities are obtained for coherent, noncoherent, and differentially coherent space–time receivers. Not only is our unified analysis applicable to more general modulation schemes and/or channel models than previously considered, but it also reveals a hitherto unrecognized eigenvalue structure that is common to all of these problems. In addition to providing an easy recipe for computing the asymptotic pairwise error rates, we make some conclusions regarding criteria for the design of signal constellations and codes such as a) the same design criteria apply for both correlated and independent and identically distributed (i.i.d.) fading processes and b) for noncoherent communications, unitary signals are optimal in the sense that they minimize the asymptotic union bound.**

***Index Terms*—Asymptotics, error analysis, fading channels, maximum-likelihood (ML) receivers, space–time modulation.**

I. Introduction

INFORMATION-theoretic results promise considerable capacity gains in wireless communication systems that employ multiple transmit and receive antennas for coherent [1], [2] and noncoherent [3], [4] reception. Motivated by these promises, several researchers have recently proposed multiantenna coding and modulation schemes for coherent (cf. [5]–[8]), noncoherent (cf. [9]), and differentially coherent (cf. [10]–[12]) multiantenna communication systems (these references are far from being exhaustive, but give a nice overview). These works are based on the performance analyses of the corresponding optimum receivers. So far, the only commonality in these analyses appears to be the Chernoff bound used to bound the pairwise error probabilities.

Manuscript received August 9, 2000; revised March 1, 2001. This work was supported in part by the National Science Foundation under Grant ANI-9725778, and in part by the U.S. Army Research Office under Grant DADD19-99-1-0291.

The authors are with the Department of Electrical and Computer Engineering, University of Colorado at Boulder, Boulder, CO 80309-0425 USA (e-mail: brehler@dsp.colorado.edu; varanasi@dsp.colorado.edu).

Communicated by M. L. Honig, Associate Editor for Communications.
Publisher Item Identifier S 0018-9448(01)06218-6.

In this paper, we adopt the point of view and terminology of modulation theory in that we consider sequences of symbols over the multiple transmit antennas in a block over which the fading is constant as a supersymbol. Note that the physical transmission of a such a supersymbol requires sending a signal matrix. The individual symbols in these signal matrices may be drawn from a regular quadrature amplitude modulation (QAM)-like alphabet (in this case, the signal matrix is a space–time "codeword") or they may be arbitrary complex numbers that result from constellation designs that do not *a priori* restrict them to a predetermined alphabet. In either case, we choose to use the more general terms "modulation" and "detection" (with respect to the constellation of signal matrices and their reception) rather than the terms coding and decoding. The latter terminology is, of course, implied whenever the transmitted signal matrices can be regarded as codewords belonging to a codebook.

This paper offers a unified approach to the asymptotic (high signal-to-noise ratio (SNR)) performance analysis of coherent, noncoherent, and differentially coherent quadratic receivers for Rayleigh-fading channels.[1] In particular, we obtain asymptotically tight expressions for pairwise error probabilities, and in the process, reveal a common eigenvalue structure in the analysis of such quadratic receivers for space–time modulation. In so doing, we also address more general modulation schemes and channel models than previously considered. For instance, our analysis applies to noncoherent modulation schemes that do not restrict the signal matrices to be of equal energy with orthonormal columns (referred to as unitary signals [9]). Moreover, we do not restrict the signals to be unitary matrices (not to be confused with unitary signals [9]) for differential modulation as do [10], [11]. As for general channel models, we consider frequency-flat as well as multipath channels in which the fading processes corresponding to the various transmit/receive antenna pairs need not be independent of each other. Such correlations do indeed arise in many practical applications [17, Secs. 1.6 and 5.3].

Since for correlated fading the noncoherent maximum-likelihood (ML) receiver requires the knowledge of the fading statistics and the average SNR, a generalized likelihood ratio test

[1] This approach was originally developed in the context of multiuser detection [13], [14], where it was inspired by [15], [16].

(GLRT) receiver is introduced and analyzed that does not require any knowledge of the fading processes.

As far as differentially coherent modulation is concerned, the asymptotic performance analysis is performed not only for constant fading over two symbol intervals (denoted as slow fading in this paper), but also for fast fading, when the fading coefficients change from one symbol interval to the next (but are still essentially constant over one symbol interval). An expression for the asymptotic error floor of the pairwise symbol error probability is obtained for the fast fading case.

The application of the general asymptotic analysis to coherent, noncoherent, and differentially coherent space–time modulation schemes underlines the commonalities in the analyses of these problems. Due to the general analysis, determining the asymptotic pairwise error probability reduces to determining the asymptotic eigenvalues of certain matrices, a slightly unusual, but by no means an overdemanding task.

As a result of our analysis, we arrive at several conclusions with regards to the problems of signal constellation design.

- Although the error probability depends on the fading correlation, the asymptotic performance criteria for designing constellations that achieve full order of diversity for the correlated fading channel are identical to those for independent and identically distributed (i.i.d.) fading. Hence, constellations that are designed for the latter will also perform well in the case of correlated fading. It is only for differentially coherent modulation and detection, that minor modifications are required in some cases.
- In the noncoherent case with a restriction to equal-energy constellations, the so-called unitary signals are shown to minimize the asymptotic union bound on the symbol error rate of the maximum-likelihood (ML) receiver.
- For noncoherent modulation, the GLRT and the ML receivers are equivalent for unitary signaling and i.i.d. fading. For unitary signaling and correlated fading, we show that the ML receiver asymptotically converges to the GLRT receiver, and consequently, that the GLRT receiver's performance asymptotically matches the ML receiver's performance. We conclude that unitary signals are also optimal for the GLRT receiver.
- For differential modulation and slow fading, the GLRT receiver is asymptotically optimum.

The paper is divided into five sections. A description of the multiple transmit and receive antenna communication system and Rayleigh-fading channel models is given in Section II; this description motivates the formulation of a general hypothesis testing problem. For this problem, an analysis of the pairwise symbol error probability is obtained in Section III, including the asymptotically tight analysis for a specific eigenvalue structure. This analysis is then seen to be applicable to the different space–time modulation schemes in Section IV. We conclude in Section V.

Throughout the paper, $\top$ denotes transpose, $*$ complex conjugate, and $\dagger$ complex conjugate transpose. The multivariate circularly symmetric, complex Gaussian distribution with mean vector $\boldsymbol{m}$ (and covariance matrix $\boldsymbol{K}$) is denoted by $\mathcal{CN}(\boldsymbol{m})$ ($\mathcal{CN}(\boldsymbol{m}, \boldsymbol{K})$). $E[\cdot]$ denotes the expected value of the expression in brackets. For any matrix $\boldsymbol{A}$ we write its determinant as $|\boldsymbol{A}|$ and its trace as $\mathrm{tr}(\boldsymbol{A})$. For any vector $\boldsymbol{a}$, we write its ℓ_2 norm as $\sqrt{\boldsymbol{a}^\dagger \boldsymbol{a}} = ||\boldsymbol{a}||$. The logarithm to the base 10 is denoted by $\lg$, the natural logarithm by $\ln$. The Kronecker product of two matrices is denoted by $\otimes$ [18]. The N-dimensional identity matrix (matrix of all zeros) is denoted by $\boldsymbol{I}_N$ ($\mathbf{0}_N$); the $M \times N$ matrix of all zeros by $\mathbf{0}_{M \times N}$.

II. System Model

In this section, we describe communication systems that employ *space-dimension* modulation (space–time modulation being one particular realization) and operate over frequency-flat or multipath Rayleigh-fading channels. We introduce the model by first assuming frequency flatness in Section II-A, and we show in Section II-B that a similar description applies for multipath fading.

Unlike previous works which start with a discrete-time model, we begin with a continuous-time description as this is necessary to develop the discrete-time model for the multipath channel. Moreover, our continuous-time description has greater generality in that it does not restrict us to a particular choice of basis functions used in the representation of the transmitted signals. In previous works on space–time modulation, it is implicitly assumed that time-translates of a single waveform are used as basis functions. Consequently, the number of dimensions D is identified as the "length of the coherence interval" (K, l, and T, respectively, in [6], [7], [9]). However, the choice of basis functions is dictated by bandwidth considerations and by the time and frequency selectivity of the channel and constraining it to be time-translates of one waveform is unnecessary and might be too restrictive. For instance, time selectivity imposes the restriction that the basis functions be time limited. Frequency selectivity has the effect of creating multiple paths when wide-band basis functions are used. Consequently, the choice of an appropriate basis is connected to the problem of time–frequency localization. In effect, our model description suggests the more general view of "space-dimension" modulation (coding), with "space–time" modulation as one possible manifestation. On the other extreme, one could use a basis of band-limited functions that are almost nonoverlapping in frequency, the so-called Chang pulses, as described in [19], that satisfy the generalized Nyquist criterion. In this case, provided the fading is frequency flat over the whole bandwidth of the basis, the usual "space–time" codes can be applied but should be more accurately called "space–frequency" codes.

In both frequency-flat and multipath fading channels, one out of S signals (a supersymbol) is sent from the M transmit antennas. The mth component of the ith ($i \in \{1, \ldots, S\}$) signal vector $\boldsymbol{s}^i(t) = [s_1^i(t), \ldots, s_M^i(t)]^\top$ in the kth supersymbol interval is sent from the mth antenna and is a superposition of D orthonormal basis functions $\boldsymbol{u}(t) = [u_1(t), \ldots, u_D(t)]^\top$, which satisfy the generalized Nyquist criterion [20]. Hence, the signal vector $\boldsymbol{s}^i(t)$ can be written as

$$\boldsymbol{s}^i(t) = \boldsymbol{S}_i^\top \boldsymbol{u}(t - kT) \tag{1}$$

where T is the supersymbol interval and $\boldsymbol{S}_i \in \mathbb{C}^{D\times M}$ contains the expansion coefficients of the signals $\boldsymbol{s}^i(t)$ with respect to the basis functions $\boldsymbol{u}(t)$.

A. Frequency-Flat Fading

In the case of frequency-flat fading, the received signal $r_n(t)$ at the nth-receiver antenna can be written as

$$\begin{aligned} r_n(t) &= \sqrt{\overline{\gamma}} \sum_{m=1}^{M} h_{mn}(k) s_m^i(t) + \eta_n(t) \\ &= \sqrt{\overline{\gamma}}\, \boldsymbol{s}^i(t)^\top \boldsymbol{h}_n(k) + \eta_n(t) \end{aligned}$$

where $\overline{\gamma}$ is the average received SNR per supersymbol per receiver antenna. The additive noise $\eta_n(t)$ is white and $\mathcal{CN}(0,\, 1)$ distributed. $h_{mn}(k)$ is the fading coefficient from the mth transmit to the nth receive antenna, assumed to be $\mathcal{CN}(0)$ distributed and constant during the (super)symbol interval k. Stacking the fading coefficients of the transmit antennas yields $\boldsymbol{h}_n(k) = [h_{1n}(k),\, \ldots,\, h_{Mn}(k)]^\top$.

The D sufficient statistics $\boldsymbol{y}_n$ for the received signal $r_n(t)$ of the nth antenna are obtained by matched filtering with respect to the basis functions

$$\boldsymbol{y}_n(k) = \int_{-\infty}^{\infty} \boldsymbol{u}^*(t-kT) r_n(t)\, dt = \sqrt{\overline{\gamma}}\, \boldsymbol{S}_i \boldsymbol{h}_n(k) + \boldsymbol{\eta}_n(k). \tag{2}$$

Note that there is no inter-supersymbol interference (ISI), because the basis functions are assumed to satisfy the generalized Nyquist criterion. Stacking these receiver antenna observations yields the DN sufficient statistics of the kth-symbol interval as

$$\boldsymbol{y}(k) = \left[\boldsymbol{y}_1^\top(k),\, \ldots,\, \boldsymbol{y}_N^\top(k)\right]^\top = \sqrt{\overline{\gamma}}\, \boldsymbol{\mathcal{S}}_i \boldsymbol{h}(k) + \boldsymbol{\eta}(k) \tag{3}$$

where

$$\begin{aligned} \boldsymbol{\mathcal{S}}_i &= \boldsymbol{I}_N \otimes \boldsymbol{S}_i \\ \boldsymbol{h}(k) &= [\boldsymbol{h}_1^\top(k),\, \ldots,\, \boldsymbol{h}_N^\top(k)]^\top \end{aligned}$$

and

$$\boldsymbol{\eta}(k) = [\boldsymbol{\eta}_1^\top(k),\, \ldots,\, \boldsymbol{\eta}_N^\top(k)]^\top.$$

While the noise processes are white in space and time ($E[\boldsymbol{\eta}(k)\boldsymbol{\eta}^\dagger(l)] = \delta(k-l)\boldsymbol{I}_{DN}$), the model allows for correlation in the fading processes, i.e., $\boldsymbol{\Sigma} = E[\boldsymbol{h}(k)\boldsymbol{h}^\dagger(k)]$ is not required to be proportional to identity. Since this work is concerned with supersymbol detection and modulation (and not coding over multiple supersymbols), it will be convenient to drop the time index k in most of what follows.

The energy of a signal $\boldsymbol{S}_i$ is denoted by $E_i = \mathrm{tr}(\boldsymbol{S}_i^\dagger \boldsymbol{S}_i)$. The average energy in the S signals is constrained to be unity, where the average is taken over the number of signals and transmit antennas, i.e.,

$$\sum_{i=1}^{S} E_i = \sum_{i=1}^{S} \mathrm{tr}\left(\boldsymbol{S}_i^\dagger \boldsymbol{S}_i\right) = MS. \tag{4}$$

Note that the normalization chosen in [7], [3], [9] would require the sum of the traces to be DMS.

The fading processes are normalized so that $\overline{\gamma}$ represents the average received SNR per supersymbol and receiver antenna, where the average is taken over the fading and all signals, i.e.,

$$E\left[\boldsymbol{h}(k)^\dagger \boldsymbol{\mathcal{S}}_i^\dagger \boldsymbol{\mathcal{S}}_i \boldsymbol{h}(k)\right] = N. \tag{5}$$

For equiprobable symbols, this leads to the condition

$$\sum_{i=1}^{S} \mathrm{tr}\left(\boldsymbol{\Sigma} \boldsymbol{\mathcal{S}}_i^\dagger \boldsymbol{\mathcal{S}}_i\right) = NS. \tag{6}$$

Note that for unitary signals (where $\boldsymbol{S}_i^\dagger \boldsymbol{S}_i = \boldsymbol{I}$, for all i) and i.i.d. fading, this amounts to $M\boldsymbol{\Sigma} = \boldsymbol{I}_{MN}$.

With these normalizations the average SNR $\overline{\gamma}$ is independent of the number of transmit antennas. At this juncture, we do not require that all the signals have the same energy (as does [9]), and moreover, the power in the transmit antennas is also allowed to be different within the signals.

B. Multipath Fading

In outdoor wireless communications, multiple signal paths from the transmitter to the receiver cannot be usually avoided, particularly for high-rate communications (which is the ultimate goal of space–time methods). We show in this section that the frequency-flat system description of the previous subsection can easily be extended to the multipath fading case by introducing the *effective* signal matrix. More specifically, we show the necessary changes to the system model when ISI masking is employed in cases where the system parameters allow us to do so. To avoid the issue of ISI, we could also consider a one-shot transmission of a supersymbol (signal), a case obviously included in our discussion but that is less general.

We assume in this section that the basis functions employed are strictly time-limited to supersymbol intervals T. With this assumption, consider a channel with maximum delay spread T_m and a D-dimensional basis so that the signal bandwidth W is roughly D/T. Thus, the number of resolvable paths L for the usual tapped-delay-line model (cf. [21, Sec. 14-5-1]) is equal to $L = DT_m/T$ and the tap delay is $1/W = T/D$. To avoid having to deal with ISI, we choose to mask-out the ISI resulting from the multiple paths (cf. [22]). The incurred loss of optimality by this masking is small, if the total delay spread $T_m = L/W$ amounts only to a small fraction of the supersymbol interval T. The requirement $T_m \ll T$ is equivalent to $L \ll D$, which limits the application of the following to high-dimensionality cases: roughly, to limit the loss in energy efficiency due to ISI masking to about 10%, the number of dimensions should be ten times higher than the number of resolvable paths. Under the stated assumptions and with the notation introduced above, the received signal in the kth supersymbol interval becomes

$$\begin{aligned} r_n(t) &= \sqrt{\overline{\gamma}} \sum_{l=1}^{L} \sum_{m=1}^{M} h_{m\,ln}(k) s_m^i\left(t - \frac{l}{W}\right) + \eta_n(t) \\ &= \sqrt{\overline{\gamma}} \sum_{l=1}^{L} \boldsymbol{u}^\top\left(t - kT - \frac{l}{W}\right) \boldsymbol{S}_i \boldsymbol{h}_{ln}(k) + \eta_n(t) \\ &= \sqrt{\overline{\gamma}}\, \hat{\boldsymbol{u}}^\top(t-kT) \left(\boldsymbol{I}_L \otimes \boldsymbol{S}_i\right) \boldsymbol{h}_n(k) + \eta_n(t) \end{aligned}$$

where

$$\hat{\boldsymbol{u}}(t) = [\boldsymbol{u}^\top(t), \boldsymbol{u}^\top(t - \frac{l}{W}), \ldots, \boldsymbol{u}^\top(t - \frac{L}{W})]^\top$$

$h_{m\,ln}(k)$ is the fading coefficient of the lth path from the mth transmit antenna to the nth receive antenna, $\boldsymbol{h}_{ln}$ contains the M fading coefficients of the lth path at receiver antenna n, and $\boldsymbol{h}_n = [\boldsymbol{h}_{1n}^\top, \ldots, \boldsymbol{h}_{Ln}^\top]^\top$. We match filter with respect to $\hat{\boldsymbol{u}}(t)$ and mask out the ISI by

$$\begin{aligned}\boldsymbol{z}_n(k) &= \int_{t=kT+L/W}^{(k+1)T} \hat{\boldsymbol{u}}^*(t-kT) r_n(t)\,dt \\ &= \sqrt{\gamma}\boldsymbol{K}_{\hat{\boldsymbol{u}}\hat{\boldsymbol{u}}}(\boldsymbol{I}_L \otimes \boldsymbol{S}_i)\boldsymbol{h}_n(k) + \boldsymbol{\eta}_n(k)\end{aligned}$$

where

$$\boldsymbol{K}_{\hat{\boldsymbol{u}}\hat{\boldsymbol{u}}} = \int_{t=kT+L/W}^{(k+1)T} \hat{\boldsymbol{u}}^*(t)\hat{\boldsymbol{u}}^\top(t)\,dt$$

(the delayed and truncated basis functions are not necessarily orthonormal to each other) and $\boldsymbol{\eta}_n(k)$ is a DL-dimensional $\mathcal{CN}(\boldsymbol{0}_{DL\times 1}, \boldsymbol{K}_{\hat{\boldsymbol{u}}\hat{\boldsymbol{u}}})$ distributed noise vector. Multiplying the sufficient statistics $\boldsymbol{z}(k)$ with $\boldsymbol{K}_{\boldsymbol{uu}}^{-(1/2)}$ yields the equivalent statistics $\boldsymbol{y}_n(k)$ as

$$\boldsymbol{y}_n(k) = \sqrt{\gamma}\boldsymbol{K}_{\hat{\boldsymbol{u}}\hat{\boldsymbol{u}}}^{1/2}(\boldsymbol{I}_L \otimes \boldsymbol{S}_i)\boldsymbol{h}_n(k) + \boldsymbol{\eta}_n(k) \tag{7}$$

where $\boldsymbol{\eta}_n(k)$ is now a zero-mean white noise vector. We note that (7) and (2) are essentially the same, with an increase in the number of observations from D to DL and the replacement of $\boldsymbol{S}_i$ of (2) by $\boldsymbol{S}_i^{\mathrm{MP}} = \boldsymbol{K}_{\hat{\boldsymbol{u}}\hat{\boldsymbol{u}}}^{1/2}(\boldsymbol{I}_L \otimes \boldsymbol{S}_i)$. We call this latter matrix the *effective* signal matrix for multipath fading. The remaining model description (stacking the per-receive antenna observations in one vector) and the analysis of the frequency-flat fading case are thus also valid for multipath fading, with the signal matrix $\boldsymbol{S}_i$ replaced by $\boldsymbol{S}_i^{\mathrm{MP}}$. The order of diversity will consequently increase from MN to LMN, if certain rank criteria are met. Multipath fading can thus be exploited to yield higher orders of diversity and must not be seen as an additional interference that has to be combated.

III. Quadratic Receivers

Since $\boldsymbol{h}$ and $\boldsymbol{\eta}$ are $\mathcal{CN}(\boldsymbol{0})$ distributed, so are the sufficient statistics $\boldsymbol{y}$. In detection problems involving such observations, the ML or the GLRT receivers are easily shown to yield decision rules that are based on quadratic forms in $\boldsymbol{y}$. In this section, we discuss very generally the pairwise symbol error probability of quadratic receivers in $\mathcal{CN}(\boldsymbol{0})$ distributed random variables.

A quadratic receiver Φ is defined by

$$\Phi\colon \hat{i} = \arg\min_{1\le i\le S} \boldsymbol{x}^\dagger \boldsymbol{F}_i \boldsymbol{x} + c_i = \arg\min_{1\le i\le S} \delta_i \tag{8}$$

where δ_i is implicitly defined, $\boldsymbol{F}_i$ is a Hermitian matrix, c_i a real constant, and $\boldsymbol{x}$ a vector of $\mathcal{CN}(\boldsymbol{0}, \boldsymbol{K}_{\boldsymbol{xx}|\boldsymbol{S}_i})$ distributed random variables. Note that $\boldsymbol{F}_i$, c_i, and $\boldsymbol{K}_{\boldsymbol{xx}|\boldsymbol{S}_i}$ depend on the transmitted signal $\boldsymbol{S}_i$.

Let $\mathcal{E}(\Phi)$ be the event that the receiver Φ decides for a wrong symbol. The evaluation of the conditional symbol error probability $\Pr\{\mathcal{E}(\Phi)|\boldsymbol{S}_i\}$, the probability that a symbol other than $\boldsymbol{S}_i$ is detected when $\boldsymbol{S}_i$ is transmitted, would require the evaluation of the probability of the union of the $S-1$ events $(\delta_j < \delta_i)$ $(1 \le j \le S,\, j \ne i)$. Since the probability of the union is usually not computable, one has to resort to a union bound, and upper-bound it by the sum of the probabilities $\Pr\{\delta_j < \delta_i\}$; these probabilities are customarily called pairwise error probabilities, and we adopt this term.[2] The symbol error probability $\Pr\{\mathcal{E}(\Phi)\}$ for equiprobable symbols is then bounded by

$$\begin{aligned}\Pr\{\mathcal{E}(\Phi)\} &= S^{-1}\sum_{i=1}^{S} \Pr\{\mathcal{E}(\Phi)|\boldsymbol{S}_i\} \\ &\le S^{-1}\sum_{i=1}^{S}\sum_{\substack{j=1\\ j\ne i}}^{S} \Pr\{\delta_j < \delta_i\}.\end{aligned}$$

In the following sections, we find expressions for $\Pr\{\delta_j < \delta_i\}$ and analyze their asymptotic (high SNR) behavior. While the expressions for the pairwise error probability are known (cf. [23]–[25]), it is the asymptotic analysis for the pairwise error probabilities that is simplified (relative to [16]) and made more accurate than the Chernoff-bound-based approaches (cf. [7], [9]). In particular, asymptotically tight expressions for $\Pr\{\delta_j < \delta_i\}$ are obtained. Due to the general analysis, the task of finding the asymptotic performance of quadratic receivers in Rayleigh-fading channels is reduced to determining the asymptotic eigenvalues of certain matrices.

A. Pairwise Symbol Error Probability

The pairwise probability of error is the probability that a Hermitian quadratic form in the $\mathcal{CN}(\boldsymbol{0})$ distributed random vector $\boldsymbol{x}$ is less than some real constant

$$\Pr\{\delta_j < \delta_i\} = \Pr\left\{\boldsymbol{x}^\dagger \boldsymbol{F}_{ij}\boldsymbol{x} < c_{ij}\right\} \tag{9}$$

where $\boldsymbol{F}_{ij} = \boldsymbol{F}_j - \boldsymbol{F}_i$ and $c_{ij} = c_i - c_j$. The cumulative distribution function of a Hermitian quadratic form in $\mathcal{CN}(\boldsymbol{0})$ random variables is derived in Appendix A. Applying the result to $\Pr\{\delta_j < \delta_i\}$ yields the following proposition, also found in [23], [24], [13]. However, our derivation in Appendix A is different from [23], [24] and is necessary for a rigorous asymptotic analysis that will follow.

Proposition 1 (Expression for $\Pr\{\delta_j < \delta_i\}$*):* Let $\{\lambda_l\}_{l=1}^{L}$ be the distinct nonzero eigenvalues of $\boldsymbol{K}_{\boldsymbol{xx}|\boldsymbol{S}_i}\boldsymbol{F}_{ij}$ with multiplicities $\{\mu_l\}_{l=1}^{L}$, and let $\{\lambda_l\}_{l=1}^{K}$ be negative and $\{\lambda_l\}_{l=K+1}^{L}$ positive, respectively. Then

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

for $c_{ij} \le 0$ and

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

for $c_{ij} > 0$.

[2]Note, however, that $\Pr\{\delta_j < \delta_i\}$ is, in general, i.e., for $S > 2$, not the probability that Φ decides in favor of symbol $\boldsymbol{S}_j$ when $\boldsymbol{S}_i$ is transmitted.

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

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

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

Since the residue at $s = 0$ is unity, the last probability may also be written by including the pole at the origin in the sum over the residues and omitting the addend unity.

Note that the above expression of the pairwise error probability is not very well-suited for numerical evaluation, because the sum of the residues tends to be numerically unstable for high multiplicities of eigenvalues. When the latter is true, a saddle-point integration technique (cf., for example, [26]) for the inverse Laplace transform of (A6) can be used for numerical calculations.

B. Asymptotic Pairwise Error Probability

The expression for the pairwise error probability $\Pr\{\delta_j < \delta_i\}$ is not insightful in terms of its dependence on system parameters like the signals employed or the fading correlation. For finite SNR, an analysis of these dependencies seems to be analytically intractable. Therefore, the dependencies in the limit as SNR $\to \infty$ are of special interest. Since $\Pr\{\delta_j < \delta_i\}$ depends on c_{ij} and the eigenvalues of $\boldsymbol{K}_{\boldsymbol{xx}|\boldsymbol{S}_i} \boldsymbol{F}_{ij}$, the asymptotic expressions for these quantities need to be obtained. Interestingly, in [16], [13], [14], and in the following sections on coherent and noncoherent space–time receivers, we find a similarity in the behavior of the eigenvalues of $\boldsymbol{K}_{\boldsymbol{xx}|\boldsymbol{S}_i} \boldsymbol{F}_{ij}$ in the limit as SNR $\to \infty$: half of the nonzero asymptotic eigenvalues are equal to minus unity, and the other half are positive and linearly proportional to the average SNR $\overline{\gamma}$.[3] It turns out that such a structure in the eigenvalues is sufficient to ensure that the order of diversity of the receiver Φ equals the number of asymptotic positive (negative) eigenvalues. After obtaining the asymptotic expressions for the eigenvalues, we address, in Proposition 2, the question as to how the pairwise error probability behaves. While [16] contains similar results for the asymptotic analysis of $\Pr\{\delta_j < \delta_i\}$, the methods used here differ entirely and the results are more compact and more easily accessed. The derivation of the next proposition is given in Appendix B.

Proposition 2 (Asymptotics of the Pairwise Symbol Error Probability): Assume that $\hat{c}_{ij} = \lim_{\overline{\gamma} \to \infty} c_{ij}$ exists and that the (asymptotic) $2K\mu$ ($\mu \in \mathbb{N}$) nonzero eigenvalues of $\boldsymbol{K}_{\boldsymbol{xx}|\boldsymbol{S}_i} \boldsymbol{F}_{ij}$ are such that

- $\mathcal{K}$ eigenvalues are of the form $\alpha_l \overline{\gamma}$ with multiplicity μ_l, where $\{\alpha_l\}_{l=1}^{\mathcal{K}}$ are distinct and positive, $\mathcal{K}$ and $\{\mu_l\}_{l=1}^{\mathcal{K}}$ such that

$$\sum_{l=1}^{\mathcal{K}} \mu_l = K\mu.$$

- -1 occurs with multiplicity $K\mu$

then

$$\lim_{\overline{\gamma} \to \infty} \overline{\gamma}^{K\mu} \Pr\{\delta_j < \delta_i\} = \frac{\mathrm{e}^{\hat{c}_{ij}}}{\prod_{l=1}^{\mathcal{K}} \alpha_l^{\mu_l}} \sum_{k=0}^{K\mu-1} \binom{2K\mu - 1 - k}{K\mu} \frac{(-\hat{c}_{ij})^k}{k!}$$

for $\hat{c}_{ij} \le 0$ and

$$\lim_{\overline{\gamma} \to \infty} \overline{\gamma}^{K\mu} \Pr\{\delta_j < \delta_i\} = \frac{1}{\prod_{l=1}^{\mathcal{K}} \alpha_l^{\mu_l}} \sum_{k=0}^{K\mu} \binom{2K\mu - 1 - k}{K\mu - 1} \frac{\hat{c}_{ij}^k}{k!}$$

for $\hat{c}_{ij} > 0$.

Note that we use the conventions $0! = 1$ and $0^0 = 1$. The proposition gives asymptotically tight approximations for the pairwise error probabilities so that resorting to a Chernoff bound becomes unnecessary. Furthermore, to evaluate these expressions, just the asymptotic eigenvalues of the given matrices have to be found.

IV. Application to Space–Time Modulation

Applying the analytical tools developed in the previous section, we calculate asymptotic pairwise error probabilities for space–time modulation. While the final results show some similarities to those obtained previously, the unifying analysis includes correlation in the fading processes, offers a common approach to all of the problems, and yields asymptotically tight expressions for the pairwise error probabilities. The latter can subsequently be combined with the asymptotic union bound, yielding an accessible and tight performance criterion for signal constellation design.

A. Coherent Detection

For coherent detection, perfect knowledge of the fading coefficients is assumed. (However, it is conceptually not hard to account for imperfect estimates in our framework as long as the difference between the true fading coefficients and their estimates is $\mathcal{CN}(\boldsymbol{0})$ distributed with known covariance.) The likelihood of the sufficient statistics $\boldsymbol{y}$ given in (3) conditioned on the transmitted signal $\boldsymbol{S}_i$ and the fading coefficients $\boldsymbol{h}$ is then given by

$$p(\boldsymbol{y}|\boldsymbol{S}_i,\, \boldsymbol{h}) = \frac{1}{\pi^{DN}|\boldsymbol{I}_{DN}|} \mathrm{e}^{-\|\boldsymbol{y} - \sqrt{\overline{\gamma}}\boldsymbol{\mathcal{S}}_i \boldsymbol{h}\|^2}. \tag{10}$$

Defining the new $(M+D)N$-dimensional decision statistic

$$\boldsymbol{z} = \begin{bmatrix} \sqrt{\overline{\gamma}}\boldsymbol{h} \\ \boldsymbol{y} \end{bmatrix} \tag{11}$$

and the matrix

$$\boldsymbol{F}_i^{\mathrm{C}} = \begin{bmatrix} \boldsymbol{\mathcal{S}}_i^\dagger \boldsymbol{\mathcal{S}}_i & -\boldsymbol{\mathcal{S}}_i^\dagger \\ -\boldsymbol{\mathcal{S}}_i & \boldsymbol{0}_{DN} \end{bmatrix} \tag{12}$$

the optimum ML receiver Φ^{C} for coherent detection is then easily seen to be

$$\Phi^{\mathrm{C}}\colon \hat{i} = \arg\min_{1 \le i \le S} \boldsymbol{z}^\dagger \boldsymbol{F}_i^{\mathrm{C}} \boldsymbol{z} = \arg\min_{1 \le i \le S} \delta_i^{\mathrm{C}} \tag{13}$$

[3]Note that the receiver Φ can be scaled by a constant without changing it, leading to an equivalent scaling of the eigenvalues, so this implies a normalization of Φ.

where δ_i^{C} is implicitly defined. This representation of Φ^{C} is different from, but equivalent to, the representations in [6], [7], [9]. However, the advantage of (13) is that it falls within the general framework to which our unified approach to asymptotic pairwise symbol error probability can be applied. Note that earlier only a Chernoff bound on the asymptotic pairwise error probability was obtained [7], [9]. (After this paper was submitted for publication, we became aware of [41], which contains an asymptotically tight bound for the coherent case similar to ours.) The bound in [9] is derived under the assumption of unitary signaling but is tighter by a factor of two when compared to [7]. Our analysis is more general in that it does not require this assumption and moreover, it is asymptotically tight.

Proposition 3 (Asymptotic Pairwise Error Probability for Coherent Detection): The probability $\Pr\{\delta_j^{\mathrm{C}} < \delta_i^{\mathrm{C}}\}$ approaches asymptotically

$$\Pr{}^{\mathrm{a}}\{\delta_j^{\mathrm{C}} < \delta_i^{\mathrm{C}}\} = \frac{\overline{\gamma}^{-MN}\binom{2MN-1}{MN}}{|\boldsymbol{\Sigma}||(\boldsymbol{S}_i-\boldsymbol{S}_j)^\dagger(\boldsymbol{S}_i-\boldsymbol{S}_j)|^N},$$

if $\boldsymbol{A}_{ij} = (\boldsymbol{S}_i-\boldsymbol{S}_j)^\dagger(\boldsymbol{S}_i-\boldsymbol{S}_j)$ has full rank M. If $\boldsymbol{A}_{ij}$ does not have full rank, the diversity order gain will be smaller than MN.

Corollary 1 (Difference Between Coherent Asymptotic Error Probability and Chernoff Bound): The asymptotic difference between the coherent asymptotic pairwise error probability and its Chernoff bound given in [9] is

$$10\lg(4) - \frac{10}{MN}\lg\binom{2MN}{MN}$$

in decibels.

For $M = N = 2$, this amounts to 1.41 dB, for $M = N = 4$ to 0.53 dB, and for $MN \to \infty$ the difference goes to zero. (To prove this limit, the duplication formula for the Gamma function [27] is useful.)

Proof: Note that the decision statistics $\boldsymbol{z}$ for Φ^{C} are the fading coefficients (assumed to be perfectly known in the coherent case) and the sufficient statistics $\boldsymbol{y}$ from the received signal and are hence $\mathcal{CN}(\boldsymbol{0}_{(D+M)N\times 1}, \boldsymbol{K}_{\boldsymbol{zz}|\boldsymbol{S}_i})$ distributed where

$$\boldsymbol{K}_{\boldsymbol{zz}|\boldsymbol{S}_i} = E\left[\boldsymbol{z}\boldsymbol{z}^\dagger\right] = \begin{bmatrix} \overline{\gamma}\boldsymbol{\Sigma} & \overline{\gamma}\boldsymbol{\Sigma}\boldsymbol{\mathcal{S}}_i^\dagger \\ \overline{\gamma}\boldsymbol{\mathcal{S}}_i\boldsymbol{\Sigma} & \overline{\gamma}\boldsymbol{\mathcal{S}}_i\boldsymbol{\Sigma}\boldsymbol{\mathcal{S}}_i^\dagger + \boldsymbol{I}\end{bmatrix}. \tag{14}$$

We study below the asymptotic eigenvalues of

$$\boldsymbol{C}_{ij}^{\mathrm{C}} = \boldsymbol{K}_{\boldsymbol{zz}|\boldsymbol{S}_i}(\boldsymbol{F}_j^{\mathrm{C}} - \boldsymbol{F}_i^{\mathrm{C}})$$

and find that half of them are equal to minus unity and the other half positive and linear in the SNR. Hence the results from Section III apply ($c_{ij} = 0$ for this case).

To find the asymptotic eigenvalues of $\boldsymbol{C}_{ij}^{\mathrm{C}}$ we define the abbreviations

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{I}_{MN} \\ \boldsymbol{\mathcal{S}}_i \end{bmatrix} \tag{15}$$

$$\boldsymbol{Y} = \boldsymbol{\Sigma}(\boldsymbol{\mathcal{S}}_i - \boldsymbol{\mathcal{S}}_j)^\dagger\left[-\boldsymbol{\mathcal{S}}_j \quad \boldsymbol{I}_{DN}\right] \tag{16}$$

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{0}_{MN} & \boldsymbol{0}_{MN\times DN} \\ \boldsymbol{\mathcal{S}}_i - \boldsymbol{\mathcal{S}}_j & \boldsymbol{0}_{DN}\end{bmatrix} \tag{17}$$

and note that

$$\boldsymbol{C}_{ij}^{\mathrm{C}} = \boldsymbol{K}_{\boldsymbol{zz}|\boldsymbol{S}_i}\left(\boldsymbol{F}_j^{C} - \boldsymbol{F}_i^{C}\right) = \overline{\gamma}\boldsymbol{X}\boldsymbol{Y} + \boldsymbol{Z}. \tag{18}$$

If we assume that

$$\boldsymbol{Y}\boldsymbol{X} = \boldsymbol{\Sigma}(\boldsymbol{\mathcal{S}}_i - \boldsymbol{\mathcal{S}}_j)^\dagger(\boldsymbol{\mathcal{S}}_i - \boldsymbol{\mathcal{S}}_j)$$

has full rank, we can apply Theorem 2 of Appendix C, and find that the nonzero eigenvalues of $\boldsymbol{C}_{ij}^{\mathrm{C}}$ are asymptotically equal to the eigenvalues of $\overline{\gamma}\boldsymbol{Y}\boldsymbol{X}$ and $\boldsymbol{Z}(\boldsymbol{I}_{(M+D)N} - \boldsymbol{X}(\boldsymbol{Y}\boldsymbol{X})^{-1}\boldsymbol{Y})$. Let $\boldsymbol{A}_{ij} = (\boldsymbol{S}_i - \boldsymbol{S}_j)^\dagger(\boldsymbol{S}_i - \boldsymbol{S}_j)$, so that by inserting the definitions for $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$, one easily finds that the nonzero eigenvalues of $\boldsymbol{Z} - \boldsymbol{Z}\boldsymbol{X}(\boldsymbol{Y}\boldsymbol{X})^{-1}\boldsymbol{Y}$ are the nonzero eigenvalues of $-(\boldsymbol{S}_i - \boldsymbol{S}_j)\boldsymbol{A}_{ij}^{-1}(\boldsymbol{S}_i - \boldsymbol{S}_j)^\dagger$ with multiplicity N, and since the latter is a negative projection matrix of rank M, the determinant yields the eigenvalue minus unity with multiplicity MN.

In case $\boldsymbol{A}_{ij}$ has rank $r < M$ the matrix $\boldsymbol{Y}\boldsymbol{X}$ has rank Nr, and it is not hard to show that the number of positive eigenvalues linear in $\overline{\gamma}$ and the multiplicity of the eigenvalue minus one will be Nr, leading to a diversity order gain $Nr < MN$.

For the corollary note the identity

$$2\binom{2MN-1}{MN} = \binom{2MN}{MN} \tag{19}$$

and take the limit $\overline{\gamma} \to \infty$ in [9, eq. (20)] ($\overline{\gamma}$ here corresponds to ρT in [9]). □

To show the usefulness of the above analysis we include two simple examples.

Example 1: BPSK With N Receive Antennas: Consider binary phase-shift keying (BPSK) received by N antennas with correlated fading with correlation matrix $\boldsymbol{\Sigma}$. Obviously, $S = 2$, $M = 1$, $D = 1$, $\boldsymbol{S}_1 = +1$, $\boldsymbol{S}_2 = -1$, and we obtain for the asymptotic error probability

$$P^{\mathrm{BPSKa}} = \frac{(4\overline{\gamma})^{-N}}{|\boldsymbol{\Sigma}|}\binom{2N-1}{N}. \tag{20}$$

The special case for i.i.d. fading reduces to [21, eq. (14-4-18)].

Example 2: Alamouti Scheme With BPSK: Consider the Alamouti scheme [28] in which signal matrices of the form

$$\boldsymbol{S}_i = \begin{bmatrix} s_0 & s_1 \\ -s_1^* & s_0^* \end{bmatrix} \tag{21}$$

are transmitted over $M = 2$ antennas. The symbols s_0 and s_1 are drawn from some predetermined alphabet, like a QAM constellation. For the example, we choose the BPSK alphabet and consequently $(s_0, s_1) = 2^{-(1/2)}(\pm 1, \pm 1)$.

For i.i.d. fading ($2\boldsymbol{\Sigma} = \boldsymbol{I}_2$) and $\boldsymbol{S}_1$ corresponding to $(+1, +1)$ and $\boldsymbol{S}_2$ to $(-1, +1)$, one can find the eigenvalues of $\boldsymbol{C}_{12}^{\mathrm{C}} = \boldsymbol{K}_{\boldsymbol{zz}|\boldsymbol{S}_1}(\boldsymbol{F}_2^{\mathrm{C}} - \boldsymbol{F}_1^{\mathrm{C}})$ analytically. There are two distinct eigenvalues, $\frac{1}{2}(\overline{\gamma} \pm \sqrt{\overline{\gamma}^2 + 4\overline{\gamma}})$, each of multiplicity two. Fig. 1 shows the behavior of the distinct eigenvalues for increasing SNR. As expected from the preceding analysis, the positive eigenvalue increases linearly with the SNR and the negative eigenvalue approaches -1.

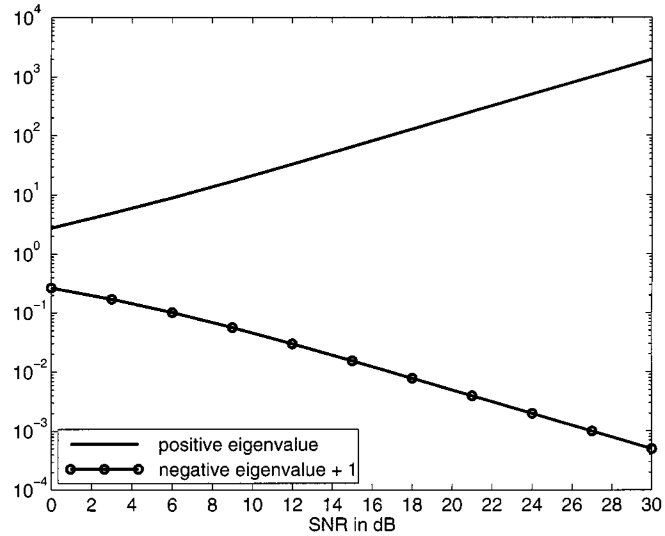

Fig. 1. The two distinct eigenvalues of $\boldsymbol{C}_{12}^{\mathrm{C}} = \boldsymbol{K}_{\boldsymbol{zz}|\boldsymbol{S}_1}(\boldsymbol{F}_2^{\mathrm{C}} - \boldsymbol{F}_1^{\mathrm{C}})$.

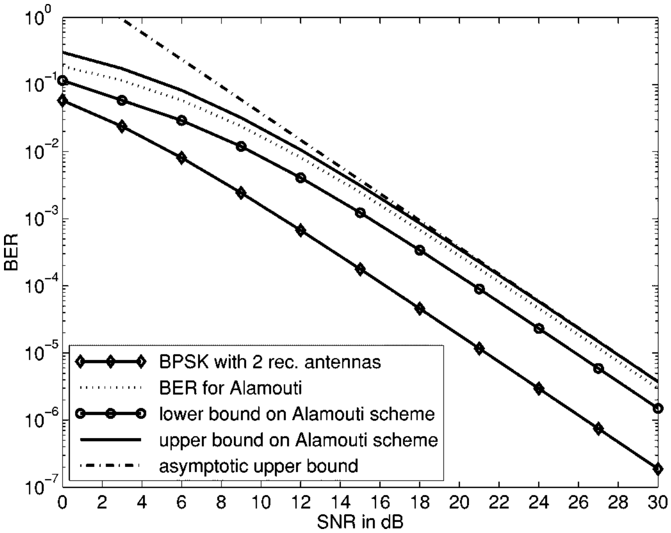

Fig. 2. Comparison of the Alamouti scheme using BPSK symbols for $M = 2$ transmit and $N = 1$ receive antenna with conventional BPSK detected by two antennas.

Due to the structure in the Alamouti scheme, it is easy for this case to obtain an asymptotic lower and upper bound on the bit error rate (BER) (in fact, for BPSK or QPSK and i.i.d. fading, the exact BER can be obtained [42]). We assume i.i.d. fading, and the bounds become

$$P_{\mathrm{LB}}^{\mathrm{Ala\text{-}a}} = \frac{1}{2}\,\overline{\gamma}^{-2N}\binom{4N-1}{2N} \tag{22}$$

$$P_{\mathrm{UB}}^{\mathrm{Ala\text{-}a}} = \left(\overline{\gamma}_b^{-2N} + (2\overline{\gamma})^{-2N}\right)\binom{4N-1}{2N}. \tag{23}$$

Fig. 2 compares the Alamouti scheme for one receive antenna with the usual BPSK detected by two receive antennas in i.i.d. fading. From the asymptotic bounds, we conclude that, with the normalization that both systems radiate the same power per *supersymbol*, the case of receiver diversity has an asymptotic advantage of at least 4.5 dB (and at most 6.5 dB) over the transmitter diversity scheme. The gap can be shown to be exactly 6 dB using [42] (thus, the upper bound is 0.5 dB loose due to use of the union bound). To make the comparison fair, however, we switch to SNR per bit ($\overline{\gamma}_b = \overline{\gamma}/2$ for the Alamouti scheme) and, consequently, the analytic gap is 3 dB in terms of SNR per bit. Note that for general space–time modulation schemes usually no analytic expressions for bit per symbol error rates are available. These general schemes can be analyzed using the presented asymptotically tight pairwise error probabilities.

B. Noncoherent Detection

For the noncoherent detection problem the knowledge of the fading coefficients is not available at the receiver. However, their statistics, the fading correlation matrix $\boldsymbol{\Sigma}$, and the average SNR $\overline{\gamma}$ are in general required by the noncoherent ML receiver $\Phi^{\mathrm{NC\text{-}ML}}$. If one wants to avoid the estimation of these statistics as well, a GLRT receiver can be employed, which turns out to be equivalent to the ML receiver for unitary signaling and i.i.d. fading.

We summarize the main results of this section in the following paragraphs.

Rather than assuming a specific signal structure, the error probability analyses of the two receivers are performed without any assumptions on the structure of the signals except for a rank requirement to guarantee full order of diversity. It is then shown that unitary signaling is optimal in the sense of minimizing the asymptotic union bound of the ML receiver for equal-energy signals.

The optimality of an equal-energy signal set in the sense of maximizing capacity was shown for i.i.d. fading for the case $M = N = 1, \overline{\gamma} \to \infty$, and the case $M = N = 1$, $D \to \infty$ in [9]; it was argued in [9] that equal-energy (across signals and transmit antennas) signaling is also optimal for more than one transmit and receive antennas. Consequently, for more than one transmit antenna, unitary signaling is argued to be optimal in the capacity sense, because the capacity-maximizing signal distribution is shown to be an isotropically distributed $D \times M$ matrix with orthonormal columns multiplied by an $M \times M$ diagonal matrix containing the powers for each antenna. The conjecture that equal-energy signaling maximizes capacity for high SNRs is proved in [29], [4]. Interestingly, therefore, unitary signaling is optimal in both the information-theoretic sense as well as the symbol error probability sense.

Furthermore, it is found that the number of dimensions has to be necessarily at least twice as large as the number of transmit antennas ($D \geq 2M$) to achieve full order of diversity MN. This requirement is also reported in [11] for unitary signaling by exploiting the observation of [9] that unity may not be a singular value of any pair $\boldsymbol{S}_i^{\dagger}\boldsymbol{S}_j$ to achieve full order of diversity; for the latter to hold [11] shows that the columns of $\boldsymbol{S}_i$ and $\boldsymbol{S}_j$ have to be linearly independent of each other and hence $D \geq 2M$. But since the Chernoff-based analysis in [9] is not asymptotically tight, Hughes [11] cannot conclude that $D \geq 2M$ is *necessary* to achieve full order of diversity. In [29] and [4], it is found that the capacity of the channel cannot be increased by increasing the number of transmit antennas if $D < 2M$. (Note that in [9], Footnote 1 should read $T \geq 2M$.)

The error probability analysis for the GLRT receiver may also be interpreted as an analysis of the ML receiver derived for unitary signaling under the i.i.d. fading assumption when the fading experienced is actually correlated. While in the case of general signals the GLRT performs asymptotically worse than the ML receiver, for unitary signals it achieves asymptotically the same symbol error probability as the ML receiver. Consequently, unitary signals are also optimum for the GLRT receiver.

1) ML Receiver: Noting that the sufficient statistics $\boldsymbol{y}$ are zero-mean $\mathcal{CN}$-distributed with covariance

$$\boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{\mathcal{S}}_i} = \overline{\gamma}\boldsymbol{\mathcal{S}}_i\boldsymbol{\Sigma}\boldsymbol{\mathcal{S}}_i^\dagger + \boldsymbol{I}_{DN} \tag{24}$$

and defining $\boldsymbol{\mathcal{R}}_{ij} = \boldsymbol{\mathcal{S}}_i^\dagger\boldsymbol{\mathcal{S}}_j$ (correspondingly, $\boldsymbol{R}_{ij} = \boldsymbol{S}_i^\dagger\boldsymbol{S}_j$, $\boldsymbol{\mathcal{R}}_{ij} = \boldsymbol{I}_N \otimes \boldsymbol{R}_{ij}$)

$$\boldsymbol{F}_i^{\rm NC} = -\boldsymbol{\mathcal{S}}_i(\boldsymbol{\mathcal{R}}_{ii} + \overline{\gamma}^{-1}\boldsymbol{\Sigma}^{-1})^{-1}\boldsymbol{\mathcal{S}}_i^\dagger$$

and $c_i = \ln|\overline{\gamma}\boldsymbol{\mathcal{R}}_{ii} + \boldsymbol{\Sigma}^{-1}|$ leads to the optimum noncoherent ML receiver

$$\Phi^{\text{NC-ML}}\colon \hat{i} = \arg\min_{1\le i\le S} \boldsymbol{y}^\dagger\boldsymbol{F}_i^{\rm NC}\boldsymbol{y} + c_i \tag{25}$$

$$= \arg\min_{1\le i\le S} \delta_i^{\text{NC-ML}} \tag{26}$$

where a formula for the matrix inverse of a small-rank adjustment ([30, Sec. 0.7.4], also known as Woodbury's identity) and a determinantal formula [30, Sec. 0.8.5] were applied. In the case of uncorrelated fading and unitary signals, $\Phi^{\text{NC-ML}}$ simplifies to the expressions presented in [9].

The unified analysis tools of the previous section allow for a rather straightforward performance evaluation.

Proposition 4 (Asymptotic Pairwise Error Probability and Union Bound for Noncoherent ML Detection): Let

$$\hat{c}_{ij} = N\ln\frac{|\boldsymbol{R}_{ii}|}{|\boldsymbol{R}_{jj}|}.$$

The probability $\Pr\{\delta_j^{\text{NC-ML}} < \delta_i^{\text{NC-ML}}\}$ approaches asymptotically

$$\Pr{}^{\rm a}\left\{\delta_j^{\text{NC-ML}} < \delta_i^{\text{NC-ML}}\right\} = \frac{\overline{\gamma}^{-MN} e^{\hat{c}_{ij}} \sum_{k=0}^{MN-1} \binom{2MN-1-k}{MN} \frac{(-\hat{c}_{ij})^k}{k!}}{|\boldsymbol{\Sigma}|\left|\boldsymbol{R}_{ii} - \boldsymbol{R}_{ij}\boldsymbol{R}_{jj}^{-1}\boldsymbol{R}_{ji}\right|^N}$$

for $\hat{c}_{ij} \le 0$ and

$$\Pr{}^{\rm a}\left\{\delta_j^{\text{NC-ML}} < \delta_i^{\text{NC-ML}}\right\} = \frac{\overline{\gamma}^{-MN} \sum_{k=0}^{MN} \binom{2MN-1-k}{MN-1} \frac{\hat{c}_{ij}^k}{k!}}{|\boldsymbol{\Sigma}|\left|\boldsymbol{R}_{ii} - \boldsymbol{R}_{ij}\boldsymbol{R}_{jj}^{-1}\boldsymbol{R}_{ji}\right|^N}$$

for $\hat{c}_{ij} > 0$, where we assumed that the matrix

$$\begin{bmatrix} \boldsymbol{S}_i^\dagger \\ \boldsymbol{S}_j^\dagger \end{bmatrix} [\,\boldsymbol{S}_i \quad \boldsymbol{S}_j\,] = \begin{bmatrix} \boldsymbol{R}_{ii} & \boldsymbol{R}_{ij} \\ \boldsymbol{R}_{ji} & \boldsymbol{R}_{jj} \end{bmatrix} \tag{27}$$

has full rank to achieve the full order of diversity MN, and hence $D \ge 2M$ is necessary. The sum of the two probabilities

$$P_{ij}^{\text{NC-MPa}} = \Pr{}^{\rm a}\left\{\delta_j^{\text{NC-ML}} < \delta_i^{\text{NC-ML}}\right\} + \Pr{}^{\rm a}\left\{\delta_i^{\text{NC-ML}} < \delta_j^{\text{NC-ML}}\right\}$$

can be written as

$$P_{ij}^{\text{NC-MLa}} = \frac{\overline{\gamma}^{-MN} \sum_{k=0}^{MN} \binom{2MN-k}{MN} \frac{\hat{c}_{ij}^k}{k!}}{|\boldsymbol{\Sigma}|\left|\boldsymbol{R}_{ii} - \boldsymbol{R}_{ij}\boldsymbol{R}_{jj}^{-1}\boldsymbol{R}_{ji}\right|^N} \tag{28}$$

for $\hat{c}_{ij} \ge 0$; the signals can always be renumbered to ensure this condition. By this, the asymptotic union bound on symbol error probability becomes

$$P_{\rm UB}^{\text{NC-MLa}} = S^{-1}\sum_{i=1}^{S-1}\sum_{j=i+1}^{S} P_{ij}^{\text{NC-MLa}} \tag{29}$$

for $|\boldsymbol{R}_{11}| \ge |\boldsymbol{R}_{22}| \ge \cdots \ge |\boldsymbol{R}_{SS}|$.

Corollary 2 (Difference Between Noncoherent Asymptotic Error Probability and Asymptotic Chernoff Bound): The difference between the asymptotic noncoherent pairwise error probability and the asymptotic Chernoff bound derived from [9] is

$$10\lg(4) - \frac{10}{MN}\lg\binom{2MN}{MN}$$

in decibels.

This is the same difference as in the coherent case (cf. Corollary 1). Note, however, that the Chernoff bound analysis in [9] depends critically on the assumption of unitary signaling. Our analysis does not require this assumption.

Proof: We have to show that the asymptotic nonzero eigenvalues of

$$\boldsymbol{C}_{ij} = \boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{\mathcal{S}}_i}(\boldsymbol{F}_j^{\rm NC} - \boldsymbol{F}_i^{\rm NC})$$

are the MN positive eigenvalues of $\overline{\gamma}\boldsymbol{\Sigma}(\boldsymbol{\mathcal{R}}_{ii} - \boldsymbol{\mathcal{R}}_{ij}\boldsymbol{\mathcal{R}}_{jj}^{-1}\boldsymbol{\mathcal{R}}_{ji})$ and minus unity with multiplicity MN. To show that $\lim_{\overline{\gamma}\to\infty} c_{ij} = \hat{c}_{ij}$ is left as an exercise.

Defining $\boldsymbol{\mathcal{S}}_{ij} = [\boldsymbol{\mathcal{S}}_i\boldsymbol{\mathcal{S}}_j]$ one finds $\boldsymbol{C}_{ij}$ given at the bottom of this page. Since the nonzero eigenvalues of $\boldsymbol{AB}$ and $\boldsymbol{BA}$ are equal [30, Theorem 1.3.20], we can equivalently find the nonzero eigenvalues of the matrix product with $\boldsymbol{\mathcal{S}}_{ij}$ to the right instead of to the left. Multiplying all matrices, one obtains $\boldsymbol{M}_{ij}$ as given at the top of the following page. Using Woodbury's identity, the upper right block of $\boldsymbol{M}_{ij}$ can be written as

$$[\boldsymbol{M}_{ij}]_{\rm ur} = \boldsymbol{\Sigma}\boldsymbol{\mathcal{R}}_{ij}\boldsymbol{\mathcal{R}}_{jj}\left(\boldsymbol{\Sigma} + \overline{\gamma}^{-1}\boldsymbol{\mathcal{R}}_{jj}^{-1}\right)^{-1}. \tag{30}$$

Neglecting terms in $\overline{\gamma}^{-1}$, and applying the results of Theorem 2 in Appendix C for the approximation of the eigenvalues of a block partitioned matrix with a diagonal block linear in $\overline{\gamma}$, we

$$\boldsymbol{C}_{ij} = \boldsymbol{\mathcal{S}}_{ij}\begin{bmatrix} (\overline{\gamma}\boldsymbol{\Sigma}\boldsymbol{\mathcal{R}}_{ii} + \boldsymbol{I}_{MN})\left(\boldsymbol{\mathcal{R}}_{ii} + \overline{\gamma}^{-1}\boldsymbol{\Sigma}^{-1}\right)^{-1} & -\overline{\gamma}\boldsymbol{\Sigma}\boldsymbol{\mathcal{R}}_{ij}\left(\boldsymbol{\mathcal{R}}_{jj} + \overline{\gamma}^{-1}\boldsymbol{\Sigma}^{-1}\right)^{-1} \\ \boldsymbol{0}_{MN} & -\left(\boldsymbol{\mathcal{R}}_{jj} - \overline{\gamma}^{-1}\boldsymbol{\Sigma}^{-1}\right)^{-1} \end{bmatrix}\boldsymbol{\mathcal{S}}_{ij}^\dagger.$$

$$\boldsymbol{M}_{ij} = \begin{bmatrix} \overline{\gamma}\boldsymbol{\Sigma}\boldsymbol{\mathcal{R}}_{ii} - \overline{\gamma}\boldsymbol{\Sigma}\boldsymbol{\mathcal{R}}_{ij}\left(\boldsymbol{\mathcal{R}}_{jj} + \overline{\gamma}^{-1}\boldsymbol{\Sigma}^{-1}\right)^{-1}\boldsymbol{\mathcal{R}}_{ji} & \overline{\gamma}\boldsymbol{\Sigma}\boldsymbol{\mathcal{R}}_{ij}\left(\boldsymbol{I}_{MN} - \left(\boldsymbol{\mathcal{R}}_{jj} + \overline{\gamma}^{-1}\boldsymbol{\Sigma}^{-1}\right)^{-1}\boldsymbol{\mathcal{R}}_{ji}\right) \\ -\left(\boldsymbol{\mathcal{R}}_{jj} - \overline{\gamma}^{-1}\boldsymbol{\Sigma}^{-1}\right)^{-1}\boldsymbol{\mathcal{R}}_{ji} & -\left(\boldsymbol{\mathcal{R}}_{jj} - \overline{\gamma}^{-1}\boldsymbol{\Sigma}^{-1}\right)^{-1}\boldsymbol{\mathcal{R}}_{jj} \end{bmatrix}.$$

find the positive and negative eigenvalues as stated above (note $|\boldsymbol{\mathcal{R}}_{ii}| = |\boldsymbol{R}_{ii}|^N$, similarly for the Schur complement). The expression for $P_{ij}^{\text{NC-MLa}}$ arises by inserting $\hat{c}_{ij} = -\hat{c}_{ji}$, $e^{\hat{c}_{ij}} = \frac{|\boldsymbol{R}_{ii}|^N}{|\boldsymbol{R}_{jj}|^N}$, and using a determinantal formula [30, Sec. 0.8.5].

The corollary is found similarly as in the coherent case, using [9, eq. (18)] and unitary signals, as assumed in that reference.□

Having obtained the pairwise symbol error probability, we now derive a result on the structure of the signals that will minimize the asymptotic union bound on the error probability. So far, we have made no assumptions concerning the structure or energy distribution of the signals, so the above analysis is completely general. However, in the following proposition, the signals are assumed to have equal energies.

Proposition 5 (Structure of the Equal-Energy Signals That Minimize $P_{\text{UB}}^{\text{NC-MLa}}$): If the energies of all signals are equal, i.e.,

$$E_l = \text{tr}(\boldsymbol{R}_{ll}) = M, \qquad \text{for all } l \in \{1, 2, \ldots, S\}$$

the signals $\{\tilde{\boldsymbol{S}}_l\}_{l=1}^S \subset \mathbb{C}^{D\times M}$ that minimize the asymptotic union bound $P_{\text{UB}}^{\text{NC-MLa}}$ have orthonormal columns, i.e., $\tilde{\boldsymbol{S}}_l^\dagger \tilde{\boldsymbol{S}}_l = \boldsymbol{I}_M$ for all l.

Proof: We will show that given any signal set $\{\boldsymbol{S}_l\}_{l=1}^S$ with equal energies, there is a set $\{\tilde{\boldsymbol{S}}_l\}_{l=1}^S$ with orthonormal columns that has lower pair probabilities $\tilde{P}_{ij}^{\text{NC-MLa}}$ for all pairs (i, j) $i \neq j$. Hence the signal set that minimizes the asymptotic union bound must have signals with orthonormal columns.

Let $\{\boldsymbol{S}_l\}_{l=1}^S$ be given. Consider the eigendecomposition

$$\boldsymbol{R}_{ll} = \boldsymbol{S}_l^\dagger \boldsymbol{S}_l = \boldsymbol{\Psi}_l \boldsymbol{\Lambda}_l \boldsymbol{\Psi}_l^\dagger$$

and define the signals of the new signal set $\{\tilde{\boldsymbol{S}}_l\}_{l=1}^S$ by

$$\tilde{\boldsymbol{S}}_l = \boldsymbol{S}_l \boldsymbol{\Psi}_l \boldsymbol{\Lambda}_l^{-(1/2)}. \tag{31}$$

It is easily verified that $\text{tr}(\tilde{\boldsymbol{S}}_l^\dagger \tilde{\boldsymbol{S}}_l) = \text{tr}(\boldsymbol{I}_M) = M$, and so the energies of the signals are not changed. After inserting $\tilde{\boldsymbol{S}}_i$ and $\tilde{\boldsymbol{S}}_j$ into (28) and some algebra, the pair probability $\tilde{P}_{ij}^{\text{NC-MLa}}$ becomes

$$\tilde{P}_{ij}^{\text{NC-MLa}} = \frac{\overline{\gamma}^{-MN}|\boldsymbol{\Lambda}_i|^N}{|\boldsymbol{\Sigma}||\boldsymbol{R}_{ii} - \boldsymbol{R}_{ij}\boldsymbol{R}_{jj}^{-1}\boldsymbol{R}_{ji}|^N}\binom{2MN}{MN}. \tag{32}$$

Since the arithmetic mean of the eigenvalues of $\boldsymbol{R}_{ii}$ is unity it follows by the arithmetic–geometric mean inequality that $|\boldsymbol{\Lambda}_i|$ is equal or less than unity. Furthermore, $\tilde{\hat{c}}_{ij} = 1$ and, consequently, the sum over nonnegative terms in (28) reduces to the first term (recall our convention $(\ln 1)^0 = 0^0 = 1$). Consequently, the pair probability between any pair of the new signals is equal to or less than the pair probability of the same pair of old signals. □

In [31], the results on the structure of the signals and the asymptotic union bound are used to propose a signal design algorithm for arbitrary numbers of antennas, dimensions, and signals. Since several numerical examples are given in [31], we only explore the simplest possible example here.

Example 3: BFSK with N Receive Antennas: Consider the usual binary frequency-shift keying (BFSK) setup, i.e., $M = 1$, $D = 2$, $\boldsymbol{S}_1 = [1\ 0]^\top$, and $\boldsymbol{S}_2 = [0\ 1]^\top$. The asymptotic symbol per bit error probability for noncoherent ML detection by N receive antennas with fading correlation matrix $\boldsymbol{\Sigma}$ then becomes

$$P^{\text{BFSKa}} = \frac{\overline{\gamma}^{-N}}{|\boldsymbol{\Sigma}|}\binom{2N-1}{N} \tag{33}$$

displaying the expected 6-dB gap to coherent modulation. The special case for i.i.d. fading reduces to [21, eq. (14-4-33)].

2) GLRT Receiver: Maximizing the likelihood of the received signal conditioned on the fading coefficients $\boldsymbol{h}$ and the transmitted signal $\boldsymbol{S}_i$ (cf. (10)) over the fading coefficients and subsequently over the transmitted signal yields the GLRT receiver as

$$\begin{aligned} \Phi^{\text{NC-G}}\colon \hat{i} &= \arg\min_{1\le i\le S} -\boldsymbol{y}^\dagger \boldsymbol{\mathcal{S}}_i \left(\boldsymbol{\mathcal{S}}_i^\dagger \boldsymbol{\mathcal{S}}_i\right)^{-1} \boldsymbol{\mathcal{S}}_i^\dagger \boldsymbol{y} \\ &= \arg\min_{1\le i\le S} -\sum_{n=1}^N \boldsymbol{y}_n^\dagger \boldsymbol{S}_i \left(\boldsymbol{S}_i^\dagger \boldsymbol{S}_i\right)^{-1} \boldsymbol{S}_i^\dagger \boldsymbol{y}_n \\ &= \arg\min_{1\le i\le S} \delta_i^{\text{NC-G}} \end{aligned} \tag{34}$$

which is equivalent to the ML receiver in i.i.d. fading and unitary signaling (cf. (25) and [9]). Since $\boldsymbol{\mathcal{S}}_i(\boldsymbol{\mathcal{S}}_i^\dagger \boldsymbol{\mathcal{S}}_i)^{-1}\boldsymbol{\mathcal{S}}_i$ is the projection matrix onto the $\boldsymbol{\mathcal{S}}_i$th signal's subspace, the detector compares the energy of the sufficient statistics in the signals' subspaces. The independence of the receiver on any kind of fading information makes it an interesting candidate for correlated fading when one does not want to estimate the correlation matrix and for nonstrictly Rayleigh fading channels.

Proposition 6 (Asymptotic Pairwise Error Probability for Noncoherent GLRT Detection): The probability $\Pr\{\delta_j^{\text{NC-G}} < \delta_i^{\text{NC-G}}\}$ approaches asymptotically

$$\Pr{}^{\text{a}}\left\{\delta_j^{\text{NC-G}} < \delta_i^{\text{NC-G}}\right\} = \frac{\overline{\gamma}^{-MN}\binom{2MN-1}{MN}}{|\boldsymbol{\Sigma}||\boldsymbol{R}_{ii} - \boldsymbol{R}_{ij}\boldsymbol{R}_{jj}^{-1}\boldsymbol{R}_{ji}|^N}$$

where we assumed that the matrix

$$\begin{bmatrix} \boldsymbol{S}_i^\dagger \\ \boldsymbol{S}_j^\dagger \end{bmatrix} [\boldsymbol{S}_i \quad \boldsymbol{S}_j] = \begin{bmatrix} \boldsymbol{R}_{ii} & \boldsymbol{R}_{ij} \\ \boldsymbol{R}_{ji} & \boldsymbol{R}_{jj} \end{bmatrix} \tag{35}$$

has full rank to achieve the full order of diversity MN and hence $D \ge 2M$ is necessary. The sum of the two probabilities $P_{ij}^{\text{NC-Ga}} = \Pr^{\text{a}}\left\{\delta_j^{\text{NC-G}} < \delta_i^{\text{NC-G}}\right\} + \Pr^{\text{a}}\left\{\delta_i^{\text{NC-G}} < \delta_j^{\text{NC-G}}\right\}$ can be written as

$$P_{ij}^{\text{NC-Ga}} = \frac{\overline{\gamma}^{-MN}\binom{2MN-1}{MN}\left(1 + \frac{|\boldsymbol{R}_{ii}|^N}{|\boldsymbol{R}_{jj}|^N}\right)}{|\boldsymbol{\Sigma}||\boldsymbol{R}_{ii} - \boldsymbol{R}_{ij}\boldsymbol{R}_{jj}^{-1}\boldsymbol{R}_{ji}|^N}. \tag{36}$$

Corollary 3: For unitary signaling the GLRT receiver achieves asymptotically the same pair probability as the ML detector, i.e.,

$$P_{ij}^{\text{NC-MLa}} = P_{ij}^{\text{NC-Ga}}.$$

Moreover, the asymptotic symbol error rate of the GLRT and ML receiver coincide for unitary signaling.

Proof: The analysis follows closely the one for the ML receiver and is even simplified by the fact that $\hat{c}_{ij} = 0$. By identical arguments to the ML case, one has to find the eigenvalues of the matrix $\boldsymbol{M}_{ij}$ as given in the first equation at the bottom of this page. By applying the results of Theorem 1 of Appendix C, the eigenvalues of $\boldsymbol{M}_{ij}$ are seen to be asymptotically the eigenvalues of $\overline{\gamma}\boldsymbol{\Sigma}\boldsymbol{\mathcal{R}}_{ii} - \overline{\gamma}\boldsymbol{\Sigma}\boldsymbol{\mathcal{R}}_{ij}\boldsymbol{\mathcal{R}}_{jj}^{-1}\boldsymbol{\mathcal{R}}_{ji}$ and $-\boldsymbol{I}_{MN}$. The proposition follows by inserting the eigenvalues into the general expressions for the asymptotic pairwise error probability (cf. Proposition 2) and applying a determinantal formula for $\Pr\{\delta_i^{\text{NC-G}} < \delta_j^{\text{NC-G}}\}$. The first part of the corollary is obvious by the identity of (19). Since the ML receiver (25) converges for $\overline{\gamma} \to \infty$ to the GLRT receiver (34) for unitary signaling, the asymptotic symbol error rates of the two receivers also coincide. Consequently, the GLRT can also be interpreted as the asymptotic expansion of the ML receiver and is thus asymptotically optimum for unitary signaling and correlated fading. ☐

C. Differentially Coherent Detection

Coherent detection requires the knowledge of the complex fading coefficients. In practice, these would have to be estimated, requiring MN estimates for a system employing M transmit and N receive antennas. This adds considerable complexity to the receiver and potential overhead for training sequences. Noncoherent modulation as presented in the previous section avoids these complications but pays a rather high price in terms of capacity, especially when the coherence time of the channel is short [3], [4]. Moreover, in the latter situation of short coherence time, the estimation of the fading coefficients for coherent detection also becomes even more difficult. In single transmit antenna situations, differential modulation and detection is known to offer a resort to this dilemma [32].

A differential detection scheme for two transmit antennas was introduced in [33]. The codes developed from orthogonal designs in (cf. [34]), originally intended for coherent signaling, were also shown to be adaptable for differential modulation [12]. Another generalization of differential modulation and detection to several transmit antennas is proposed in [10] and [11]. In these papers, the information carrying signals are constrained to be unitary matrices. Optimal constellations under this paradigm with group constraints are presented in [35]. A general theory of group designs suitable for coherent and differentially coherent modulation is developed in [36].

In this work, we consider the differential modulation scheme of [10] and [11]. By applying the analytic tools developed above, the assumption of i.i.d. fading processes can be relaxed and an analysis for both, slow and fast fading is performed. In the slowly fading case, the fading coefficients are assumed to be constant over the duration of two consecutive symbol intervals, whereas in the fast fading case, the fading coefficients vary from one symbol to the next (but still remain constant within a symbol interval).

For this section, we reintroduce a time index; since we consider differential modulation/detection it will be convenient to denote the $D \times M$ signal matrix sent from the M transmit antennas in the current symbol interval by $\boldsymbol{S}(0)$, while the one transmitted in the previous symbol interval is denoted by $\boldsymbol{S}(-1)$. Both [10] and [11] generalize differential modulation such that the information for the current symbol interval is encoded into one out of S unitary matrices $\{\boldsymbol{V}_l\}_{l=1}^{S}$ of size $D \times D$. The signal sent in the previous symbol interval, $\boldsymbol{S}(-1)$, is multiplied with the information-carrying unitary matrix and transmitted, i.e.,

$$\boldsymbol{S}(0) = \boldsymbol{V}_l \boldsymbol{S}(-1). \tag{37}$$

References [10], [11], and [35] design constellations with (cyclic) group properties. For the purpose of this paper, the specific constellation employed is not relevant; we do not impose the restrictions that the signal matrices $\boldsymbol{S}(k)$ have orthonormal columns and/or $D = M$ (cf. [37]), but require only that $D \le M$.

At the receiver, the sufficient statistics for the previous and current symbol interval become

$$\boldsymbol{y}(-1) = \sqrt{\overline{\gamma}}\,\boldsymbol{\mathcal{S}}(-1)\boldsymbol{h}(-1) + \boldsymbol{\eta}(-1) \tag{38}$$

$$\boldsymbol{y}(0) = \sqrt{\overline{\gamma}}\,\boldsymbol{\mathcal{S}}(0)\boldsymbol{h}(0) + \boldsymbol{\eta}(0) \tag{39}$$

where, similarly as above, $\boldsymbol{\mathcal{S}}(0) = \boldsymbol{I}_N \otimes \boldsymbol{S}(0) = \boldsymbol{\mathcal{V}}_l \boldsymbol{\mathcal{S}}(-1)$ and $\boldsymbol{\mathcal{V}}_l = \boldsymbol{I}_N \otimes \boldsymbol{V}_l$. Letting $\boldsymbol{y} = [\boldsymbol{y}(-1)^\top \ \boldsymbol{y}(0)^\top]^\top$ the matrix correlation of the sufficient statistics $\boldsymbol{y}$ conditioned on symbol $\boldsymbol{V}_i$ being transmitted is $\boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{V}_i}$ and is given in the second equation at the bottom of the page. where we assumed wide-sense stationarity of the fading processes and consequently

$$\boldsymbol{\Sigma}(0) = E[\boldsymbol{h}(0)\boldsymbol{h}^\dagger(0)] = E[\boldsymbol{h}(-1)\boldsymbol{h}^\dagger(-1)]$$

and

$$\boldsymbol{\Sigma}(1) = E[\boldsymbol{h}(-1)\boldsymbol{h}^\dagger(0)].$$

To simplify notation we introduce the abbreviations $\boldsymbol{\mathcal{S}}(-1)\boldsymbol{\Sigma}(0)\boldsymbol{\mathcal{S}}^\dagger(-1) = \boldsymbol{\Omega}(0)$ and $\boldsymbol{\mathcal{S}}(-1)\boldsymbol{\Sigma}(1)\boldsymbol{\mathcal{S}}^\dagger(-1) = \boldsymbol{\Omega}(1)$. Note that the correlation matrix $\boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{V}_i}$ does not only in general depend on the transmitted information (the unitary matrix $\boldsymbol{V}_i$), but also on the previously transmitted signal $\boldsymbol{S}(-1)$. Thus, the ML, differentially coherent receiver $\Phi^{\text{DC-ML}}$ needs to

$$\boldsymbol{M}_{ij} = \begin{bmatrix} \overline{\gamma}\boldsymbol{\Sigma}\boldsymbol{\mathcal{R}}_{ii} + \boldsymbol{I}_{MN} - \overline{\gamma}\boldsymbol{\Sigma}\boldsymbol{\mathcal{R}}_{ij}\boldsymbol{\mathcal{R}}_{jj}^{-1}\boldsymbol{\mathcal{R}}_{ji} & \overline{\gamma}\boldsymbol{\Sigma}\boldsymbol{\mathcal{R}}_{ij} + \boldsymbol{\mathcal{R}}_{ii}^{-1}\boldsymbol{\mathcal{R}}_{ij} - \overline{\gamma}\boldsymbol{\Sigma}\boldsymbol{\mathcal{R}}_{ij} \\ -\boldsymbol{\mathcal{R}}_{jj}^{-1}\boldsymbol{\mathcal{R}}_{ji} & -\boldsymbol{I}_{MN} \end{bmatrix}.$$

$$\boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{V}_i} = E\left[\boldsymbol{y}\boldsymbol{y}^\dagger\right] = \begin{bmatrix} \overline{\gamma}\boldsymbol{\mathcal{S}}(-1)\boldsymbol{\Sigma}(0)\boldsymbol{\mathcal{S}}^\dagger(-1) + \boldsymbol{I}_{DN} & \overline{\gamma}\boldsymbol{\mathcal{S}}(-1)\boldsymbol{\Sigma}(1)\boldsymbol{\mathcal{S}}^\dagger(-1)\boldsymbol{\mathcal{V}}_i^\dagger \\ \overline{\gamma}\boldsymbol{\mathcal{V}}_i\boldsymbol{\mathcal{S}}(-1)\boldsymbol{\Sigma}^\dagger(1)\boldsymbol{\mathcal{S}}^\dagger(-1) & \overline{\gamma}\boldsymbol{\mathcal{V}}_i\boldsymbol{\mathcal{S}}(-1)\boldsymbol{\Sigma}(0)\boldsymbol{\mathcal{S}}^\dagger(-1)\boldsymbol{\mathcal{V}}_i^\dagger + \boldsymbol{I}_{DN} \end{bmatrix}.$$

maximize over both, $\boldsymbol{V}_i$, and all possible transmitted previous signals $\boldsymbol{S}(-1)$

$$\Phi^{\text{DC-ML}}\colon \hat{i} = \arg \min_{\substack{1\le i\le S\\ \boldsymbol{S}(-1)}} \boldsymbol{y}^\dagger \boldsymbol{K}^{-1}_{\boldsymbol{yy}|\boldsymbol{V}_i}\boldsymbol{y} + \ln\left|\boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{V}_i}\right| = \arg \min_{\substack{1\le i\le S\\ \boldsymbol{S}(-1)}} \delta_i^{\text{DC-ML}}.$$

Consequently, this detector has, in general, a complexity that is proportional to the product of the number of information symbols S and the number of admissible previously sent signals, which might be infinity if no further structure is enforced on the signals. Fortunately, there are two ways of avoiding this complexity.

First, if the fading correlation matrices $\boldsymbol{\Sigma}(0)$ and $\boldsymbol{\Sigma}(1)$ have certain structures, for example, if they are proportional to the identity matrix, then the correlation matrix $\boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{V}_i}$ (and hence $\Phi^{\text{DC-ML}}$) does not depend on $\boldsymbol{S}(-1)$ for $D = M$ and the right choice of signals. Some such fading correlation matrices and signals are explored in Appendix D.

Second, we consider the asymptotic expansion of the ML detector for slow fading, i.e., $\boldsymbol{h}(-1) = \boldsymbol{h}(0)$ and hence $\boldsymbol{\Sigma}(0) = \boldsymbol{\Sigma}(1) = \boldsymbol{\Sigma}$ (correspondingly, $\boldsymbol{\Omega}(0) = \boldsymbol{\Omega}(1) = \boldsymbol{\Omega}$). To this end, we apply the formula for the inverse of a partitioned matrix to $\boldsymbol{K}^{-1}_{\boldsymbol{yy}|\boldsymbol{V}_i}$ and a determinantal formula to $|\boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{V}_i}|$ (cf. [30, Secs. 0.7.3 and 0.8.5]). After some algebra, including the application of Woodbury's identity, one finds

$$\lim_{\overline{\gamma}\to\infty} \boldsymbol{K}^{-1}_{\boldsymbol{yy}|\boldsymbol{V}_i} = \frac{1}{2}\begin{bmatrix} \boldsymbol{I}_{DN} & -\boldsymbol{\mathcal{V}}_i^\dagger \\ -\boldsymbol{\mathcal{V}}_i & \boldsymbol{I}_{DN}\end{bmatrix}$$

and $\lim_{\overline{\gamma}\to\infty} \overline{\gamma}^{2DN}|\boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{V}_i}| = 1$. Because of the latter limit, for SNR $\to \infty$, the determinantal term is equal for all signals and can thus be neglected by the asymptotically optimum (AO) receiver. The AO, differentially coherent receiver for slow fading can then be written as

$$\Phi^{\text{DC-AO}}: \hat{i} = \arg \min_{1\le i\le S} \boldsymbol{y}^\dagger \boldsymbol{F}_i^{\text{DC-AO}}\boldsymbol{y} = \arg \min_{1\le i\le S} \delta_i^{\text{DC-AO}} \tag{40}$$

where

$$\boldsymbol{F}_i^{\text{DC-AO}} = \lim_{\overline{\gamma}\to\infty} \boldsymbol{K}^{-1}_{\boldsymbol{yy}|\boldsymbol{V}_i} - \frac{1}{2}\boldsymbol{I}_{2DN}.$$

For noncoherent detection, we introduced a GLRT receiver and realized it was AO (for unitary signals). Since differentially coherent detection in slow fading can be viewed as noncoherent detection over two consecutive symbol intervals, we expect a similar result, i.e., $\Phi^{\text{DC-AO}}$ should be equivalent to the differentially coherent GLRT receiver $\Phi^{\text{DC-G}}$. For unitary $\boldsymbol{S}(k)$ the equivalence of $\Phi^{\text{DC-AO}}$ to the noncoherent GLRT receiver $\Phi^{\text{NC-G}}$ can be seen by inserting

$$\boldsymbol{S}_i = [\boldsymbol{S}^\top(-1)\ \boldsymbol{S}^\top(-1)\boldsymbol{V}_i^\top]^\top$$

into $\Phi^{\text{NC-G}}$, (34). To show that $\Phi^{\text{DC-AO}}$ and $\Phi^{\text{DC-G}}$ are also equivalent for nonunitary $\boldsymbol{S}(k)$, consider the likelihood of the sufficient statistics $\boldsymbol{y} = [\boldsymbol{y}^\top(-1)\ \boldsymbol{y}^\top(0)]^\top$, given the information symbol $\boldsymbol{V}_i$, the previously transmitted signal $\boldsymbol{S}(-1)$, and the fading coefficients $\boldsymbol{h}(-1) = \boldsymbol{h}(0) = \boldsymbol{h}$

$$p(\boldsymbol{y}|\boldsymbol{V}_i,\, \boldsymbol{S}(-1),\, \boldsymbol{h}) = \frac{e^{-\|\boldsymbol{y}(-1)-\sqrt{\overline{\gamma}}\boldsymbol{\mathcal{S}}(-1)\boldsymbol{h}\|^2} e^{-\|\boldsymbol{y}(0)-\sqrt{\overline{\gamma}}\boldsymbol{\mathcal{V}}_i\boldsymbol{\mathcal{S}}(-1)\boldsymbol{h}\|^2}}{\pi^{2DN}}.$$

Maximizing $p(\boldsymbol{y}|\boldsymbol{V}_i,\, \boldsymbol{S}(-1),\, \boldsymbol{h})$ first over $\boldsymbol{x} = \sqrt{\overline{\gamma}}\boldsymbol{\mathcal{S}}(-1)\boldsymbol{h}$ and subsequently over $\boldsymbol{V}_i$ yields the GLRT receiver. The maximizing $\hat{\boldsymbol{x}}$ is found by a technique similar to completing the square and is

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y}(-1) - \boldsymbol{x}\|^2 + \|\boldsymbol{y}(-1) - \boldsymbol{\mathcal{V}}_i\boldsymbol{x}\|^2 = \frac{1}{2}\left(\boldsymbol{y}(-1) + \boldsymbol{\mathcal{V}}_i^\dagger\boldsymbol{y}(0)\right).$$

We insert $\hat{\boldsymbol{x}}$ into $p(\boldsymbol{y}|\boldsymbol{V}_i,\, \boldsymbol{S}(-1),\, \boldsymbol{h})$, which, consequently, only depends on $\boldsymbol{V}_i$. Maximizing over $\boldsymbol{V}_i$ yields the GLRT receiver as

$$\Phi^{\text{DC-G}}\colon \hat{i} = \arg\min_{1\le i\le S} \boldsymbol{y}^\dagger \begin{bmatrix} \boldsymbol{0}_{DN} & -\frac{1}{2}\boldsymbol{\mathcal{V}}_i^\dagger \\ -\frac{1}{2}\boldsymbol{\mathcal{V}}_i & \boldsymbol{0}_{DN}\end{bmatrix}\boldsymbol{y} = \arg\min_{1\le i\le S} \boldsymbol{y}^\dagger \boldsymbol{F}_i^{\text{DC-G}}\boldsymbol{y} = \arg\min_{1\le i\le S} \delta^{\text{DC-G}}.$$

Obviously, $\boldsymbol{F}_i^{\text{DC-G}} = \boldsymbol{F}_i^{\text{DC-AO}}$ ($\delta_i^{\text{DC-G}} = \delta_i^{\text{DC-AO}}$) and the GLRT detector equals the asymptotically optimum (AO) receiver.

Since we assumed slow fading for the GLRT/AO receiver, we will analyze them for slow fading only. The ML detector is, however, also applicable for fast fading, and we derive the pairwise error floor. For analytic simplicity, we will assume that $\boldsymbol{\Sigma}(1) = \rho^*\boldsymbol{\Sigma}(0) = \rho^*\boldsymbol{\Sigma}$ (hence $\boldsymbol{\Omega}(1) = \rho^*\boldsymbol{\Omega}$), where ρ is a fade rate with absolute value between zero and one. The fade rate can, for example, be connected to the normalized Doppler bandwidth B_DT by $\rho = J_0(\pi B_DT)$ for fading with power spectrum density according to Jakes's model [17]

1) Slow Fading: For uncorrelated fading and differential detection, a 3-dB performance loss when compared to coherent detection was identified in [10] and [11] for optimal ML detection. We expect the result to hold also for correlated fading and find the exact asymptotic pairwise symbol error probability of the ML and GLRT/AO receivers.

Proposition 7 (Asymptotic Pairwise Error Probability for Differentially Coherent Detection in Slow Fading): The probability $\Pr\{\delta_j^{\text{DC-ML}} < \delta_i^{\text{DC-ML}}\}$ asymptotically approaches

$$\Pr\left\{\delta_j^{\text{DC-G}} < \delta_i^{\text{DC-G}}\right\} = \Pr\left\{\delta_j^{\text{DC-AO}} < \delta_i^{\text{DC-AO}}\right\}$$

which in turn approaches $\Pr^{\text{a}}\{\delta_j^{\text{DC}} < \delta_i^{\text{DC}}\}$, i.e.,

$$\Pr\left\{\delta_j^{\text{DC-ML}} < \delta_i^{\text{DC-ML}}\right\} \to \Pr^{\text{a}}\left\{\delta_j^{\text{DC}} < \delta_i^{\text{DC}}\right\} = \frac{\overline{\gamma}^{-DN}\binom{2DN-1}{DN}}{|\boldsymbol{\mathcal{S}}(-1)\boldsymbol{\Sigma}\boldsymbol{\mathcal{S}}^\dagger(-1)||0.5(\boldsymbol{V}_i - \boldsymbol{V}_j)^\dagger(\boldsymbol{V}_i - \boldsymbol{V}_j)|^N}$$

if $\boldsymbol{A}_{ij} = (\boldsymbol{V}_i - \boldsymbol{V}_j)^\dagger(\boldsymbol{V}_i - \boldsymbol{V}_j)$ has full rank D. If $\boldsymbol{A}_{ij}$ does not have full rank, the diversity order gain will be smaller than DN.

As one might expect, for $D < M$ the total order of diversity is only DN, so that it is not advisable to use fewer dimensions than transmit antennas. Note that although the detectors (at least asymptotically) do not depend on the previously transmitted signal $\boldsymbol{S}(-1)$, the asymptotic error probability seemingly does. However, since for square matrices $\boldsymbol{A}$, $\boldsymbol{B}$ we have $|\boldsymbol{AB}| = |\boldsymbol{BA}|$, the probability $\Pr^{\text{a}}\{\delta_j^{\text{DC}} < \delta_i^{\text{DC}}\}$ really only depends on the first, unmodulated $\boldsymbol{S}(k)$. For cases in which $|\boldsymbol{\mathcal{S}}(-1)\boldsymbol{\Sigma}\boldsymbol{\mathcal{S}}^\dagger(-1)| = |\boldsymbol{\Sigma}|$ (as in i.i.d. fading and unitary $\boldsymbol{S}(-1)$, or fading and signal structure as presented in Appendix D) there is obviously the conjectured 3-dB gap between the coherent and differentially coherent detection of $\boldsymbol{V}_i$.

Proof: We already showed that asymptotically the ML receiver is equivalent with the GLRT/AO receiver, so it is sufficient to analyze the latter by finding the asymptotic eigenvalues of $\boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{V}_i}\boldsymbol{F}^{\text{DC-G}}$. This task is simplified by comparing it with the coherent case presented in Section IV-A. We replace $\boldsymbol{S}_i$ and $\boldsymbol{z}$ of that analysis with $\boldsymbol{V}_i$ and $\boldsymbol{y}$ of this section and see that the same analysis applies with only minor modifications. Notably, the abbreviations $\boldsymbol{X}, \boldsymbol{Y}$, and $\boldsymbol{Z}$ are now defined as

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{I}_{DN} \\ \boldsymbol{\mathcal{V}}_i \end{bmatrix} \tag{41}$$

$$\boldsymbol{Y} = \frac{1}{2}\boldsymbol{\Omega}(\boldsymbol{\mathcal{V}}_i - \boldsymbol{\mathcal{V}}_j)^\dagger [-\boldsymbol{\mathcal{V}}_j \quad \boldsymbol{I}_{DN}] \tag{42}$$

$$\boldsymbol{Z} = \frac{1}{2}\begin{bmatrix} \boldsymbol{0}_{DN} & \boldsymbol{\mathcal{V}}_i^\dagger - \boldsymbol{\mathcal{V}}_j^\dagger \\ \boldsymbol{\mathcal{V}}_i - \boldsymbol{\mathcal{V}}_j & \boldsymbol{0}_{DN} \end{bmatrix}. \tag{43}$$

Exactly as in the coherent case, one finds the positive eigenvalues linear in the SNR as the eigenvalues of $\overline{\gamma}\boldsymbol{YX}$. The remaining eigenvalues are the eigenvalues of

$$\boldsymbol{M} = \boldsymbol{Z} - \boldsymbol{ZX}[\boldsymbol{YX}]^{-1}\boldsymbol{Y}.$$

Making use of the fact that $\boldsymbol{V}_i$ and $\boldsymbol{V}_j$ (thus, $\boldsymbol{\mathcal{V}}_i$ and $\boldsymbol{\mathcal{V}}_j$) are unitary and inserting the definitions of $\boldsymbol{X}, \boldsymbol{Y}$, and $\boldsymbol{Z}$, one finds after some algebra

$$\boldsymbol{M} = \frac{1}{2}\begin{bmatrix} -\boldsymbol{I}_{DN} & \boldsymbol{\mathcal{V}}_i^\dagger \\ \boldsymbol{\mathcal{V}}_i & -\boldsymbol{I}_{DN} \end{bmatrix} \tag{44}$$

where we assumed that $\boldsymbol{A}_{ij} = (\boldsymbol{V}_i - \boldsymbol{V}_j)^\dagger(\boldsymbol{V}_i - \boldsymbol{V}_j)$ has full rank D. The only nonzero eigenvalue of $\boldsymbol{M}$ is -1 with multiplicity DN, as is easily obtained from $\boldsymbol{TMT}^{-1}$ with

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{I}_{DM} & -\boldsymbol{\mathcal{V}}_i^\dagger \\ \boldsymbol{0}_{DM} & \boldsymbol{I}_{DM} \end{bmatrix}. \tag{45}$$

□

The asymptotic optimality of the GLRT receiver extends the results of [32], because the one transmit antenna case considered therein is a special case of the result obtained here. However, to show the application of the results here, we show in an example the computation of the asymptotic error rate for a single transmit antenna.

Example 4: Differential Phase-Shift Keying (DPSK) with N Receive Antennas: Consider the usual DPSK setup, i.e., $S = 2$, $M = 1$, and $D = 1$, received by N antennas suffering from correlated fading with correlation matrix $\boldsymbol{\Sigma}$. The BER then asymptotically approaches

$$P^{\text{DPSKa}} = \frac{(2\overline{\gamma})^{-N}}{|\boldsymbol{\Sigma}|}\binom{2N-1}{N} \tag{46}$$

as expected from [32] we see an asymptotic 3-dB gap between coherent BPSK detection and noncoherent DPSK detection, independent of the fading correlation. The reader is invited to obtain the more general result of [32] that includes waveform (for example, multipath) diversity with signal correlations by applying the effective signal matrix concept as presented in Section II-B.

Example 5: Binary Cyclic Group (BCG) Code: Consider the BCG code presented in [11] (it can also be obtained from the results in [10]) with $\boldsymbol{V}_1 = \boldsymbol{I}_2$ and $\boldsymbol{V}_2 = -\boldsymbol{I}_2$. The exact asymptotic symbol/bit error probability for i.i.d. fading ($2\boldsymbol{\Sigma} = \boldsymbol{I}_{2N}$) becomes

$$P^{\text{BCGa}} = \overline{\gamma}^{-2N}\binom{4N-1}{2N}. \tag{47}$$

While this is the exact bit error probability, the best Chernoff bound based asymptotic error probability (using [9, eq. (18)]) is

$$P^{\text{BCG-Ca}} = \frac{1}{2}\left(\frac{\overline{\gamma}}{4}\right)^{-2N} \tag{48}$$

which is 2.13 dB off from the exact asymptote for $N = 1$. Also, for $N = 1$ the transmit diversity scheme looses 3 dB in energy efficiency *and* 0.5 b/s/Hz in spectral efficiency when compared to the receiver diversity.

In [11], it is pointed out that the differentially encoded version of Alamouti's scheme (presented in [33]) is a group code for binary symbols (s_0, s_1). We leave it to the reader to verify the 1.5- to 3.5-dB penalty (in terms of SNR per bit) in comparison to $N = 2$ receive antenna diversity.

2) Fast Fading: For fast fading we consider only the ML receiver, which is asymptotically not equivalent in this case to the GLRT. For high SNR, the ML receiver reaches an error floor. For $N = 1$ receiver antenna, a formula for the floor is given. For more receive antennas, the formula involves derivatives.

Proposition 8 (Error Floor for Differentially Coherent Detection in Fast Fading): For N receive antennas the error floor reached for $\Pr\{\delta_j^{\text{DC-ML}} < \delta_i^{\text{DC-ML}}\}$ is

$$\Pr{}^{\text{f}}\left\{\delta_j^{\text{DC-ML}} < \delta_i^{\text{DC-ML}}\right\} = \frac{1}{(N-1)!}\sum_{k=1}^{D}\frac{d^{N-1}}{ds^{N-1}}\left.\frac{-1}{s\prod_{\substack{l=1\\l\neq k}}^{2D}\lambda_l^N\left(s+\frac{1}{\lambda_l}\right)^N}\right|_{s_k=-\lambda_k^{-1}}$$

where

$$\lambda_l = \frac{|\rho|^2}{1-|\rho|^2}\,\sigma_l^R - \frac{|\rho|}{1-|\rho|^2}\,\sqrt{\sigma_l^{R^2} + (1-|\rho|^2)\sigma_l^{I^2}}$$

and

$$\lambda_{l+D} = \frac{|\rho|^2}{1-|\rho|^2}\,\sigma_l^R + \frac{|\rho|}{1-|\rho|^2}\,\sqrt{\sigma_l^{R^2} + (1-|\rho|^2)\sigma_l^{I^2}}$$

for $1 \le l \le D$. σ_l^R and σ_l^I are the real and imaginary parts of the D nonzero eigenvalues of $\boldsymbol{I} - \boldsymbol{V}_j\boldsymbol{V}_i^\dagger$ assumed to be distinct. For $N = 1$ receive antenna, the error floor reached for detecting symbol $\boldsymbol{V}_j$ if symbol $\boldsymbol{V}_i$ is transmitted is

$$\Pr{}^f\left\{\delta_j^{\text{DC-ML}} < \delta_i^{\text{DC-ML}}\right\} = \sum_{k=1}^{D}\frac{1}{\prod_{\substack{l=1\\l\neq k}}^{2D}\left(1-\frac{\lambda_l}{\lambda_k}\right)}.$$

Proof: For fast fading, the asymptotic expression for $\boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{V}_i}^{-1}$ can be found as

$$\lim_{\overline{\gamma}\to\infty} \overline{\gamma}\boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{V}_i}^{-1} = \begin{bmatrix} \frac{1}{1-|\rho|^2}\,\Omega^{-1} & \frac{\rho^*}{1-|\rho|^2}\,\Omega^{-1}\boldsymbol{\mathcal{V}}_i^\dagger \\ \frac{\rho}{1-|\rho|^2}\,\boldsymbol{\mathcal{V}}_i\Omega^{-1} & \frac{1}{1-|\rho|^2}\,\Omega^{-1} \end{bmatrix}.$$

The asymptotic expression for $\boldsymbol{C}_{ij}^{\rm DC}$ can subsequently be found as

$$\begin{aligned} \boldsymbol{C}_{ij}^{\rm DCa} &= \lim_{\overline{\gamma}\to\infty} \boldsymbol{C}_{ij}^{\rm DC} = \lim_{\overline{\gamma}\to\infty} \boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{V}_i}\left(\boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{V}_j}^{-1} - \boldsymbol{K}_{\boldsymbol{yy}|\boldsymbol{V}_i}^{-1}\right) \\ &= \begin{bmatrix} \frac{|\rho|^2}{1-|\rho|^2}\,\boldsymbol{\mathcal{V}}_i^\dagger(\boldsymbol{\mathcal{V}}_i - \boldsymbol{\mathcal{V}}_j) & \frac{\rho^*}{1-|\rho|^2}\,(\boldsymbol{\mathcal{V}}_i - \boldsymbol{\mathcal{V}}_j)^\dagger \\ \frac{\rho}{1-|\rho|^2}\,(\boldsymbol{\mathcal{V}}_i - \boldsymbol{\mathcal{V}}_j) & \frac{|\rho|^2}{1-|\rho|^2}\,\boldsymbol{\mathcal{V}}_i(\boldsymbol{\mathcal{V}}_i - \boldsymbol{\mathcal{V}}_j)^\dagger \end{bmatrix}. \end{aligned}$$

Since the eigenvalues of $\boldsymbol{C}_{ij}^{\rm DCa}$ and $\boldsymbol{T}\boldsymbol{C}_{ij}^{\rm DCa}\boldsymbol{T}^{-1}$ are the same, we may choose

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{\mathcal{V}}_i & \boldsymbol{0}_{DN} \\ \boldsymbol{0}_{DN} & \boldsymbol{I}_{DN} \end{bmatrix} \tag{49}$$

and obtain

$$\begin{aligned} &\boldsymbol{T}\boldsymbol{C}_{ij}^{\rm DCa}\boldsymbol{T}^{-1} \\ &= \begin{bmatrix} \frac{|\rho|^2}{1-|\rho|^2}\left(\boldsymbol{I}_{DN} - \boldsymbol{\mathcal{V}}_j\boldsymbol{\mathcal{V}}_i^\dagger\right) & \frac{\rho^*}{1-|\rho|^2}\left(\boldsymbol{I}_{DN} - \boldsymbol{\mathcal{V}}_i\boldsymbol{\mathcal{V}}_j^\dagger\right) \\ \frac{\rho}{1-|\rho|^2}\left(\boldsymbol{I}_{DN} - \boldsymbol{\mathcal{V}}_j\boldsymbol{\mathcal{V}}_i^\dagger\right) & \frac{|\rho|^2}{1-|\rho|^2}\left(\boldsymbol{I}_{DN} - \boldsymbol{\mathcal{V}}_i\boldsymbol{\mathcal{V}}_j^\dagger\right) \end{bmatrix}. \end{aligned}$$

Since $\boldsymbol{I}_{DN} - \boldsymbol{\mathcal{V}}_j\boldsymbol{\mathcal{V}}_i^\dagger$ is normal, one can find a unitary matrix $\boldsymbol{Q}$ such that $\boldsymbol{Q}^\dagger(\boldsymbol{I} - \boldsymbol{\mathcal{V}}_j\boldsymbol{\mathcal{V}}_i^\dagger)\boldsymbol{Q} = \boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues [30]. Applying another similarity transformation with $\boldsymbol{Q}^\dagger$ as diagonal blocks, we are finally left with the matrix

$$\begin{bmatrix} \frac{|\rho|^2}{1-|\rho|^2}\,\boldsymbol{\Lambda} & \frac{\rho^*}{1-|\rho|^2}\,\boldsymbol{\Lambda}^* \\ \frac{\rho}{1-|\rho|^2}\,\boldsymbol{\Lambda} & \frac{|\rho|^2}{1-|\rho|^2}\,\boldsymbol{\Lambda}^* \end{bmatrix} \tag{50}$$

whose eigenvalues can be easily calculated as the $\{\lambda_l\}_{l=1}^{2D}$ as given in the proposition. □

Unfortunately, the formula for the asymptotic error floor $\Pr^f\{\delta_j^{\text{DC-ML}} < \delta_i^{\text{DC-ML}}\}$ does not seem to lend itself to an easy interpretation in terms of signal design for the fast fading channel.

V. Conclusion

In this paper, we present a general asymptotic error rate analysis of quadratic receivers for Rayleigh-fading channels. This analysis permits a unified approach to the analysis of coherent, differentially coherent, and noncoherent space–time receivers for general multiantenna modulation schemes and/or fading channel models. Although we have confined our attention to space–time systems in this paper, we have found that our analysis methodology is also applicable for multiuser communications problems (cf. [16], [13], [14]). While the same asymptotic eigenvalue structure also occurs in the latter problems, finding this structure requires a more delicate algebraic analysis in the multiuser setting. So far, we do not have an insight into why this eigenvalue structure arises in so many, seemingly different problems. However, it is clear that the number of positive asymptotic eigenvalues that are proportional to the SNR is always equal to the diversity order. It would be interesting to discover other applications in which similar problems arise.

Appendix A
Derivation of the Cumulative Distribution Function of a Hermitian Quadratic Form in Zero-Mean, Complex Gaussian Random Variables

The calculation of the cumulative distribution function (cdf) in terms of residues seems to be originally performed in [23]. Reference [25] considers only the probability that the quadratic form is less or equal to zero. In [24], details of the derivation are given, that were omitted in [23]. However, the derivation here differs in some steps from [24] in that it pays more attention to some issues. Furthermore, it sets the stage for the derivations in Appendix B.

The residues arise from the calculation of complex contour integrals, which are shown to be equivalent to the integral of the inverse two-sided Laplace transform used to obtain the characteristic function of the quadratic form.

Consider a quadratic form

$$f = \boldsymbol{x}^\dagger\boldsymbol{F}\boldsymbol{x} \tag{A1}$$

in V linearly independent, $\mathcal{CN}(\boldsymbol{0}_{V\times 1}, \boldsymbol{K})$ distributed random variables x_v $(1 \le v \le V, \boldsymbol{x} = [x_1, x_2, \ldots, x_V]^\top)$.[4] The covariance matrix $\boldsymbol{K}$ of the random vector $\boldsymbol{x}$ is defined as

$$\boldsymbol{K} = E\left[\boldsymbol{x}\boldsymbol{x}^\dagger\right]. \tag{A2}$$

The resulting complex Gaussian probability density function (pdf) is

$$p_{\boldsymbol{x}}(\boldsymbol{x}) = \frac{1}{\pi^V|\boldsymbol{K}|}\,\mathrm{e}^{-\boldsymbol{x}^\dagger\boldsymbol{K}^{-1}\boldsymbol{x}}. \tag{A3}$$

Note that other authors may define $\boldsymbol{K}$ differently from (A2), leading to changes in (A3) and in what follows.

If $\boldsymbol{F}$ in (A1) is Hermitian, it is easily seen that f is real. In this paper, we will always consider a Hermitian matrix $\boldsymbol{F}$, not necessarily positive or negative semidefinite. In [38, Appendix B], the characteristic function of a Hermitian quadratic form f in $\mathcal{CN}(\boldsymbol{m}, \boldsymbol{K})$ distributed random variables is derived by diagonalizing $\boldsymbol{K}$ and $\boldsymbol{F}$. The simplified result for zero-mean $\boldsymbol{x}$ is

$$G_f(s) = E[\mathrm{e}^{-sf}] = \frac{1}{\prod_{l=1}^{l=L}\lambda_l^{\mu_l}(s+\lambda_l^{-1})^{\mu_l}}, \tag{A4}$$

where L is the number of the distinct nonzero eigenvalues λ_l of $\boldsymbol{K}\boldsymbol{F}$ with multiplicity μ_l. In what follows, it will be convenient to assume that the K eigenvalues $\{\lambda_l\}_{l=1}^{l=K}$ are negative and that the remaining $L-K$ eigenvalues $\{\lambda_l\}_{l=K+1}^{L}$ are positive. The notation employed here differs from [38, Appendix B],

[4]For zero-mean random variables, this assumption can be easily dropped without changing the resulting characteristic function or the results of this and the following appendix.

because $\boldsymbol{K}$ is twice the complex conjugate of the correlation matrix used therein ($\boldsymbol{K} = 2\boldsymbol{R}^*$) and the factor of 2 in the characteristic function vanishes. Furthermore, we state the characteristic function $G_f(s)$ only in the nonzero eigenvalues of $\boldsymbol{KF}$ and use a two-sided Laplace transform rather than the usual Fourier transform or one-sided Laplace transform for positive-definite $\boldsymbol{F}$ as in [38, Appendix B]. Note that other authors prefer to call the Laplace transform of the pdf the moment-generating function, because the coefficients of its Taylor expansion around zero are proportional to the moments of the random variable f [26].

The pdf of f is the inverse two-sided Laplace transform of $G_f(s)$ and the cdf is easily obtained from it by integration. To use contour integration, the two cases $c \leq 0$ and $c > 0$ have to be distinguished and $P_f(c)$, the cdf of f, becomes

$$P_f(c) = \begin{cases} \int_{-\infty}^{c} \int_{-j\infty+\epsilon}^{j\infty+\epsilon} \mathrm{e}^{sf} G_f(s) \frac{ds\, df}{2\pi j}, & \text{for } c \leq 0 \\ 1 - \int_{c}^{\infty} \int_{-j\infty-\epsilon}^{j\infty-\epsilon} \mathrm{e}^{sf} G_f(s) \dfrac{ds\, df}{2\pi j}, & \text{for } c > 0 \end{cases} \tag{A5}$$

where $\epsilon > 0$ is such that $\pm\epsilon$ is within the regularity domain of the Laplace transform (a vertical strip including the imaginary axis). Details for the further derivation are given for the case $c \leq 0$, for $c > 0$ only some particularities to this case are presented. The first crucial step in the derivation is the interchange of the integrals over s and f, which needs to be justified. In [24], it only appears that this issue is circumvented by invoking an integration by parts. However, it is not clear that $\mathrm{e}^{-(\epsilon+jt)x} P_f(x)|_{x=-\infty}^{\infty}$ is always zero as stated in the derivation without constraints on $P_f(x)$ or other justifications.

Consider the Case $c \leq 0$: The interchange of integrals is justified by Fubini's theorem [39, Theorem 7.8], because after a change of variable $s = jt + \epsilon$ the absolute value of $\mathrm{e}^{(jt+\epsilon)f} G_f(jt + \epsilon)$ is easily seen to be integrable over t and f. Consequently, we can write

$$\begin{aligned} P_f(c) &= \int_{-j\infty+\epsilon}^{j\infty+\epsilon} \int_{-\infty}^{c} \mathrm{e}^{sf} G_f(s) \frac{df\, ds}{2\pi j} \\ &= \int_{-j\infty+\epsilon}^{j\infty+\epsilon} s^{-1} \mathrm{e}^{sc} G_f(s) \frac{ds}{2\pi j}. \end{aligned} \tag{A6}$$

The last integral in (A6) will be solved using a complex contour integral. The path of integration in (A6) lies in the right half plane, parallel to the imaginary axis, from minus infinity to plus infinity. This integral can be replaced with the integral along the path $\Gamma_1(R)$ according to Fig. 3. If the radius R of the half-circle goes to infinity, we obviously get the same path of integration as in (A6) plus an additional integral along the half-circle. As this latter integral gives zero as R goes to infinity for $G_f(s)$ as in (A4), we can replace the integral in (A6) by the complex contour integral for $R \to \infty$. A proof that the integral along the half-circle actually is zero for $R \to \infty$ is not too hard and depends critically on $c \leq 0$. From the theory of complex contour integrals and residues it is well known that the value of a contour integral along the contour $\Gamma_1(R)$ only depends on the poles of the integrand enclosed by the contour. For sufficiently large R, all poles will be enclosed in the right half plane. In equations we can summarize

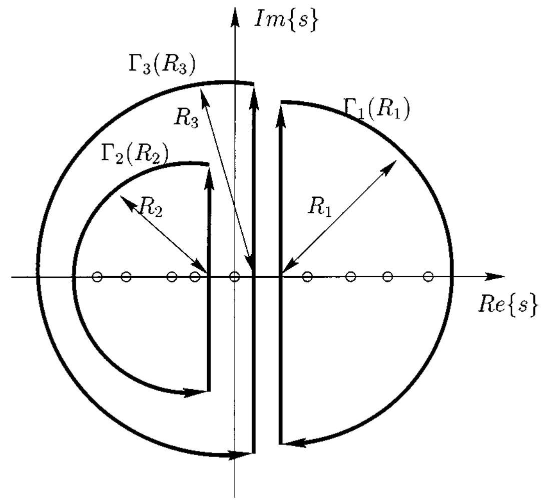

Fig. 3. Paths of integration.

$$\begin{aligned} P_f(c) &= \int_{-j\infty+\epsilon}^{j\infty+\epsilon} s^{-1} e^{sc} G_f(s) \frac{ds}{2\pi j} \\ &= \lim_{R\to\infty} \int_{\Gamma_1(R)} s^{-1} \mathrm{e}^{sc} G_f(s) \frac{ds}{2\pi j} \\ &= \int_{\Gamma_1(R_1)} s^{-1} \mathrm{e}^{sc} G_f(s) \frac{ds}{2\pi j} \end{aligned}$$

for $c \leq 0$ and R_1 sufficiently large to enclose all poles of the right half-plane. The integral along the contour $\Gamma_1(R_1)$ can be solved using Cauchy's theorem [39]

$$\begin{aligned} P_f(c) &= \int_{\Gamma_1(R_1)} s^{-1} \mathrm{e}^{sc} G_f(s) \frac{ds}{2\pi j} \\ &= -\sum_{k=1}^{k=K} \mathrm{Res}(s^{-1} \mathrm{e}^{sc} G_f(s),\, s_k > 0) \end{aligned}$$

where $s_k > 0$ denotes the kth of the K positive poles and a negative sign is necessary, because $\Gamma_1(R_1)$ is oriented clockwise. All poles $s_l = -\lambda_l^{-1}$ are real, because the eigenvalues λ_l of $\boldsymbol{KF}$ are always real (cf. [38]); this is indicated in Fig. 3 by the circles on the real axis. Inserting $G_f(s)$ and recalling that the K eigenvalues $\{\lambda_l\}_{l=1}^{l=K}$ are negative yields the negative sum over the residues of the M positive poles as given in Proposition 1.

Consider the Case $c > 0$: For similar reasons as before

$$\begin{aligned} &\int_{c}^{\infty} \int_{-j\infty-\epsilon}^{j\infty-\epsilon} \mathrm{e}^{sf} G_f(s) \frac{ds\, df}{2\pi j} \\ &\quad = -\int_{\Gamma_2(R_2)} s^{-1} \mathrm{e}^{sc} G_f(s) \frac{ds}{2\pi j} \\ &\quad = -\sum_{k=K+1}^{k=L} \mathrm{Res}(s^{-1} \mathrm{e}^{sc} G_f(s),\, s_k < 0). \end{aligned}$$

Inserting the definition of $G_f(s)$ yields the expression of the proposition. Since the residue at the origin is by definition unity, we can write

$$\begin{aligned} P_f(c) &= \sum_{k=K+1}^{k=L+1} \operatorname{Res}(s^{-1}e^{sc}G_f(s),\, s_k \le 0) \\ &= \int_{\Gamma_3(R_3)} s^{-1}e^{sc}G_f(s)\frac{ds}{2\pi j} \end{aligned}$$

where the last equation will be of interest in Appendix B.

Appendix B
Properties of the cdf

In this section, the behavior of $P_f(c)$ for specific asymptotic eigenvalues λ_l of $\boldsymbol{KF}$ is examined. We are interested in expressions for

$$\lim_{\overline{\gamma}\to\infty} \overline{\gamma}^H P_f(c) = \lim_{\overline{\gamma}\to\infty} \overline{\gamma}^H \int_{\Gamma_p} s^{-1}e^{sc_{\overline{\gamma}}}G_{f\overline{\gamma}}(s)\frac{ds}{2\pi j}$$

where $H \in \mathbb{N}$, $p \in \{1, 3\}$ selects the path of integration according to $c \le 0$ or $c > 0$ and we added a subscript $\overline{\gamma}$ on the terms dependent on the SNR $\overline{\gamma}$. The case $H = 0$ has been studied in [24] in the context of asymptotic error floors. We generalize to $H > 0$. Naturally, H should be such that the limit is neither zero nor infinity and hence captures the asymptotic behavior of $P_f(c)$.

The dependence of $G_{f\overline{\gamma}}(s)$ on $\overline{\gamma}$ is indirect through the eigenvalues $\{\lambda_l\}_{l=1}^{L}$ that depend on the average SNR $\overline{\gamma}$. We assume that in the limit for $\overline{\gamma} \to \infty$ $K\mu$ eigenvalues are -1 and $\mathcal{K}$ eigenvalues are distinct, positive with multiplicity μ_l of the form $\lambda_l = \alpha_l\overline{\gamma}$, where α_l is a positive constant and $\overline{\gamma}$ the average SNR. The multiplicities are assumed to be such that $\sum_{l=1}^{l=\mathcal{K}} \mu_l = K\mu$ (often we find $\mathcal{K} = K$ and $\mu_l = \mu$ for all l). Consequently, the number of nonzero eigenvalues L is $2K\mu$. Furthermore, it is assumed that $\lim_{\overline{\gamma}\to\infty} c_{\overline{\gamma}} = c$, a constant independent of $\overline{\gamma}$. It is then obvious that

$$\begin{aligned} G_\infty^{K\mu} &= \lim_{\overline{\gamma}\to\infty} \overline{\gamma}^{K\mu} s^{-1}e^{sc_{\overline{\gamma}}}G_{f\overline{\gamma}}(s) \\ &= \frac{e^{sc}}{s(s-1)^{K\mu}(-1)^{K\mu}\prod_{l=1}^{\mathcal{K}} \alpha_l^{\mu_l}s^{\mu_l}} \end{aligned}$$

where the limit is pointwise for almost all s. Since none of the poles lies on either of the contours Γ_1 and Γ_3, it is not hard to see that the absolute value of $s^{-1}e^{sc}G_f(s)$ is bounded by some positive number. Moreover, the paths of integration Γ_1 and Γ_2 are of finite length. By Lebesgue's dominated convergence theorem [39, Theorem 1.34], the limit and the integral commute

$$\begin{aligned} \lim_{\overline{\gamma}\to\infty} \overline{\gamma}^L P_f(c_{\overline{\gamma}}) &= \int_{\Gamma_p} \lim_{\overline{\gamma}\to\infty} \overline{\gamma}^{K\mu} s^{-1}e^{sc_{\overline{\gamma}}}G_{f\overline{\gamma}}(s)\frac{ds}{2\pi j} \\ &= \int_{\Gamma_p} G_\infty^{K\mu}. \end{aligned}$$

In [24], where the case $K\mu = 0$ is considered, Lebesgue's dominated convergence theorem is invoked separately for the real and the imaginary part; this separation is unnecessary. The conditions given in [24] for the real and imaginary part should probably read $O(w)$ and $o(w)$, respectively.

The contour intervals along Γ_1 and Γ_3 can once more be calculated using residues, where there is a pole of multiplicity $K\mu + 1$ at the origin and a pole of multiplicity $K\mu$ at unity.

Consider $c \le 0$ First: Apply the definition for the residue of the pole at $s = 1$ of multiplicity $K\mu$ to arrive at

$$\begin{aligned} &\lim_{\overline{\gamma}\to\infty} \overline{\gamma}^{K\mu} P_f(c) \\ &= \operatorname{Res}\left(G_\infty^{K\mu},\, s_k = 1\right) \\ &= \frac{-1}{(K\mu-1)!\prod_{l=1}^{\mathcal{K}} \alpha_l^{\mu_l}} \frac{d^{K\mu-1}}{ds^{K\mu-1}} \frac{e^{sc}}{s^{K\mu+1}(-1)^{K\mu}}\Bigg|_{s=1} \\ &= \frac{e^c}{\prod_{l=1}^{\mathcal{K}} \alpha_l^{\mu_l}} \sum_{k=0}^{K\mu-1} \binom{2K\mu-1-k}{K\mu} \frac{(-c)^k}{k!} \end{aligned} \qquad \text{(B1)}$$

which is obviously constant and greater than zero, because every term in the sum is at least nonnegative. The $(K\mu-1)$th derivative is calculated using Leibniz's rule of differentiation. Equation (B1) proves our claim for $c \le 0$. (The case $c = 0$ is included by defining $0^0 = 1$.)

Now Consider $c > 0$: Apply the definition for the residue of the pole at $s = 0$ of multiplicity $K\mu + 1$ to arrive at

$$\begin{aligned} &\lim_{\overline{\gamma}\to\infty} \overline{\gamma}^{K\mu} P_f(c_{\overline{\gamma}}) \\ &= \operatorname{Res}\left(G_\infty^{K\mu},\, s_k = 0\right) \\ &= \frac{1}{(K\mu)!\prod_{l=1}^{\mathcal{K}} \alpha_l^{\mu_l}} \frac{d^{K\mu}}{ds^{K\mu}} \frac{e^{sc}}{(s-1)^{K\mu}(-1)^{K\mu}}\Bigg|_{s=0} \\ &= \frac{1}{\prod_{l=1}^{\mathcal{K}} \alpha_l^{\mu_l}} \sum_{k=0}^{K\mu} \binom{2K\mu-1-k}{K\mu-1} \frac{c^k}{k!} \end{aligned} \qquad \text{(B2)}$$

which is obviously constant and greater than zero, because each term in the sum is greater than zero.

Appendix C
Approximation of Eigenvalues

A. Approximation of Eigenvalues of a Block-Partioned Matrix

Theorem 1: The eigenvalues of a matrix

$$\boldsymbol{M} = \begin{bmatrix} \overline{\gamma}\boldsymbol{A} + \tilde{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}$$

with $\boldsymbol{A}$ and $\boldsymbol{D}$ square are arbitrarily close to the eigenvalues of the matrices $\overline{\gamma}\boldsymbol{A}$ and $\boldsymbol{D}$ as $\overline{\gamma}$ goes to infinity if $\boldsymbol{A}$ has full rank.

Proof: We write the determinant to find the eigenvalues, apply a formula for the determinant of a block-partitioned matrix, and neglect terms in $\overline{\gamma}^{-1}$

$$\begin{aligned}|\boldsymbol{M}-\lambda\boldsymbol{I}| &= \left|\begin{bmatrix}\overline{\gamma}\boldsymbol{A}+\tilde{\boldsymbol{A}} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}-\lambda\boldsymbol{I}\right|\\ &= \left|\left(\overline{\gamma}\boldsymbol{A}+\tilde{\boldsymbol{A}}\right)-\lambda\boldsymbol{I}\right|\\ &\quad\times\left|\boldsymbol{D}-\lambda\boldsymbol{I}-\boldsymbol{C}\left(\left(\overline{\gamma}\boldsymbol{A}+\tilde{\boldsymbol{A}}\right)-\lambda\boldsymbol{I}\right)^{-1}\boldsymbol{B}\right|.\end{aligned}$$

The first determinant can be approximated for full-rank and square $\boldsymbol{A}$ as

$$\left|\left(\overline{\gamma}\boldsymbol{A}+\tilde{\boldsymbol{A}}\right)-\lambda\boldsymbol{I}\right| \approx |\overline{\gamma}\boldsymbol{A}-\lambda\boldsymbol{I}|, \qquad \text{for } \overline{\gamma} \ll 1. \tag{C1}$$

The second determinant is approximated as follows, which requires also a full-rank and square (hence invertible) $\boldsymbol{A}$ and completes the proof

$$\begin{aligned}&\left|\boldsymbol{D}-\lambda\boldsymbol{I}-\boldsymbol{C}\left(\left(\overline{\gamma}\boldsymbol{A}+\tilde{\boldsymbol{A}}\right)-\lambda\boldsymbol{I}\right)^{-1}\boldsymbol{B}\right|\\ &\quad= \left|\boldsymbol{D}-\lambda\boldsymbol{I}-\overline{\gamma}^{-1}\boldsymbol{C}\left(\left(\boldsymbol{A}+\overline{\gamma}^{-1}\tilde{\boldsymbol{A}}\right)-\overline{\gamma}^{-1}\lambda\boldsymbol{I}\right)^{-1}\boldsymbol{B}\right|\\ &\quad\approx \left|\boldsymbol{D}-\lambda\boldsymbol{I}-\overline{\gamma}^{-1}\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right|\\ &\quad\approx |\boldsymbol{D}-\lambda\boldsymbol{I}| \qquad \text{for } \overline{\gamma} \ll 1.\end{aligned}$$

B. Approximation of Eigenvalues of a Sum of Matrices

Theorem 2: The nonzero eigenvalues of the matrix

$$\boldsymbol{M} = \overline{\gamma}\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{Z}$$

are arbitrarily close to the nonzero eigenvalues of $\overline{\gamma}\boldsymbol{Y}\boldsymbol{X}$ and $\boldsymbol{Z}(\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{Y}\boldsymbol{X})^{-1}\boldsymbol{Y})$ as $\overline{\gamma}$ goes to infinity if $\boldsymbol{Y}\boldsymbol{X}$ has full rank.

Proof: Denoting the nonzero eigenvalues of any matrix $\boldsymbol{M}$ with $\lambda_{\mathsf{NZ}}(\boldsymbol{M})$ we can write

$$\begin{aligned}&\lambda_{\mathsf{NZ}}(\overline{\gamma}\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{Z})\\ &\quad= \lambda_{\mathsf{NZ}}\left(\begin{bmatrix}\overline{\gamma}\boldsymbol{X}\boldsymbol{Y} & \boldsymbol{0}\\ \overline{\gamma}\boldsymbol{Y} & \boldsymbol{0}\end{bmatrix}+\begin{bmatrix}\boldsymbol{Z} & \boldsymbol{0}\\ \boldsymbol{Y} & \boldsymbol{0}\end{bmatrix}\right)\\ &\quad= \lambda_{\mathsf{NZ}}\left(\boldsymbol{T}^{-1}\begin{bmatrix}\overline{\gamma}\boldsymbol{X}\boldsymbol{Y} & \boldsymbol{0}\\ \overline{\gamma}\boldsymbol{Y} & \boldsymbol{0}\end{bmatrix}\boldsymbol{T}+\boldsymbol{T}^{-1}\begin{bmatrix}\boldsymbol{Z} & \boldsymbol{0}\\ \boldsymbol{Y} & \boldsymbol{0}\end{bmatrix}\boldsymbol{T}\right).\end{aligned}$$

We choose the transformation matrix $\boldsymbol{T}$ to be

$$\boldsymbol{T} = \begin{bmatrix}\boldsymbol{I} & \boldsymbol{X}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix} = \begin{bmatrix}\boldsymbol{I} & -\boldsymbol{X}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}^{-1} \tag{C2}$$

and obtain

$$\lambda_{\mathsf{NZ}}(\overline{\gamma}\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{Z}) = \lambda_{\mathsf{NZ}}\left(\begin{bmatrix}\boldsymbol{Z}-\boldsymbol{X}\boldsymbol{Y} & (\boldsymbol{Z}-\boldsymbol{X}\boldsymbol{Y})\boldsymbol{X}\\ (\overline{\gamma}+1)\boldsymbol{Y} & (\overline{\gamma}+1)\boldsymbol{Y}\boldsymbol{X}\end{bmatrix}\right).$$

The nonzero eigenvalues of the last matrix will be obtained from the calculation of the determinant

$$\begin{aligned}&\left|\begin{bmatrix}\boldsymbol{Z}-\boldsymbol{X}\boldsymbol{Y} & (\boldsymbol{Z}-\boldsymbol{X}\boldsymbol{Y})\boldsymbol{X}\\ (\overline{\gamma}+1)\boldsymbol{Y} & (\overline{\gamma}+1)\boldsymbol{Y}\boldsymbol{X}\end{bmatrix}-\lambda\boldsymbol{I}\right|\\ &\quad= |(\overline{\gamma}+1)\boldsymbol{Y}\boldsymbol{X}-\lambda\boldsymbol{I}|\\ &\qquad\times\Big|\boldsymbol{Z}-\boldsymbol{X}\boldsymbol{Y}-\lambda\boldsymbol{I}-(\boldsymbol{Z}-\boldsymbol{X}\boldsymbol{Y})\boldsymbol{X}\\ &\qquad\qquad\times[(\overline{\gamma}+1)\boldsymbol{Y}\boldsymbol{X}-\lambda\boldsymbol{I}]^{-1}(\overline{\gamma}+1)\boldsymbol{Y}\Big|.\end{aligned}$$

For large values of $\overline{\gamma}$ we can hence approximate the first of these two determinants by

$$|(\overline{\gamma}+1)\boldsymbol{Y}\boldsymbol{X}-\lambda\boldsymbol{I}| \approx |\overline{\gamma}\boldsymbol{Y}\boldsymbol{X}-\lambda\boldsymbol{I}| \tag{C3}$$

which completes one part of Theorem 2, and the other determinant by

$$\begin{aligned}&\Big|\boldsymbol{Z}-\boldsymbol{X}\boldsymbol{Y}-\lambda\boldsymbol{I}-(\boldsymbol{Z}-\boldsymbol{X}\boldsymbol{Y})\boldsymbol{X}\\ &\quad\times[(\overline{\gamma}+1)\boldsymbol{Y}\boldsymbol{X}-\lambda\boldsymbol{I}]^{-1}(\overline{\gamma}+1)\boldsymbol{Y}\Big|\\ &\quad\approx \left|\boldsymbol{Z}-\boldsymbol{X}\boldsymbol{Y}-\lambda\boldsymbol{I}-(\boldsymbol{Z}-\boldsymbol{X}\boldsymbol{Y})\boldsymbol{X}[\boldsymbol{Y}\boldsymbol{X}]^{-1}\boldsymbol{Y}\right|\\ &\quad= \left|\boldsymbol{Z}-\boldsymbol{Z}\boldsymbol{X}[\boldsymbol{Y}\boldsymbol{X}]^{-1}\boldsymbol{Y}-\lambda\boldsymbol{I}\right|\end{aligned}$$

where in both approximations we neglected terms in $\overline{\gamma}^{-1}$. Note that in general $\boldsymbol{X}$ and $\boldsymbol{Y}$ are not square, which complicates the analysis. □

APPENDIX D
EXAMPLE FOR FADING CORRELATIONS

Without any further assumptions on the fading processes or signals $\boldsymbol{S}(k)$ the correlation matrix (and hence the optimal ML receiver) depends on $\boldsymbol{V}_i$ and the previously transmitted signal $\boldsymbol{S}(-1)$. The matrices

$$\boldsymbol{\mathcal{S}}(k)\boldsymbol{\Sigma}(0)\boldsymbol{\mathcal{S}}^{\dagger}(k) = \boldsymbol{\Omega}(0)$$

and

$$\boldsymbol{\mathcal{S}}(k)\boldsymbol{\Sigma}(-1)\boldsymbol{\mathcal{S}}^{\dagger}(k) = \boldsymbol{\Omega}(-1)$$

are trivially independent of $\boldsymbol{S}(k)$ for all k if $D=M$, the fading is i.i.d. ($\boldsymbol{\Sigma} = M^{-1}\boldsymbol{I}_{MN}$), and the columns of the $\boldsymbol{S}(k)$ are orthonormal [$\boldsymbol{S}^{\dagger}(k)\boldsymbol{S}(k) = \boldsymbol{I}_M$, unitary signaling, the case so far considered in the literature].

Nontrivially, if one assumes, for example, the abstract "one-ring" fading model of [17, Sec. 1.6] valid in a fixed wireless communications context and brought to attention recently in the context of space–time coding in [40], the correlation matrices of the fading processes can be described by $\boldsymbol{\Sigma}_{\mathrm{UL}}(0) = \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}_M(0)$ for the uplink and $\boldsymbol{\Sigma}_{\mathrm{DL}}(0) = \boldsymbol{\Sigma}_M(0) \otimes \boldsymbol{I}_N$ for the downlink. For the uplink, if we assume knowledge of the fading correlations (not the realizations) at the transmitter, the signal matrices can be postmultiplied with $\boldsymbol{\Sigma}_M^{-(1/2)}(0)$ and hence $\boldsymbol{\mathcal{S}}(k)\boldsymbol{\Sigma}(0)\boldsymbol{\mathcal{S}}^{\dagger}(k) = \boldsymbol{I}$ for $D=M$ and previously unitary signals. For the downlink, if the number of transmitter antennas equals the number of receiver antennas ($M=D=N$), then

$\boldsymbol{S}(k)$ with orthonormal columns is sufficient to guarantee the independence of the correlation matrix from the $\boldsymbol{S}(k)$.

REFERENCES

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

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. 1999. Originally, a Bell Labs, Lucent Technology Tech. Rep., Oct. 1995..

[3] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 139–157, Jan. 1999.

[4] L. Zheng and D. N. C. Tse, "Packing spheres into the Grassmann manifold: A geometric approach to noncoherent multi-antenna channels," *IEEE Trans. Inform. Theory*, submitted for publication.

[5] 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.

[6] 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, pp. 527–537, Apr. 1999.

[7] 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, pp. 744–765, Mar. 1998.

[8] A. R. J. Hammons and H. El Gamal, "On the theory of space-time codes for PSK modulation," *IEEE Trans. Inform. Theory*, vol. 46, pp. 524–542, Mar. 2000.

[9] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, pp. 543–564, Mar. 2000.

[10] B. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun*, vol. 48, pp. 2041–2052, Dec. 2000.

[11] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2567–2578, Nov. 2000.

[12] H. Jafarkhani and V. Tarokh, "Multiple transmit antenna differential detection from generalized orthogonal designs," *IEEE Trans. Inform. Theory*, submitted for publication.

[13] M. Brehler, "A systematic approach to noncoherent detection for DPSK modulation in multiuser correlated diversity Rayleigh fading channels," Dipl.-Ing. thesis, Tech. Univ. München, München, Germany, May 1998.

[14] M. Brehler and M. K. Varanasi, "Optimum multiuser noncoherent DPSK detection in generalized diversity Rayleigh fading channels," *IEEE Trans. Inform. Theory*, submitted for publication.

[15] A. Russ and M. K. Varanasi, "Noncoherent multiuser detection for nonlinear modulation: An asymptotic analysis of minimum probability of error for the Rayleigh fading channel," in *Proc. Commun. Theory Mini-Conf. (IEEE GLOBECOM)*, Phoenix, AZ, Nov. 1997, pp. 195–199.

[16] ——, "Noncoherent multiuser detection for nonlinear modulation over the Rayleigh fading channel," *IEEE Trans. Inform. Theory*, vol. 47, pp. 295–307, Jan. 2001.

[17] W. C. Jakes, *Microwave Mobile Communications*. New York: IEEE, 1993. An IEEE Press Classic Reissue. Originally, an American Telephone and Telegraph Company publication, 1974.

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

[19] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," *Bell Syst. Tech. J*, vol. 45, no. 10, pp. 1775–1796, Dec. 1966.

[20] E. A. Lee and D. G. Messerschmitt, *Digital Communication*, 2nd ed. Boston, MA: Kluwer, 1994.

[21] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[22] M. K. Varanasi, "Parallel group detection for synchronous CDMA communication over frequency-selective Rayleigh fading channels," *IEEE Trans. Inform. Theory*, vol. 42, pp. 116–128, Jan. 1996.

[23] J. K. Cavers and P. Ho, "Analysis of the error performance of trellis-coded modulations in Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 40, pp. 74–83, Jan. 1992.

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

[25] M. J. Barrett, "Error probability for optimal and suboptimal quadratic detectors in rapid Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. 5, pp. 302–304, Feb. 1987.

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

[27] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions*. New York: Dover, 1965. National Bureau of Standards.

[28] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[29] L. Zheng and D. N. C. Tse, "Packing spheres in the Grassmann manifold: A geometric approach to the noncoherent multi-antenna channel," in *Proc. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Sept. 1999.

[30] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1993.

[31] M. L. McCloud, M. Brehler, and M. K. Varanasi, "Signal constellations for noncoherent space-time communications," in *Proc. Allerton Conf. on Communication, Control, and Computing*, Monticello, IL, Oct. 2000.

[32] M. K. Varanasi, "A systematic approach to the design and analysis of optimum DPSK receivers for generalized diversity communications over Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 1365–1375, Sept. 1999.

[33] V. Tarokh and Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1169–1174, July 2000.

[34] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.

[35] B. L. Hughes, "Optimal space-time constellations from groups," *IEEE Trans. Inform. Theory*, submitted for publication.

[36] B. Hassibi, B. M. Hochwald, A. Shokrollahi, and W. Sweldens, "Representation theory for high-rate multiple-antenna code design," *IEEE Trans. Inform. Theory*, submitted for publication.

[37] B. L. Hughes, "Extensions to the theory of differential space-time modulation," in *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, June 2000, p. 285.

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

[39] W. Rudin, *Real and Complex Analysis*, 1st ed. New York: McGraw-Hill, 1966.

[40] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.

[41] M. P. Fitz, J. Grimm, and S. Siwamogsatham, "A new view of performance analysis techniques in correlated Rayleigh fading," in *Proc. IEEE Wireless Communications and Networking Conf.*, New Orleans, LA, Sept. 1999.

[42] G. Bauch, J. Hagenauer, and N. Seshadri, "Turbo processing in transmit antenna diversity systems," *Ann. Telecommun. (Special Issue on Turbo Codes and Related Topics)*, Aug. 2001.