On the Random Coding Exponent for Differential Modulation and Detection

Siddharth K. Srinivasan and Lutz Lampe
Department of Electrical and Computer Engineering
University of British Columbia, Canada
Email: {sids.Lampe}@ece.ubc.ca

Volker Pauli
Institute for Information Transmission
Universität Erlangen-Nürnberg, Germany,
Email: pauli@LNT.de

Abstract—The random coding exponent for differential modulation with finite signal sets over block-fading channels and multiple-symbol differential detection with an observation window size of \( N \) symbols is considered. Different from and in extension to previous work, the channel coherence interval is independent of \( N \) and the case of spatial transmit diversity is included. In this context, we also devise bounds and approximations of the random coding exponent, which allow for efficient numerical evaluation. The presented numerical results provide useful information on the performance of coded differential transmission with short to moderate code lengths.

I. INTRODUCTION

Error exponent analysis is a powerful tool to study the achievable performance for coded transmission over fading channels with short to moderate code lengths due to e.g. delay constraints. In particular, random coding exponents have been used to bound error and/or outage probability for certain types of channels and different degrees of channel state information (CSI) being available, cf. e.g. the discussion and references provided in [1, Section III.C.7]. In this context the block-fading channel model is often employed to (approximately) represent realistic fading channel scenarios with possibly non-ideal interleaving [2], [3], [4]. More recently, random coding exponents for space-time (ST) modulation over block fading channels have been evaluated in [5], [6].

In this paper, as in [5], we consider Gallager's random coding exponent [7, Ch. 5] for ST transmission over block fading channels with no CSI at both the transmitter and the receiver. Different from [5], we consider bandwidth-efficient differential space-time modulation (DSTM) with practical signal constellations (e.g. [8], [9], [10]) at the transmitter and power-efficient multiple-symbol differential detection (MSDD) [11] at the receiver. The main contributions of this paper are as follows.

- We derive an expression for the random coding exponent for DSTM and MSDD with observation window size \( N \) for a block-fading channel with arbitrary coherence length. This extends previous work [12] for single-antenna differential phase-shift keying (DPSK), where transmit diversity has not been considered and where the channel coherence time was assumed equal to \( N \). The latter extension is important as it allows for a fair comparison of MSDD with different window sizes \( N \) but fixed coded diversity.

- We devise, respectively, bounds and approximations for the random coding exponent, whose numerical evaluation is feasible also for larger values of \( N \). For this purpose, we exploit that MSDD can be formulated as tree-search problem with an additive metric and apply a version of the stack algorithm.

- The presented numerical results show (a) the usefulness of the derived expressions for performance analysis for coded differential transmission with MSDD and (b) the particular benefits of spatial transmit diversity for this scenario and limited code length.

Organization: The system model is introduced in Section II, and expressions for the random coding exponent are derived in Section III. Numerical results are presented and discussed in Section IV, and conclusions are given in Section V.

Notation: \((\cdot)^{\dag}\), \((\cdot)^{T}\), tr\{\cdot\}, \|\cdot\|, \det\{\cdot\}\), and \(\mathcal{E}_{X}\{\cdot\}\) denote Hermitian transposition, transposition, trace, Frobenius norm, determinant of a matrix, and expectation with respect to \( X \), respectively. \( I_{L} \) is the \( L \times L \) identity matrix.

II. SYSTEM MODEL

A. Block-Fading Channel

We consider space-time transmission with \( N_{T} \) transmit and \( N_{R} \) receive antennas over a frequency-nonselective block-fading channel with \( L \) independent channel realizations per coding frame (cf. e.g. [2], [3], [4] for similar single-antenna scenarios). Since MSDD at the receiver is based on blocks of \( N \) received samples (see Section II-C), the stream of transmitted symbols per each fading channel is organized in blocks of \( N \) symbols. \( D \) such blocks are transmitted over one fading realization.

The transmitted symbols are \( N_{T} \times N_{T} \) matrices denoted by \( S_{\ell,d,n}[k] \) (see Section II-B), where subscripts \( \ell, d, \) and \( n, 1 \leq \ell \leq L, 1 \leq d \leq D, 1 \leq n \leq N, \) specify the fading subchannel, the position of the block transmitted over this subchannel, and the symbol position within this block, respectively, and \( k \in \mathbb{Z} \) is the time-index with respect to frames of \( n_{s} \approx LDN \) symbols. The \( \ell \)th subchannel is represented by the \( N_{T} \times N_{R} \) channel matrix \( H_{\ell}[k] \).

For a compact notation, it is convenient to introduce the matrix notation

\[
X_{\ell,d}[k] \triangleq [X_{\ell,d,1}[k] \ldots X_{\ell,d,N}[k]]^{T},
\]

\[
X_{\ell}[k] \triangleq [X_{\ell,1}[k] \ldots X_{\ell,D}[k]]^{T},
\]

where \( X_{\ell,d}[k] \) is the transmitted signal vector over the \( \ell \)th subchannel and \( d \)th fading realization at time \( k \).

The transmitted symbols are \( N_{T} \times N_{T} \) matrices denoted by \( S_{\ell,d,n}[k] \) (see Section II-B), where subscripts \( \ell, d, \) and \( n, 1 \leq \ell \leq L, 1 \leq d \leq D, 1 \leq n \leq N, \) specify the fading subchannel, the position of the block transmitted over this subchannel, and the symbol position within this block, respectively, and \( k \in \mathbb{Z} \) is the time-index with respect to frames of \( n_{s} \approx LDN \) symbols. The \( \ell \)th subchannel is represented by the \( N_{T} \times N_{R} \) channel matrix \( H_{\ell}[k] \).

For a compact notation, it is convenient to introduce the matrix notation

\[
X_{\ell,d}[k] \triangleq [X_{\ell,d,1}[k] \ldots X_{\ell,d,N}[k]]^{T},
\]

\[
X_{\ell}[k] \triangleq [X_{\ell,1}[k] \ldots X_{\ell,D}[k]]^{T},
\]
corresponding to blocks of $N$ and $DN$ matrix-symbol transmissions, respectively, and $X \in \{S, R, N\}$. Then, the $DNN_T \times N_T$ matrix $R_\ell[k]$ of $DN$ received samples can be written as

$$R_\ell[k] = S_\ell[k]H_\ell[k] + N_\ell[k], \quad 1 \leq \ell \leq L,$$  

where $N_\ell[k]$ denotes additive spatially and temporally white circularly-symmetric Gaussian noise (AWGN) with variance $\sigma_n^2$ per complex scalar component. We also assume that the elements of $H_\ell[k]$ are circularly-symmetric Gaussian distributed, i.e., block Rayleigh fading, with common variance $\sigma_h^2$ and possible spatial correlation, which is assumed identical for all subchannels $\ell$.

We note that this transmission model represents or approximates many practical systems employing coding and interleaving over channels with a finite number $L$ of degrees of freedom. For example, assuming interleaving to be limited to frames of $LDN$ space-time symbols, we have $L \approx 2f_DT_s \cdot DNN_T$ for a narrowband system and transmission over a time-selective fading channel with normalized Doppler bandwidth $f_D T_s$, and $L \approx B_s d_s$ for a static multicarrier system with bandwidth $B_s$ and transmission over a frequency-selective channel with delay spread $d_s$, respectively, cf. [3, Ch. 3].

### B. Differential Encoding

We assume that the channel states $H_\ell[k]$ are unknown to the transmitter and the receiver. To enable detection without CSI, DSTM is applied at the transmitter, cf. e.g. [8], [9], [10]. We consider DSTM with differential $N_T \times N_T$ matrix symbols $V_\ell,d,n[k]$ taken from an $M$-ary set $\mathcal{V}$ of unitary matrices and matrix-wise differential encoding given by

$$S_{\ell,d,n+1}[k] = V_{\ell,d,n}[k]S_{\ell,d,n}[k], \quad 1 \leq n \leq N - 1,$$  

where $S_{\ell,d,1}[k]$ is a unitary matrix and thus $S_{\ell,d,n}[k]$ are also unitary. Classical DPSK results as special case for $N_T = 1$.

Due to the unknown channel, the first (reference) symbol $S_{\ell,d,1}[k]$ of each block can be chosen arbitrarily. For improved bandwidth and power efficiency of DSTM, a previously transmitted symbol is typically chosen as reference symbol $S_{\ell,d,1}[k]$. For example $S_{\ell,d,1}[k] = S_{\ell,d,N}[k - 1]$ could be the case in multicarrier modulation with differential encoding among subcarriers. Considering single-carrier transmission, $S_{\ell,d,1}[k] = S_{\ell,d,n}[k]$ is typically chosen. The latter implies $N_{\ell,d,1}[k] = N_{\ell,d,n}[k]$ for the elements of the noise matrix $N_\ell[k]$ in (1). In the following, however, we neglect possible dependencies between noise samples due to such differential encoding and assume that the elements of $N_\ell[k]$ are independent identically distributed (i.i.d.). This considerably simplifies the numerical analysis presented in Section III.

In the following, without loss of generality and for the sake of clarity, we omit the frame index $k$.

### C. Coding and Decoding

At the transmitter, we consider binary channel coding, which is performed independently for frames of $n_s = LDN$ space-time transmit symbols. Assuming bandwidth-efficient differential transmission as described in Section II-B, one codeword spans $n_v \triangleq LD(N - 1)$ $M$-ary differential symbols, i.e., the code length is $n_c = n_v \log_2(M)$.

At the receiver, MSDD with an observation window size $N$ is applied. This means that the decoder input is based on independent processing of blocks of $N$ received samples collected in $R_{\ell,d}$ corresponding to $N - 1$ differential symbols

$$V_{\ell,d} = [V_{\ell,d,1}^T \ldots V_{\ell,d,N-1}^T]^T.$$  

Let $R \triangleq [R_1^T \ldots R_L^T]^T$ and $V \triangleq [V_1^T \ldots V_L^T]^T \triangleq [V_{1,1}^T V_{1,2}^T \ldots V_{L,D}^T]^T$ collect all received samples and differential symbols corresponding to the transmission of one codeword. The corresponding maximum-likelihood (ML) decoding metric is of the form

$$\lambda(R|V) = \sum_{\ell=1}^L \sum_{d=1}^D \lambda(R_{\ell,d}|V_{\ell,d}),$$  

where the metric increments $\lambda(R_{\ell,d}|V_{\ell,d})$ are log-likelihood expressions derived from the probability density function (pdf) $p(R_{\ell,d}|V_{\ell,d})$ (see Section III). The rationale for (3) is that since the size of the effective signal constellation and hence the complexity of soft-output MSDD increases exponentially with the window size $N$, the value for $N$ has often to be chosen (much) smaller than the channel coherence interval of length $DN$, i.e., $D > 1$ (cf. e.g. [11]).

### III. RANDOM CODING EXPONENT ANALYSIS

In this section, we derive the random coding exponent for the transmission system described above. To this end, we first give expressions for the necessary pdfs and the incremental metric $\lambda(R_{\ell,d}|V_{\ell,d})$ in Section III-A. Then, the exact random coding exponent is presented in Section III-B, while bounds and approximations are devised in Section III-C.

### A. Probability Density Function and Decoding Metric

The pdf for the received sequence $R$ given the sequence of differential symbols $V$ assuming coherent detection with CSI can be decomposed into the product $(H \triangleq [H_1^T \ldots H_L^T]^T)$

$$p(R|V, H) = \prod_{\ell=1}^L \prod_{d=1}^D p(R_{\ell,d}|V_{\ell,d}, H_\ell),$$  

with the $NN_T N_R$-dimensional pdf

$$p(R_{\ell,d}|V_{\ell,d}, H_\ell) = \mathcal{E}_{S_{\ell,d},1} \left\{ \frac{\exp \left( -\frac{||R_{\ell,d} - S_{\ell,d}H_\ell||^2}{\sigma_n^2} \right)}{\left( \pi \sigma_n^2 \right)^{NN_T N_R}} \right\}.$$  

As discussed in Section II-C, decoding at the receiver is based on MSDD with window size $N$ and the decoding metric is based on the noncoherent pdf $p(R_{\ell,d}|V_{\ell,d})$. More specifically, for ease of practical implementation, we consider

1 Throughout this paper, we use the short-hand notation $p(X)$ for the pdf of a random variable $X$, i.e., we do not distinguish between the variable and its realization.
a decoder designed for spatially uncorrelated fading. In this case, the pdf for $R_{\ell,d}$ conditioned on $V_{\ell}$ reads

$$
p(R_{\ell,d}|V_{\ell,d}) = \frac{\exp \left( -\text{tr} \left\{ R_{\ell,d}^H \Psi_R^{-1} R_{\ell,d} \right\} \right)}{(\pi N R_T)^{N R_T} |\det(\Psi_R)|^{N R_T}} ,
$$

where

$$
\Psi_R = \mathcal{E}_{R_{\ell,d}} \{ R_{\ell,d}^H R_{\ell,d} | V_{\ell,d} \} = N_R (\sigma_n^2 S_{\ell,d} S_{\ell,d}^H + \sigma_n^2 I_{N R_T})
$$

(7)

with some arbitrarily chosen $S_{\ell,d}$. We note that $\det{\Psi_R}$ is independent of $V_{\ell,d}$. Inserting (7) into (6) the decoding metric follows as

$$
\lambda(R_{\ell,d}|V_{\ell,d}) = \log(p(R_{\ell,d}|V_{\ell,d})) = -\|R_{\ell,d}\|^2 / N_R \sigma_n^2 + \frac{\sigma_n^2 \|R_{\ell,d}^H S_{\ell,d}\|^2}{N_R \sigma_n^2 (\sigma_n^2 + N_N \sigma_n^2)} - N_R \log(\pi N R_T |\det(\Psi_R)|) .
$$

(8)

B. Random Coding Exponent

We define $R$ as the data rate in bits per scalar channel use, i.e., $N_T R$ is the rate in bits per $M$-ary space-time channel

$$
V_{\ell,d,n}.\text{ Given coding with code length } n_c \text{, spanning } n_v \text{ space-time symbols chosen uniformly and i.i.d. from } \mathcal{V} \text{, and adapting the derivation in [7, Sec. 5.6] to the problem at hand, the average word error rate can be bounded as}
$$

$$
P_e \leq 2^{-\rho N_T n_c} \int_{H \in \mathcal{C}_{N_T \times N_R}} \int_{R \in \mathcal{C}_{N_T \times N_R}} \frac{1}{M^{n_c}} \sum_{V \in \mathcal{V}^{n_v}} p(R|V,H) [\exp(\lambda(R|V))]^{-\frac{1}{\rho}} \left( \frac{1}{M^{n_v}} \sum_{V' \in \mathcal{V}^{n_v}} \left[ \exp(\lambda(R|V')) \right]^{-\frac{1}{\rho}} \right) d(H,R)
$$

(9)

for $0 \leq \rho \leq 1$. Optimizing the parameter $\rho$ such that the upper bound (9) becomes tightest, we can write

$$
P_e \leq 2^{-n_c E_{\ell}(R,N,D)} ,
$$

(10)

where $E_{\ell}(R,N,D)$ is the random coding exponent given by

$$
E_{\ell}(R,N,D) = \max_{0 \leq \rho \leq 1} \{ E_0(\rho,N,D) - \rho N_T R \} .
$$

(11)

After straightforward manipulations using (3), (4), and (9), the Gallager function $E_0(\rho,N,D)$ is obtained as

$$
E_0(\rho,N,D) = -\frac{1}{(N-1) D} \log \left[ \mathcal{E}_{R_{\ell,d}} \{ \prod_{d=1}^{D} \left[ \exp(\lambda(R_{\ell,d}|V_{\ell,d})^{-\frac{1}{\rho}} p(R_{\ell,d}|V_{\ell,d},H_{\ell}) \right] \} \right] \left[ \exp(\lambda(R_{\ell,d}|V_{\ell,d})^{-\frac{1}{\rho}} p(R_{\ell,d}|V_{\ell,d},H_{\ell}) \right] \left[ \exp(\lambda(R_{\ell,d}|V_{\ell,d})^{-\frac{1}{\rho}} p(R_{\ell,d}|V_{\ell,d},H_{\ell}) \right]^{-1} .
$$

(12)

Some comments are in order at this point. First, expressing the right-hand side of (12) in terms of expectations allows for efficient Monte-Carlo integration, which we used to obtain numerical results presented in Section IV. Second, different

from the case of coherent detection with CSI (cf. e.g. [4, Eq. (19)], [6, Eq. (6)]) the expressions in (12) cannot further be simplified due to the applied MSDD metric. We should note that $E_0(\rho,N,D)$ could also be written in terms of the noncoherent pdf $p(R_{\ell}|V_{\ell})$ since the channel is memoryless with respect to $\ell$ for the case of decoding without CSI. However, averaging over the set $\mathcal{V}^{D(N-1)}$ of size $M^{D(N-1)}$ would be required, which renders computation of $E_0(\rho,N,D)$ using Monte-Carlo techniques quickly infeasible even for moderate values of $D$. Finally, the bound could be tightened when optimizing $\rho$ as a function of $H$ (cf. [4, Eq. (17)]), which however considerably increases computational complexity.

Special case of $D = 1$ and spatially uncorrelated fading: For the interesting special case of $D = 1$ and spatially uncorrelated channels, i.e., each block of $N$ transmitted space-time symbols experiences independent fading and (8) is the correct ML decoding metric, the Gallager function can be simplified to

$$
E_0(\rho,N,D = 1) = -\frac{1}{(N-1)} \log \left[ \mathcal{E}_{R_{\ell,1}} \left\{ \frac{p(R_{\ell,1}|V_{\ell,1})^{\frac{1}{\rho}}}{\exp(\lambda(R_{\ell,1}|V_{\ell,1})^{\frac{1}{\rho}})} \right\} \right]^{1+\rho} .
$$

(13)

C. Bounds and Approximation for $E_0(\rho,N,D)$

Different from the case of coherent detection with CSI, the MSDD metric $\lambda(R_{\ell,d}|V_{\ell,d})$ cannot be split into individual terms depending only on $V_{\ell,d,n}, 1 \leq n \leq N - 1$. Thus, averaging over the set $\mathcal{V}^{N-1}$ instead of $\mathcal{V}$ is necessary in (12), which becomes computationally expensive for large values of $N$. We therefore consider an approximation of $E_0(\rho,N,D)$ using only a subset $\mathcal{V}_s \subset \mathcal{V}^{N-1}$ of differential symbols to reduce computational complexity when numerically evaluating $E_0(\rho,N,D)$. To this end, we start with the special case of $D = 1$ and spatially uncorrelated fading and then treat the general block-fading case.

1) $D = 1$ and spatially uncorrelated fading: Considering (13) it is reasonable to choose $\mathcal{V}_s$ such that the ratio

$$
r(C_s) \triangleq \frac{\sigma_s}{\sigma} < 1
$$

(14)

with

$$
\sigma \triangleq \mathcal{E}_{R_{\ell,1}} \left\{ \frac{p(R_{\ell,1}|V_{\ell,1})^{\frac{1}{\rho}}}{\exp(\lambda(R_{\ell,1}|V_{\ell,1})^{\frac{1}{\rho}})} \right\}
$$

(15)

and

$$
\sigma_s \triangleq \sum_{V_{\ell,1} \in \mathcal{V}_s} \frac{1}{M^{N-1}} \left( p(R_{\ell,1}|V_{\ell,1}) \right)^{\frac{1}{\rho}}
$$

(16)

is maximized for a given cardinality

$$
C_s \triangleq |\mathcal{V}_s| < M^{N-1} .
$$

(17)

This means we only consider the dominating terms when averaging over $V_{\ell,1}$ and to do so, we need to sort matrices $V_{\ell,1} \in \mathcal{V}^{N-1}$ according to decreasing values of $p(R_{\ell,1}|V_{\ell,1})$.

a) Sorting: An efficient way to accomplish sorting without actually calculating $p(R_{\ell,1}|V_{\ell,1})$ for all $V_{\ell,1} \in \mathcal{V}^{N-1}$ is to run a detection algorithm which returns the best (ML) and successively the next best estimates for $V_{\ell,1}$ according to
and \( p(R_{\ell,1}|V_{\ell,1}) \) and given \( R_{\ell,1} \). As this detection requires a tree search and the search metric is additive, a modified version of the stack algorithm has been developed, which continuously finds the next best signal point. Due to space limitations we omit the implementation details for the applied stack algorithm and refer to [13] for the application of tree-search algorithms to MSDD for DSTM.

b) Choice of \( C_s \): For a given \( C_s \) the ratio \( r(C_s) \) in (14) significantly varies depending on the realization \( R_{\ell,1} \). It is therefore advisable not to fix \( C_s \) but to choose it adaptively when Monte-Carlo integrating (13). To this end, let \( V(C_s)_{\ell,1} \) be the \( C_s \)th best term found by the stack algorithm. The upper bound

\[
\sigma_u \triangleq \sigma_s + \left(1 - C_s/M^{N-1}\right) \left(p(R_{\ell,1}|V_{\ell,1}(C_s))\right)^{\frac{1}{M}} \geq \sigma \quad (18)
\]

for \( \sigma \) allows to lower bound \( r(C_s) \) by

\[
r(C_s) \geq \frac{\sigma_u}{\sigma_s}. \quad (19)
\]

We then can choose \( C_s \) such that the lower bound exceeds a certain threshold value of, e.g., 0.9.

c) Bounds for \( E_0(\rho, N, D) \): Let us define

\[
\eta \triangleq \left\{ \mathcal{E}_{\ell,1}\left[p(R_{\ell,1}|V_{\ell,1})\right]\right\}^{\frac{1}{M}} \quad (20)
\]

\[
\eta_s \triangleq \left( \sum_{V_{\ell,1} \in \mathcal{V}_s} \frac{1}{M^{N-1}} p(R_{\ell,1}|V_{\ell,1}) \right)^{\frac{1}{M}} \quad (21)
\]

\[
\eta_u \triangleq \left( \sum_{V_{\ell,1} \in \mathcal{V}_u} \frac{1}{M^{N-1}} p(R_{\ell,1}|V_{\ell,1}) + \left(1 - C_s/M^{N-1}\right) p(R_{\ell,1}|V_{\ell,1}(C_s)) \right)^{\frac{1}{M}}. \quad (22)
\]

We note that \((\sigma/\eta)^{1+\rho}\) is used in (13). Since the function \( f(x) = x^{1/(1+\rho)} \) is concave for \( 0 \leq \rho \leq 1 \), the inequalities

\[
\frac{\sigma_s}{\eta_s} \leq \frac{\sigma}{\eta} \leq \frac{\sigma_u}{\eta_u}
\]

hold. Hence, \( E_0(\rho, N, D = 1) \) in (13) is, respectively, upper and lower bounded when using the subset \( \mathcal{V}_s \) and replacing the exact ratio \( \sigma/\eta \) with \( \sigma_s/\eta_s \) and \( \sigma_u/\eta_u \), respectively.

2) Transmission with \( D > 1 \) or spatially correlated fading: Although in case of (a) spatially correlated fading \( \lambda(R_{\ell,d}|V_{\ell,d}) \) is not the ML metric or (b) transmission with \( D > 1 \) the pdf \( p(R_{\ell,d}|V_{\ell,d}, H_{\ell}) \) is used for Monte-Carlo integrating (12), and thus \( r(C_s) \) does not appear in the expression for Gallager’s function, we extend the application of the bound (19) to determine the set \( \mathcal{V}_s \) to these cases. The reasons for this are that (a) \( r(C_s) \) (or its bound (19)) is still a good indicator whether the terms dominating the argument in (12) for given \( R_{\ell} \) and \( H_{\ell} \) are taken into account and (b) sorting \( V_{\ell,d} \) according to a mixed criterion of \( \exp(\lambda(R_{\ell,d}|V_{\ell,d}))^{\rho/(1+\rho)} \) and \( p(R_{\ell,d}|V_{\ell,d}) \) or \( p(R_{\ell,d}|V_{\ell,d}, H_{\ell}) \) would depend on \( \rho \).

We note that in this case, using only the set \( \mathcal{V}_s \) does not result in bounds for the Gallager function, but rather approximates \( E_0(\rho, N, D) \).

IV. NUMERICAL RESULTS AND DISCUSSION

In this section, we present numerical results using (a) the expression (12) for \( E_0(\rho, N, D) \) to obtain \( E_r(R, N, D) \) in (11), which we refer to as “exact” random coding exponent, and (b) the approximations for \( D > 1 \) corresponding to bounds when \( D = 1 \) as introduced in Section III-C with subsets \( \mathcal{V}_s \), chosen such that \( r(C_s) \geq 0.9 \) (see (19)). We consider 4DPSK for single-antenna transmission and, respectively, 16-ary orthogonal designs (16OD) and 16-ary diagonal codes (16DI) for DSTM with \( N_T = 2 \) (cf. [8], [9], [10]). For simplicity, and since we focus on transmit diversity, \( N_R = 1 \) in the presented examples. The spatial correlation coefficient is denoted by \( \rho \). When we compare schemes with different values for \( N \) and/or \( N_T \), we use the normalized coherence length \( n_c/L = (N - 1)D \log_2(M) \) as parameter to specify the amount of coded diversity (either temporally or spectrally) available.

Fig. 1 shows the \( E_r(R, N, D) \) as function of \( R \) for 4DPSK transmission, SNR \( 10 \log_{10}(E_s/N_0) = 10 \) dB, and fixed normalized coherence length \( n_c/L = 40 \). As can be seen from Fig. 1(a), and as expected from MSDD performance results in e.g. [11], the random coding exponent for MSDD approaches that of coherent detection with CSI, which is included as reference curve, with increasing \( N \). Since numerical evaluation of (11) and (12) for \( N = 11 \) is not feasible, approximations corresponding to the upper and lower bound are shown in Fig. 1 for this case. The usefulness of these approximations is underlined by the curves in Fig. 1(b), where approximations for \( E_r(R, N, D) \) are plotted (in logarithmic scale) together with the exact numerical result for MSDD with \( N = 6 \). Especially the approximation derived from the upper bound is fairly tight.

\(^2E_s\), \( E_b \), and \( N_0 \) denote, respectively, the average received energy per symbol and per information bit and the two-sided complex noise power spectral density.
Fig. 2 illustrates the effect of transmit diversity on the random coding exponent assuming MSDD with $N = 2$, different normalized coherence lengths $n_c/L$, and $10 \log_{10}(E_b/N_0) = 10$ dB. We observe from Fig. 2(a), where spatially uncorrelated channels are assumed, that spatial diversity results in a significantly increased exponent $E_r(R, N, D)$. MSDD consistently outperforms 16DI (only shown for $n_c/L = 20$). Interestingly, comparing the curves for (4DPSK, $n_c/L = 20$) and (16OD, $n_c/L = 40$), which have identical aggregate coded and transmit diversity, shows that the benefit of spatial diversity increases for rates $R$ closer to channel capacity, where the effect of coded diversity saturates. On the other hand, given a certain target rate $R$ and error rate $P_e$, spatial diversity allows for significant reduction in code length. Moreover, from Fig. 2(b) it can be seen that these gains are also achieved if channels are highly correlated with $\rho \lesssim 0.8$. A similar trend, considering coherent transmission with receive diversity, has been reported in [14, Ch. 3].

Finally, Fig. 3 shows the SNR $10 \log_{10}(E_b/N_0)$ required to achieve $P_e \leq 10^{-3}$ according to (9) as function of code length $n_c$ for transmission with rate $R = 1$ bit per scalar channel use. The normalized coherence length is $n_c/L = 100$, i.e., degrees of freedom $L$ and thus coded diversity increase linearly with code length $n_c$. Two things are interesting to note. First, increasing $N$ consistently improves power efficiency regardless of $n_c$. This is decidedly different from [12], where coded diversity was decreasing with $N$. Second, the gain due to transmit diversity increases with decreasing $n_c$. This result reiterates the benefits of spatial transmit diversity if coded diversity is limited, e.g., due to delay constraints.

V. CONCLUSIONS

In this paper, we derived an expression for the random coding exponent for DSTM over block-fading channels and MSDD. Furthermore, we devised bounds or approximations for this expression, which allow for faster Monte-Carlo integration. Presented numerical results showed (a) the usefulness of the proposed approximations, (b) the improvements with increasing observation size $N$, and (c) the benefits of spatial transmit diversity even in highly correlated channels for coded DSTM transmission with short or moderate code lengths.

REFERENCES


