Multiple-Symbol Differential Detection With Interference Suppression

Debang Lao, *Student Member, IEEE*, and Alexander M. Haimovich, *Senior Member, IEEE*

***Abstract*—A multiple-symbol differential detector is formulated for M-ary differential phase-shift keying modulation where the channel state information is unknown to the receiver. The maximum-likelihood decision statistic is derived for the detector, and its performance is demonstrated by analysis and simulation. Under the Gaussian assumption for the aggregate interference plus noise, an exact expression for the symbol pairwise error probability is developed for M-ary differential phase-shift keying modulation over a diversity, slow-fading Rayleigh channel in the presence of an interference source. A simpler expression of the pairwise error probability is developed for the asymptotic case of large signal-to-noise ratio and small signal-to-interference ratio. It is shown that with an increasing observation interval, the performance of the differential detector over an unknown channel approaches that of optimum combining with known channel.**

***Index Terms*—Differential detection, diversity reception, error probability performance, interference suppression, optimum combining.**

I. Introduction

MULTIPLE-SYMBOL differential detection (MSDD) was first proposed for detecting multiple phase-shift keying (M-PSK) signals transmitted over an additive white Gaussian noise (AWGN) channel [1]. The main advantage of MSDD is that it does not require a coherent phase reference at the receiver (it does require, however, the ability to measure relative phase differences). MSDD performs maximum-likelihood detection of a block of information symbols based on a corresponding observation interval. The method was presented as a bridge of the gap between the performance of coherent detection of M-PSK and conventional differential detection of M-ary differential phase-shift keying (M-DPSK) [1]. The channel phase was assumed to be unknown to the receiver but constant over multiple symbol intervals. In [1] it was shown that for a long observation interval, the performance of MSDD [in terms of the required signal-to-noise ratio (SNR) for a given bit-error probability (BEP)] approaches that of coherent detection (with differential encoding at the transmitter). MSDD was extended to trellis coded M-PSK in [2]. MSDD for the fading channel was analyzed in [3], and for correlated fading in [4]. MSDD application to multiuser code-division multiple access (CDMA) was considered in [5]. Performance of MSDD with narrowband interference over a nonfading channel was discussed in [6]. A system with MSDD and reception diversity was formulated in [7], while [8] considered MSDD with transmit diversity. A class of algorithms different from MSDD is based on approximations of the optimal noncoherent maximum-likelihood sequence detector [9], [10].

In this paper, we derive an extension to MSDD for communication in the presence of a single interference source. The channel of the desired signal is a diversity Rayleigh channel with multiple outputs. For an antenna array at the receiver of a communication system operating over a slow-fading channel, any signal source is spatially correlated. The channel realizations at each output are mutually independent, constant over the observation interval, and unknown to the receiver. The Gaussian assumption is made with respect to the aggregate of interference plus noise. The covariance matrix of the interference plus noise is assumed known. The MSDD decision statistic is derived based on the principle of maximum-likelihood sequence detection (MLSD). A closed-form expression for the pairwise error probability (PEP) is derived. A closed-form expression for the BEP is intractable; however, one is obtained for an approximation to the union bound. The approximation utilizes only dominant terms in the union bound and it is shown to be a good approximation of the BEP.

The coherent counterpart of MSDD is the optimum combining detector [11]–[13], which requires a coherent reference and the channel information of the desired signal. In this paper, we show that with an increasing number of symbols in the observation interval, the performance of MSDD approaches that of optimum combining (with differential encoding at the transmitter).

In the course of designing simulations for evaluating MSDD, we realized that there was no efficient MSDD algorithm available for MSDD with diversity. The computational complexity of direct computation of the decision statistic grows exponentially with the number of symbols in the observation interval. For single-channel MSDD, an optimum algorithm was proposed in [14]. Suboptimal decision-feedback algorithms for the single-channel case were suggested in [15]–[17]. In this paper, we modify the suboptimal decision-feedback algorithm in [17] for application to MSDD with diversity. The main improvement over published algorithms is the introduction of iterations for symbol detection.

This paper is organized as follows. Section II presents the signal model. The MSDD decision statistic is derived in Sec-

Paper approved by X. Dong, the Editor for Modulation and Signal Design of the IEEE Communications Society. Manuscript received December 5, 2001; revised July 26, 2002. This work was supported in part by the National Science Foundation under Award CCR-0085846, and in part by the Air Force Office of Scientific Research under Grant F49620-00-1-0107. This paper was presented in part at the 2001 Conference on Information Sciences and Systems, Johns Hopkins University, Baltimore, MD, March, 2001.

The authors are with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: haimovic@njit.edu).

Digital Object Identifier 10.1109/TCOMM.2003.809285

0090-6778/03$17.00 © 2003 IEEE

tion III. The error analysis is developed in Section IV, while Section V presents the numerical results. Conclusions are drawn in Section VI.

II. System Model

Consider a wireless communications system operating over L independent branches. Each of the branches is a slow-fading channel that attenuates, phase shifts, and adds noise to the signal. Assuming perfect time synchronization, the sampled output of the matched filter corresponding to time k and the ℓth branch is

$$r_{k,\ell} = \sqrt{P_s} c_\ell s_k + z_{k,\ell}, \qquad \ell = 1, 2, \ldots, L \tag{1}$$

where P_s is the power of the desired signal, c_ℓ is the channel gain of the ℓth branch, s_k is the transmitted M-DPSK symbol, and $z_{k,\ell}$ is Gaussian correlated noise. For M-DPSK modulation, the transmitted signals can be expressed as $s_k = e^{j\theta_k}$, $\theta_k = 2\pi(i_k - 1)/M$, $i_k = 1, 2, \ldots, M$. The transmitted symbols are differentially encoded, i.e., $\theta_k = \theta_{k-1} + \Delta\theta_k$, where $\Delta\theta_k$ is the phase representing the transmitted information at time k.

The signal model in vector notation is

$$\mathbf{r}_k = \sqrt{P_s}\mathbf{c}s_k + \mathbf{z}_k, \tag{2}$$

where $\mathbf{r}_k = [r_{k,1}, \ldots, r_{k,L}]^T$, $\mathbf{c} = [c_1, \ldots, c_L]^T$, $\mathbf{z}_k = [z_{k,1}, \ldots, z_{k,L}]^T$ and the superscript T denotes vector transposition.

The channel gains c_ℓ are assumed to be independent and identically distributed (i.i.d.), zero-mean, circularly symmetric, complex Gaussian random variables (Rayleigh fading), with variance $\Omega_\ell/2$ per dimension. The correlated noise term $\mathbf{z}_k$ is the aggregate of an interference source and AWGN and it is assumed to be complex-valued, zero-mean, circularly symmetric, and governed by a Gaussian distribution with covariance matrix $\mathbf{R}_z = E[\mathbf{z}_k\mathbf{z}_k^H]$. For a single interference source and AWGN, the covariance matrix can be expressed as

$$\mathbf{R}_z = E\left[\mathbf{z}_k\mathbf{z}_k^H\right] = P_I\mathbf{c}_I\mathbf{c}_I^H + \mathrm{diag}\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_L^2\right) \tag{3}$$

where P_I is the interference power, $\mathbf{c}_I$ is the interference channel vector, the superscript H denotes the Hermitian transpose and $(\sigma_1^2, \sigma_2^2, \ldots, \sigma_L^2)$ is the power profile of the AWGN.

Consider a sequence of K symbols running from time $k - (K-1)$ to k. Assume the channel is static over the duration of this sequence. Using vector notation

$$\underline{\mathbf{r}}_k = \sqrt{P_s}\mathbf{H}\mathbf{s}_k + \underline{\mathbf{z}}_k \tag{4}$$

where $\underline{\mathbf{r}}_k = [\mathbf{r}_{k-(K-1)}^T, \ldots, \mathbf{r}_k^T]^T$, $\mathbf{s}_k$, and $\underline{\mathbf{z}}_k$ are vectors defined similar to $\underline{\mathbf{r}}_k$, and $\mathbf{H} = \mathbf{I}_K \otimes \mathbf{c}$ is the channel matrix for the signal of interest, where $\otimes$ denotes the Kronecker product and $\mathbf{I}_K$ is the identity matrix of rank K.

III. Decision Statistic

We formulate the decision statistic for the symbol sequence $\mathbf{s}_k = [s_{k-(K-1)}, \ldots, s_k]^T$ based on an observation interval of length K as embodied by the vector $\underline{\mathbf{r}}_k$. The maximum-likelihood detector for the sequence $\mathbf{s}_k$ is given by

$$\widehat{\mathbf{s}}_k = \arg\max_{\mathbf{s}_k} p\left(\underline{\mathbf{r}}_k|\mathbf{s}_k\right) \tag{5}$$

where $p(\underline{\mathbf{r}}_k|\mathbf{s}_k)$ is the likelihood of the observed data $\underline{\mathbf{r}}_k$ given the transmitted symbol sequence $\mathbf{s}_k$. Under the Gaussian assumption for the aggregate of interference and noise, the observation $\underline{\mathbf{r}}_k$ conditioned on the transmitted sequence $\mathbf{s}_k$ and on the channel $\mathbf{c}$ has a multivariate Gaussian distribution. The conditional probability $p(\underline{\mathbf{r}}_k|\mathbf{s}_k)$ can then be expressed as

$$\begin{aligned} p\left(\underline{\mathbf{r}}_k|\mathbf{s}_k, \mathbf{c}\right) =& \pi^{-KL}\left|\mathbf{R}_z\right|^{-K} \\ &\cdot \exp\left\{-\sum_{i=0}^{K-1}\left(\mathbf{r}_{k-i} - \sqrt{P_s}\mathbf{c}s_{k-i}\right)^H\right. \\ &\left.\cdot \mathbf{R}_z^{-1}\left(\mathbf{r}_{k-i} - \sqrt{P_s}\mathbf{c}s_{k-i}\right)\right\}. \end{aligned} \tag{6}$$

Diagonalize the interference-plus-noise covariance matrix $\mathbf{R}_z$ as $\mathbf{R}_z = \mathbf{U}_z\mathbf{\Lambda}_z\mathbf{U}_z^H$, where $\mathbf{\Lambda}_z = \mathrm{diag}\left(\lambda_1, \ldots, \lambda_L\right)$, $\lambda_1, \ldots, \lambda_L$ are the eigenvalues of $\mathbf{R}_z$, and $\mathbf{U}_z$ is a unitary matrix whose columns are the eigenvectors of $\mathbf{R}_z$. It follows that (6) can be written as

$$\begin{aligned} p\left(\underline{\mathbf{r}}_k|\mathbf{s}_k, \mathbf{g}\right) =& \pi^{-KL}\left|\mathbf{R}_z\right|^{-K} \\ &\cdot \exp\left\{-\sum_{i=0}^{K-1}\left(\mathbf{x}_{k-i} - \sqrt{P_s}\mathbf{g}s_{k-i}\right)^H\right. \\ &\left.\cdot \mathbf{\Lambda}_z^{-1}\left(\mathbf{x}_{k-i} - \sqrt{P_s}\mathbf{g}s_{k-i}\right)\right\} \end{aligned} \tag{7}$$

where $\mathbf{x}_{k-i} = \mathbf{U}_z^H\mathbf{r}_{k-i}$ is the whitened received signal vector and $\mathbf{g} = \mathbf{U}_z^H\mathbf{c}$ is the modified channel vector. Note that since $\mathbf{U}_z$ is unitary, the modified channel vector $\mathbf{g}$ has the same distribution as the original channel vector $\mathbf{c}$. Let the components of the modified channel vector $\mathbf{g}$ be expressed as $g_\ell = \alpha_\ell e^{j\phi_\ell}$, $\ell = 1, \ldots, L$. Likewise, let the ℓth component of $\mathbf{x}_{k-i}$ be $x_{k-i,\ell}$. Expanding the exponent in (7) and grouping terms that do not depend on $\mathbf{g}$ or s_{k-i}, we obtain

$$\begin{aligned} p\left(\underline{\mathbf{r}}_k|\mathbf{s}_k, \mathbf{g}\right) =& \pi^{-KL}\left|\mathbf{R}_z\right|^{-K}\exp\left\{-C_0\right\} \\ &\cdot \prod_{\ell=1}^{L}\exp\left\{-KP_s\lambda_\ell^{-1}\alpha_\ell^2 + 2\sqrt{P_s}\lambda_\ell^{-1}\left|y_\ell\left(\mathbf{s}_k\right)\right|\right. \\ &\left.\cdot \alpha_\ell\cos\left(\phi_\ell - \theta_\ell\left(\mathbf{s}_k\right)\right)\right\} \end{aligned} \tag{8}$$

where

$$C_0 = \sum_{i=0}^{K-1}\sum_{\ell=1}^{L}\lambda_\ell^{-1}\left|x_{k-i,\ell}\right|^2 \tag{9}$$

$$y_\ell\left(\mathbf{s}_k\right) = \left(\sum_{i=0}^{K-1}x_{k-i,\ell}s_{k-i}^*\right) = \left|y_\ell\left(\mathbf{s}_k\right)\right|e^{j\theta_\ell(\mathbf{s}_k)}. \tag{10}$$

Note that $y_\ell(\mathbf{s}_k)$ is a function of both the transmitted sequence $\mathbf{s}_k$ and the observed sequence $\underline{\mathbf{r}}_k$.

Recalling that the components of the modified channel vector $\mathbf{g}$ have the same distribution as the components of the channel vector $\mathbf{c}$, it follows that α_ℓ is Rayleigh with $E[\alpha_\ell^2] = \Omega_\ell$, and ϕ_ℓ is uniformly distributed in the interval [0, 2π). To average the conditional distribution $p(\underline{\mathbf{r}}_k|\mathbf{s}_k, \mathbf{g})$ over the modified channel $\mathbf{g}$, we need to evaluate the integral

$$\begin{aligned} p(\underline{\mathbf{r}}_k|\mathbf{s}_k) =& \int p(\underline{\mathbf{r}}_k|\mathbf{s}_k, \mathbf{g})\, p_{\mathbf{g}}(\mathbf{g}) d\mathbf{g} \\ =& \pi^{-KL} |\mathbf{R}_z|^{-K} \exp\{-C_0\} \\ &\cdot \prod_{\ell=1}^{L} \left\{ \frac{2}{\Omega_\ell} \int_0^\infty \exp\left\{ -\left(K P_s \lambda_\ell^{-1} + \frac{1}{\Omega_\ell} \right) \alpha_\ell^2 \right\} \right. \\ &\cdot \alpha_\ell d\alpha_\ell \frac{1}{2\pi} \int_0^{2\pi} \exp\left\{ 2\sqrt{P_s} \lambda_\ell^{-1} |y_\ell(\mathbf{s}_k)| \alpha_\ell \right. \\ &\left.\left. \cdot \cos(\phi_\ell - \theta_\ell) \right\} d\phi_\ell \right\} \end{aligned} \tag{11}$$

where $p_{\mathbf{g}}(\mathbf{g})$ is the probability density function of $\mathbf{g}$. Averaging over the uniform distribution of ϕ_ℓ, we obtain

$$\begin{aligned} p(\underline{\mathbf{r}}_k|\mathbf{s}_k) =& \pi^{-KL} |\mathbf{R}_z|^{-K} \exp\{-C_0\} \\ &\cdot \prod_{\ell=1}^{L} \frac{2}{\Omega_\ell} \int_0^\infty \exp\left\{ -\left(K P_s \lambda_\ell^{-1} + \frac{1}{\Omega_\ell} \right) \alpha_\ell^2 \right\} \\ &\cdot I_0\left(2\sqrt{P_s} \lambda_\ell^{-1} |y_\ell(\mathbf{s}_k)| \alpha_\ell \right) \alpha_\ell d\alpha_\ell \end{aligned} \tag{12}$$

where $I_0(x)$ is the zeroth-order modified Bessel function of the first kind. Now, averaging over the fading statistics of α_ℓ, we obtain

$$\begin{aligned} p(\underline{\mathbf{r}}_k|\mathbf{s}_k) =& \pi^{-KL} |\mathbf{R}_z|^{-K} \exp\{-C_0\} \left(\prod_{\ell=1}^{L} \frac{\lambda_\ell}{K P_s \Omega_\ell + \lambda_\ell} \right) \\ &\cdot \exp\left\{ P_s \sum_{\ell=1}^{L} \frac{\Omega_\ell |y_\ell(\mathbf{s}_k)|^2}{\lambda_\ell (K P_s \Omega_\ell + \lambda_\ell)} \right\}. \end{aligned} \tag{13}$$

In (13), only the argument of the exponential function is dependent on the transmitted sequence $\mathbf{s}_k$, since only the terms $y_\ell(\mathbf{s}_k)$ are functions of $\mathbf{s}_k$. Due to the monotonicity of the exponential function, maximizing $p(\underline{\mathbf{r}}_k|\mathbf{s}_k)$ with respect to $\mathbf{s}_k$ is equivalent to maximizing the following decision statistic:

$$\eta(\mathbf{s}_k) = \sum_{\ell=1}^{L} \frac{\Omega_\ell |y_\ell(\mathbf{s}_k)|^2}{\lambda_\ell (K P_s \Omega_\ell + \lambda_\ell)}. \tag{14}$$

From the previous relation and (10), it follows that the optimum MSDD for multiple-channel branches and in the presence of interference is a weighted sum of correlations of whitened observations and hypothesis symbols. Note that this decision statistic does not require knowledge of the signal channel vector.

The decision statistic is ambiguous with respect to an arbitrary phase θ'. Indeed, let $\mathbf{s}'_k = e^{j\theta'} \mathbf{s}_k$, then

$$\begin{aligned} |y_\ell(\mathbf{s}'_k)| =& \left| \sum_{i=0}^{K-1} x_{k-i,\ell} \left(e^{j\theta'} s_{k-i} \right)^* \right| \\ =& |y_\ell(\mathbf{s}_k)|. \end{aligned} \tag{15}$$

Differential encoding at the transmitter is required to resolve this ambiguity.

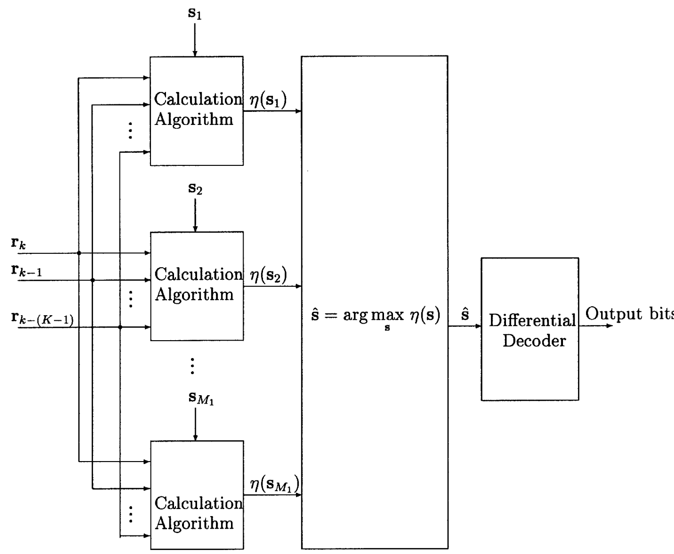


Fig. 1. Diagram of MSDD. For M-DPSK with M symbols, $M_1 = M^{K-1}$.

The decision statistic in (14) provides MSDD for an M-DPSK sequence transmitted over multiple, independent fading channels in the presence of correlated Gaussian noise.

From (5), the MSDD decision rule is

$$\widehat{\mathbf{s}}_k = \arg\max_{\mathbf{s}_k} \eta(\mathbf{s}_k). \tag{16}$$

For M-DPSK symbols, the relative complexity of this operation is proportional to M^{K-1}. A diagram of the MSDD receiver is shown in Fig. 1.

Some special cases provide insights into the operation of MSDD. For a channel with a flat gain profile $\Omega_\ell = 1$ and a flat AWGN profile $\sigma_\ell^2 = \sigma^2$ for $\ell = 1, 2, \ldots, L$, (14) can be expressed as

$$\eta(\mathbf{s}_k) = \sum_{\ell=1}^{L} \frac{|y_\ell(\mathbf{s}_k)|^2}{\lambda_\ell (K P_s + \lambda_\ell)}. \tag{17}$$

We further specialize (17) to the following special cases.

A. No Interference

For this case, $P_I = 0$, the noise covariance matrix $\mathbf{R}_z = \sigma^2 \mathbf{I}_L$, eigenvalues $\lambda_\ell = \sigma^2$, $\ell = 1, 2, \ldots, L$, and $\mathbf{U}_z = \mathbf{I}_L$. Then (10) simplifies to

$$y_\ell(\mathbf{s}_k) = \sum_{i=0}^{K-1} r_{k-i,\ell} s_{k-i}^*. \tag{18}$$

The decision statistic in (17) becomes

$$\eta(\mathbf{s}_k) = \frac{1}{\sigma^2 (K P_s + \sigma^2)} \sum_{\ell=1}^{L} |y_\ell(\mathbf{s}_k)|^2. \tag{19}$$

Since the term outside the sum is independent of $\mathbf{s}_k$, the above decision statistic is equivalent to

$$\eta(\mathbf{s}_k) = \sum_{\ell=1}^{L} |y_\ell(\mathbf{s}_k)|^2 = \sum_{\ell=1}^{L} \left| \sum_{i=0}^{K-1} r_{k-i,\ell} s_{k-i}^* \right|^2. \tag{20}$$

This decision statistic is the same as that in [7, eq. (8)]. Indeed, (14) is the generalization of [7, eq. (8)] to MSDD in the presence of interference.

B. Interference $\gg$ Noise

For single interferer and a uniform AWGN power profile, the eigenvalues of the interference-plus-noise covariance matrix (3) are [13] $\lambda_1 = P_I \sum_{\ell=1}^{L} |c_{I,\ell}|^2 + \sigma^2$, $\lambda_2 = \cdots = \lambda_L = \sigma^2$. For a high interference-to-noise ratio, $\lambda_1 \gg \lambda_\ell$, $\ell \neq 1$. It follows that the decision statistic in (17) can be approximated by the expression

$$\eta(\mathbf{s}_k) \approx \sum_{\ell=2}^{L} \frac{|y_\ell(\mathbf{s}_k)|^2}{\sigma^2 (KP_s + \sigma^2)}. \tag{21}$$

The interpretation of this result is that for a strong interference source, the decision statistic is similar to that of MSDD without interference and one fewer degree of freedom. This result will be further demonstrated in the ensuing error probability analysis.

IV. Error Probability Analysis

An exact expression for the BEP for differential detection can be obtained only for differential binary PSK (DPSK) modulation and the special case of $K = 2$ symbols. The exact error analysis is intractable for the general case of MSDD with M-DPSK modulation over diversity channels and in the presence of interference. The alternative approach is to obtain an analytical approximate upper bound. In this section, we first derive an exact expression for the PEP under the Gaussian assumption for the aggregate interference plus noise. Then, using this expression, we derive the union bound of the BEP. From the union bound, an approximate upper bound is derived. The approximate upper bound consists of relatively simple algebraic expressions. Even simpler expressions are obtained for large SNR and small signal-to-interference ratio (SIR). In the numerical results section, it is shown that the approximate upper bound is very close to the BEP obtained by simulation.

A. PEP Analysis

In the derivation of the PEP, we assume a uniform flat power profile for the channel of the desired signal, $\Omega_\ell = 1$, and a flat AWGN profile with $\sigma_\ell^2 = \sigma^2$ for $\ell = 1, 2, \ldots, L$. The PEP is developed for correlated noise characterized by the covariance matrix in (3).

In general, the interference source is subject to effects of the fading channel (similar to the desired source). It follows that analysis using the covariance matrix $\mathbf{R}_z$ in (3) is conditional on the interference random channel $\mathbf{c}_I$. Results obtained from such analysis need to be averaged over the distribution of $\mathbf{c}_I$. Fortunately, this complication can be avoided by recognizing that when the detector acts to suppress the interference, there is only a small penalty in using in the analysis the average value of the interference power $P_I E[\mathbf{c}_I^H \mathbf{c}_I]$ in lieu of the instantaneous power $P_I \mathbf{c}_I^H \mathbf{c}_I$ (see [12]). Assuming that the interference channel $\mathbf{c}_I$ is complex-valued, zero mean, and with variance $\Omega_{I,\ell}/2 = 1/2$ per dimension, it follows that the average eigenvalues of $\mathbf{R}_z$ are

$$\lambda_1 = P_I E\left[\mathbf{c}_I^H \mathbf{c}_I\right] + \sigma^2 = LP_I + \sigma^2 \tag{22}$$

and $\lambda_\ell = \sigma^2$ for $\ell = 2, 3, \ldots, L$.

Let $\mathbf{s}_k$ and $\mathbf{s}'_k$ denote two sequences, each containing K M-DPSK symbols. The PEP that $\mathbf{s}_k$ is transmitted but $\mathbf{s}'_k$ is detected ($\mathbf{s}'_k \neq s_0 \mathbf{s}_k$, where s_0 is an arbitrary M-PSK symbol) is denoted as $P(\mathbf{s}_k \rightarrow \mathbf{s}'_k)$. An error event occurs when $\eta(\mathbf{s}_k) < \eta(\mathbf{s}'_k)$. Define the random variable D

$$D = \eta(\mathbf{s}_k) - \eta(\mathbf{s}'_k). \tag{23}$$

Note that D is random due to both random interference plus noise and the random channel. We seek to evaluate the probability that $D < 0$.

Using steps similar to [18, App. B], it can be shown that

$$\begin{aligned} P(\mathbf{s}_k \rightarrow \mathbf{s}'_k) &= -\frac{1}{2\pi j} \int_{-\infty + j\varepsilon}^{\infty + j\varepsilon} \frac{\Phi_D(j\omega)}{\omega} d\omega \\ &= -\sum_{\mathrm{Im}(\omega_\ell) > 0} \mathrm{Res}\left[\frac{\Phi_D(j\omega)}{\omega}; \omega_\ell\right] \end{aligned} \tag{24}$$

where ε is a small positive number, $\Phi_D(j\omega)$ is the characteristic function of D, $\mathrm{Res}[(\Phi_D(j\omega))/\omega; \omega_\ell]$ denotes the residue of $(\Phi_D(j\omega))/\omega$ at pole ω_ℓ, and the summation is taken over the poles in the upper half of the complex plane.

In Appendix A, the following expression is derived for the characteristic function of the random variable D:

$$\begin{aligned} \Phi_D(j\omega) = &\frac{1}{(1 - j\mu_1\omega)} \frac{1}{(1 - j\mu_2\omega)} \\ &\cdot \frac{1}{(1 - j\mu_3\omega)^{L-1}} \frac{1}{(1 - j\mu_4\omega)^{L-1}} \end{aligned} \tag{25}$$

where

$$\mu_i = \begin{cases} \frac{1}{2} b_1^2 \left[\zeta P_s \pm \sqrt{\zeta^2 P_s^2 + 4(KP_s + \lambda_1)\zeta\lambda_1}\right], & i = 1, 2 \\ \frac{1}{2} b_2^2 \left[\zeta P_s \pm \sqrt{\zeta^2 P_s^2 + 4(KP_s + \lambda_2)\zeta\lambda_2}\right], & i = 3, 4 \end{cases} \tag{26}$$

$$b_\ell = \sqrt{\frac{1}{\lambda_\ell (KP_s + \lambda_\ell)}}, \qquad \ell = 1, 2 \tag{27}$$

$$\zeta = K^2 - |v(\mathbf{s}_k, \mathbf{s}'_k)|^2 \tag{28}$$

and $v(\mathbf{s}_k, \mathbf{s}'_k) = \mathbf{s}'^H_k \mathbf{s}_k$ is the correlation coefficient between the transmitted sequence $\mathbf{s}_k$ and the detected sequence $\mathbf{s}'_k$. Note that $0 \leq v(\mathbf{s}_k, \mathbf{s}'_k) < K$ for $\mathbf{s}'_k \neq s_0 \mathbf{s}_k$, where s_0 is an arbitrary M-PSK symbol.

For $\Phi_D(j\omega)$, since $\mu_1, \mu_3 > 0$, and $\mu_2, \mu_4 < 0$, only the poles $-j/\mu_2$ and $-j/\mu_4$ are in the upper half of the complex plane. Substituting (25) into (24) and carrying out the calculation of residues [19], we obtain the PEP as

$$\begin{aligned} P(\mathbf{s}_k \rightarrow \mathbf{s}'_k) = &-\frac{(\mu_2)^{2L-1}}{(\mu_1 - \mu_2)(\mu_3 - \mu_2)^{L-1}(\mu_4 - \mu_2)^{L-1}} \\ &- \frac{(-1)^L \mu_4^{L-1}}{\mu_3^{L-1}} \frac{1}{(L-2)!} \sum_{k=0}^{L-2} \frac{(L-2+k)}{k!} \\ &\times \left[1 - \frac{\mu_1}{(\mu_1 - \mu_2)} \frac{\mu_1^{L-1-k}}{(\mu_1 - \mu_4)^{L-1-k}} \right. \\ &\left. + \frac{\mu_2}{(\mu_1 - \mu_2)} \frac{\mu_2^{L-1-k}}{(\mu_2 - \mu_4)^{L-1-k}}\right] \\ &\cdot \frac{\mu_3^{L-1+k}}{(\mu_3 - \mu_4)^{L-1+k}}. \end{aligned} \tag{29}$$

The former expression is the exact PEP of MSDD with diversity branches and a rank one interference source. The PEP is a function of the transmitted sequence $\mathbf{s}_k$ and the detected sequence $\mathbf{s}'_k$. Note that the expression in (29) already incorporates statistical information on both the channel and interference. This form of the PEP is quite complicated and does not afford much insight. It is of interest to obtain simpler expressions for special cases. In the ensuing analysis, the symbol SNR is denoted as $\gamma = P_s/\sigma^2$ and the SIR is P_s/P_I.

1) No Interference: For this case $P_I = 0$, $\lambda_1 = \lambda_2 = \sigma^2$. From (26) we have

$$\mu_i = \begin{cases} \frac{1}{2P_s}\frac{\gamma^2}{(K\gamma+1)}\left[\zeta + \sqrt{\zeta^2 + 4\left(K+\frac{1}{\gamma}\right)\zeta\frac{1}{\gamma}}\right], & i = 1,3 \\ \frac{1}{2P_s}\frac{\gamma^2}{(K\gamma+1)}\left[\zeta - \sqrt{\zeta^2 + 4\left(K+\frac{1}{\gamma}\right)\zeta\frac{1}{\gamma}}\right], & i = 2,4. \end{cases} \tag{30}$$

Substituting (30) in (24), we can get the PEP as

$$P(\mathbf{s}_k \to \mathbf{s}'_k) = \frac{(-\mu_2)^L}{(\mu_1-\mu_2)^L (L-1)!}\left[\sum_{\ell=0}^{L-1}\frac{(L-1+\ell)!}{\ell!}\frac{\mu_1^\ell}{(\mu_1-\mu_2)^\ell}\right]. \tag{31}$$

For all the cases we tried, (31) yielded the same numerical results as the PEP developed in [7]. However, (31) has the advantage that it provides the PEP in closed form without the need of integration.

The case of no interference can be further simplified for large SNR $\gamma \gg 1$. In this case, (30) simplifies to

$$\mu_i \approx \begin{cases} \frac{\gamma\zeta}{KP_s}, & \text{i=1,3} \\ -\frac{1}{P_s}, & \text{i=2,4} \end{cases}. \tag{32}$$

Substituting these results in (31), noticing that $\mu_1 \gg \mu_2$, and using [20, eq. (0.151.1)], we have

$$P(\mathbf{s}_k \to \mathbf{s}'_k) \approx \binom{2L-1}{L}\frac{1}{\left(\frac{\zeta}{K}\gamma\right)^L}. \tag{33}$$

This expression clearly exhibits the L-order diversity of the system.

2) SIR $\ll 1$, SNR $\gg 1$: This special case is of theoretical interest since it can show the ability of MSDD to suppress a large cochannel interferer. By assumption, $P_I \gg P_s \gg \sigma^2$, therefore, $\lambda_1 \gg P_s \gg \sigma^2$. After some manipulations, we have $\mu_{1,2} \approx \pm\sqrt{\zeta}/\lambda_1$, $\mu_3 \approx \zeta/(K\sigma^2)$, and $\mu_4 \approx -1/P_s$. Substituting these approximate values into (29) and keeping only the dominant term, we obtain

$$P(\mathbf{s}_k \to \mathbf{s}'_k) \approx \binom{2(L-1)-1}{(L-1)}\frac{1}{\left(\frac{\zeta}{K}\gamma\right)^{L-1}}. \tag{34}$$

Comparing (34) with (33), we can see that the PEP for systems with diversity L and a large interference is equal to the PEP for systems with diversity $(L-1)$ and without interference. This result is well known for interference suppression using optimum combining. This analysis proves that the loss of degree of freedom due to interference suppression carries over to MSDD over a diversity Rayleigh channel.

B. BEP Approximate Upper Bound

The sequence $\mathbf{s}_k$ of M-DPSK symbols corresponds to $(K-1)\log_2 M$ information bits (with differential encoding, the first symbol is known). Let $\mathbf{u}_k$ be the sequence of $(K-1)\log_2 M$ information bits encoded as $\mathbf{s}_k$ and let $\mathbf{u}'_k$ be the sequence of information bits which results from the detection of $\mathbf{s}'_k$. The pairwise BEP associated with transmitting a sequence $\mathbf{u}_k$ and detecting another sequence $\mathbf{u}'_k$ is given by

$$P_b(\mathbf{s}_k \to \mathbf{s}'_k) = \frac{1}{(K-1)\log_2 M}h(\mathbf{u}_k, \mathbf{u}'_k)P(\mathbf{s}_k \to \mathbf{s}'_k) \tag{35}$$

where $h(\mathbf{u}_k, \mathbf{u}'_k)$ denotes the Hamming distance between $\mathbf{u}_k$ and $\mathbf{u}'_k$.

The BEP that $\mathbf{s}_k$ is transmitted, but an error sequence (any error sequence) is detected, and is upper bounded by the union of all pairwise bit-error events. Since $\mathbf{s}_k$ can be any input sequence (e.g., the null sequence $\mathbf{s}_k = [1,0,0,\ldots,0]^T$), we drop the dependency on $\mathbf{s}_k$ from the notation. The union bound on the BEP can then be written as

$$\begin{aligned} P_b &\le \sum_{\mathbf{u}'_k \ne \mathbf{u}_k} P_b(\mathbf{s}_k \to \mathbf{s}'_k) \\ &= \frac{1}{(K-1)\log_2 M}\sum_{\mathbf{u}'_k \ne \mathbf{u}_k} h(\mathbf{u}_k, \mathbf{u}'_k)P(\mathbf{s}_k \to \mathbf{s}'_k) \end{aligned} \tag{36}$$

where the summation is taken over all the sequences $\mathbf{u}'_k$, which are different from the transmitted sequence of information bits $\mathbf{u}_k$.

Direct application of (36) does not shed light on the mechanisms affecting MSDD performance. A clearer picture is obtained by developing an approximation to the union bound. Note that the union bound in (36) is a function of the PEPs, which, in turn, are determined by μ_1, μ_2, μ_3, and μ_4 [see (29)]. From (26), μ_1, μ_2, μ_3, and μ_4 are functions of the quantity $|v(\mathbf{s}_k, \mathbf{s}'_k)|^2$ through the relation $\zeta = K^2 - |v(\mathbf{s}_k, \mathbf{s}'_k)|^2$. In [1], it is shown that on the AWGN channel, for large SNR, the dominant terms in the BEP occur for sequences for which the quantity $|v(\mathbf{s}_k, \mathbf{s}'_k)|^2$ is maximum. Carrying over the same approach to the fading channel, keeping only the dominant terms and noticing that $P(\mathbf{s}_k \to \mathbf{s}'_k)$ is constant if $|v(\mathbf{s}_k, \mathbf{s}'_k)|$ is constant, we obtain the following approximation to the union bound:

$$\begin{aligned} A = \frac{1}{(K-1)\log_2 M}&\left[\sum_{\substack{\mathbf{u}'_k \ne \mathbf{u}_k \\ |v(\mathbf{s}_k,\mathbf{s}'_k)| = |v(\mathbf{s}_k,\mathbf{s}'_k)|_{\max}}} h(\mathbf{u}_k, \mathbf{u}'_k)\right] \\ &\times \left[P(\mathbf{s}_k \to \mathbf{s}'_k)\,|_{|v(\mathbf{s}_k,\mathbf{s}'_k)| = |v(\mathbf{s}_k,\mathbf{s}'_k)|_{\max}}\right]. \end{aligned} \tag{37}$$

The maximum value of $|v(\mathbf{s}_k, \mathbf{s}'_k)|$ for $\mathbf{s}'_k \ne s_0\mathbf{s}_k$ was shown in [1, eq. (38)] to be

$$|v(\mathbf{s}_k, \mathbf{s}'_k)|_{\max} = \sqrt{(K-1)^2 + 2(K-1)\left(1 - 2\sin^2\frac{\pi}{M}\right) + 1}. \tag{38}$$

Also from [1, App. B], for sequences such that $|v(\mathbf{s}_k, \mathbf{s}'_k)| = |v(\mathbf{s}_k, \mathbf{s}'_k)|_{\max}$, the accumulated Hamming distances are

$$\sum_{\substack{\mathbf{u}'_k \ne \mathbf{u}_k \\ |v(\mathbf{s}_k,\mathbf{s}'_k)| = |v(\mathbf{s}_k,\mathbf{s}'_k)|_{\max}}} h(\mathbf{u}_k, \mathbf{u}'_k) = \begin{cases} 1, & K = 2 \\ 2(K-1), & K > 2 \end{cases} \tag{39}$$

TABLE I
COMPARISON OF APPROXIMATE UPPER BOUND FROM (37) AND SIMULATED BEP FOR DPSK, $L = 2$ BRANCHES, SIR = 3 dB, SNR = 10 dB

K	2	7	12	40
Appr. upper bound ($\times 10^{-2}$)	3.13	2.74	2.41	2.13
Simulated BEP ($\times 10^{-2}$)	3.15	2.29	2.14	1.71

TABLE II
COMPARISON OF APPROXIMATE UPPER BOUND AND SIMULATED BEP FOR DQPSK, $L = 4$ BRANCHES, SIR = -6 dB, SNR = 6 dB

K	2	7	12	40
Appr. upper bound ($\times 10^{-3}$)	7.67	4.13	3.43	2.87
Simulated BEP ($\times 10^{-3}$)	7.85	3.77	3.87	2.89

for binary modulation, $M = 2$ and

$$\sum_{\substack{\mathbf{u}'_k \neq \mathbf{u}_k \\ |v(\mathbf{s}_k, \mathbf{s}'_k)| = |v(\mathbf{s}_k, \mathbf{s}'_k)|_{\max}}} h(\mathbf{u}_k, \mathbf{u}'_k) = \begin{cases} 2, & K = 2 \\ 4(K-1), & K > 2 \end{cases} \tag{40}$$

for multilevel modulation, $M \geqslant 4$.

Strictly speaking, (37) is not an upper bound of the BEP. Numerical results (such as those in Tables I and II), however, show that it is close to the BEP obtained by simulation. Therefore, we will use (37) to study the performance of MSDD in the presence of interference.

Next, we evaluate the approximate upper bound for DPSK and M-DPSK ($M \geq 4$) modulations.

1) DPSK ($M = 2$): For this case, from (38) we have

$$|v(\mathbf{s}_k, \mathbf{s}'_k)|_{\max} = K - 2. \tag{41}$$

For conventional differential detection, the observation interval is $K = 2$ symbols, $|v(\mathbf{s}_k, \mathbf{s}'_k)|_{\max} = 0$. In this case, there is only one error sequence, therefore, the PEP is also the BEP

$$P_b = P(\mathbf{s}_k \to \mathbf{s}'_k)\big|_{|v(\mathbf{s}_k, \mathbf{s}'_k)|=0}. \tag{42}$$

Substituting $P(\mathbf{s}_k \to \mathbf{s}'_k)$ from (31) into (42), we obtain the exact BEP for DPSK over L diversity fading channels without interference. For high SNR $\gg 1$, using (33), we get

$$P_{b,\mathrm{DPSK}} \approx \binom{2L-1}{L} \frac{1}{(2\gamma)^L}. \tag{43}$$

This expression is the same as the one in [18, eq. (14-4-48)] and it demonstrates that familiar expressions for differential detection can be obtained as a special case of the general case treated in this paper.

For a longer observation interval $K > 2$, substitute (41) and (39) into (37) to obtain the approximate BEP upper bound for DPSK as

$$A_{\mathrm{DPSK}} = 2\, P(\mathbf{s}_k \to \mathbf{s}'_k)\big|_{|v(\mathbf{s}_k, \mathbf{s}'_k)|=K-2}. \tag{44}$$

Substituting (34) in (44), we obtain simplified expression for $K > 2$, SIR $\ll 1$, and SNR $\gg 1$ as

$$A_{\mathrm{DPSK}} = 2 \binom{2(L-1)-1}{L-1} \frac{1}{4^{L-1}\gamma^{L-1}\left(1-\frac{1}{K}\right)^{L-1}}. \tag{45}$$

2) M-DPSK ($M \geqslant 4$): For M-DPSK, substituting (40) and (38) into (37), we obtain the following approximate upper bound:

$$A_{\mathrm{M-DPSK}} = \frac{2}{\log_2 M}\, P(\mathbf{s}_k \to \mathbf{s}'_k)\big|_{|v(\mathbf{s}_k, \mathbf{s}'_k)|=2|\cos \pi/M|} \tag{46}$$

for $K = 2$ symbols. See (47) at the bottom of the page for observation intervals of length $K > 2$.

Substituting (34) in (46) and (47) we obtain a simplified expression for SIR $\ll 1$ and SNR $\gg 1$ as

$$A_{M-\mathrm{DPSK}} = \frac{2}{\log_2 M} \binom{2(L-1)-1}{L-1} \cdot \frac{1}{2^{L-1}\gamma^{L-1}\sin^{2(L-1)}\left(\frac{\pi}{M}\right)} \tag{48}$$

for $K = 2$, and

$$A_{M-\mathrm{DPSK}} = \frac{4}{\log_2 M} \binom{2(L-1)-1}{L-1} \cdot \frac{1}{4^{L-1}\gamma^{L-1}\sin^{2(L-1)}\left(\frac{\pi}{M}\right)\left(1-\frac{1}{K}\right)^{L-1}} \tag{49}$$

for $K > 2$.

C. Comparison With Optimum Combining

It is of interest to compare the performance of MSDD, which does not require a coherent reference and knowledge of the channel, with that of optimum combining, which requires both a coherent phase reference and channel information. With both methods, it is assumed that transmitted symbols are differentially encoded. In [21], it is shown that for SIR $\ll 1$ and SNR $\gg 1$, the BEP for optimum combining with differential encoding at the transmitter is approximated by the expressions

$$P_{b,\mathrm{DPSK}} \approx 2 \binom{2(L-1)-1}{(L-1)-1} \frac{1}{4^{L-1}\gamma^{L-1}} \tag{50}$$

for DPSK, and by

$$P_{b,M-\mathrm{DPSK}} \approx \frac{4}{\log_2 M} \binom{2(L-1)-1}{(L-1)-1} \cdot \frac{1}{4^{L-1}\gamma^{L-1}\sin^{2(L-1)}\left(\frac{\pi}{M}\right)} \tag{51}$$

for M-DPSK.

$$A_{\mathrm{M-DPSK}} = \frac{4}{\log_2 M}\, P(\mathbf{s}_k \to \mathbf{s}'_k)\big|_{|v(\mathbf{s}_k, \mathbf{s}'_k)|=\sqrt{(K-1)^2+2(K-1)(1-2\sin^2(\pi/M))+1}} \tag{47}$$

The ratio of the BEP for optimum combining and the approximate upper bound of MSDD ($K > 2$) is given by the ratios of (50) to (45) and (51) to (49), respectively

$$\frac{P_{b,\text{DPSK}}}{A_{\text{DPSK}}} = \frac{P_{b,M-\text{DPSK}}}{A_{M-\text{DPSK}}} = \left(1 - \frac{1}{K}\right)^{L-1}. \tag{52}$$

This expression holds for the asymptotic case of SIR $\ll 1$ and SNR $\gg 1$. We conclude that for SIR $\ll 1$ and SNR $\gg 1$, when the observation interval of MSDD increases to infinity, i.e., $K \to \infty$, the performance of MSDD with noncoherent detection approaches that of optimum combining with differential encoding.

V. Numerical Results

Numerical results presented in this section include Monte Carlo simulation results and analysis results. In all cases, the channel branches and noise power profiles are assumed to be uniform, i.e., $\Omega_\ell = 1$ and $\sigma_\ell^2 = \sigma^2$ for $\ell = 1, 2, \ldots, L$. The bit SNR is $(P_s/\log_2(M))/\sigma^2$. For comparison purposes, we also provide BEP curves for optimum combining with differential encoding. Simulation results were generated based on the Gaussian assumption.

As mentioned in Section III, the complexity of MSDD for M-DPSK with a K-symbol observation interval increases with M^{K-1}. For large K, simulations are impractical. To overcome this difficulty, a practical suboptimal algorithm that uses decision feedback was implemented. The basic idea of the algorithm is to make symbol-by-symbol decisions rather than testing the full sequence of symbols simultaneously. The algorithm proceeds from symbol to symbol along the sequence of K symbols; at symbol i it maximizes a decision statistic, assuming that the other $(K-1)$ symbols have been detected and are known. Several iterations can be carried out to improve performance. The algorithm was implemented as the following procedure:

1) Initialization:
 a) Initialize iteration index $m = 0$.
 b) Initialize $\mathbf{s}_k^{(m)} \triangleq [1, s_{k-(K-2)}^{(m)}, s_{k-(K-3)}^{(m)}, \ldots, s_k^{(m)}]^T = [1, 0, 0, \ldots, 0]^T$.
 c) Initialize time index $i = K-2$.
2) Increase iteration index $m + 1 \to m$.
3) For $i = (K-2)$ to 0,
 Evaluate

$$s_{k-i}^{(m)} = \arg\max_{s_{k-i}^{(m)}} \eta \cdot \left(\left[1, s_{k-(K-2)}^{(m)}, \ldots, s_{k-i}^{(m)}, s_{k-i+1}^{(m-1)}, \ldots, s_k^{(m-1)}\right]^T \right).$$

End loop i.
4) If m is not equal to the required iteration number (which is determined empirically), go back to step 2.
5) Differentially decode $\mathbf{s}_k^{(m)}$ to get the final output.

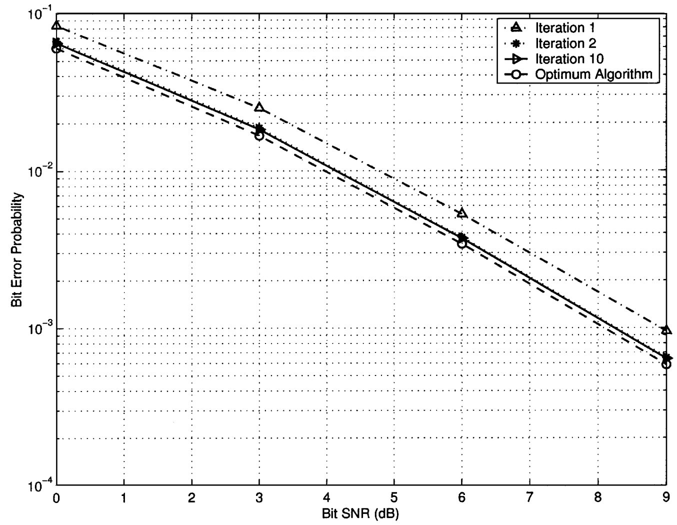


Fig. 2. Comparison of optimum algorithm and iterative decision feedback algorithm for $L = 4$ branches, DPSK modulation, SIR $= -6$ dB.

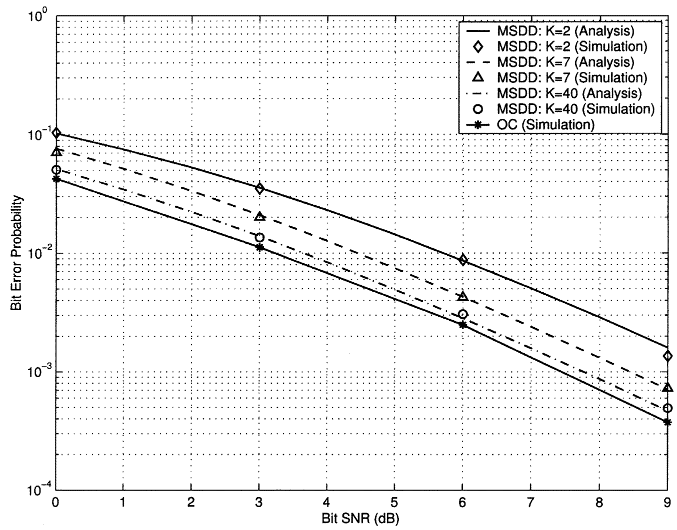


Fig. 3. BEP versus SNR for $L = 4$ branches, DPSK modulation, SIR $= -6$ dB.

To demonstrate the performance of this suboptimal algorithm, Fig. 2 compares the suboptimal and optimal [based on (16)] algorithms. The comparison is for the case of $L = 4$ diversity branches, DPSK modulation, and SIR $= -6$ dB. For an observation interval of $K = 12$ symbols, with just two iterations, the performance of the suboptimal algorithm is within just 0.2 dB of that of the optimum algorithm. In the results reported below, this suboptimal iterative decision feedback algorithm was applied to detection with observation interval $K = 40$.

All the ensuing figures are for $L = 4$ diversity branches. Fig. 3 shows the BEP versus SNR for DPSK at SIR $= -6$ dB. The diamonds, triangles and circles labeled "Simulation" represent simulation results, while curves labeled "Analysis" show analytical results as yielded by the approximate upper bounds (42) (for $K = 2$) and (44) (for $K > 2$). In all cases, PEPs were exact as computed by (29). The interference-plus-noise term was generated such that its covariance matrix followed (3).

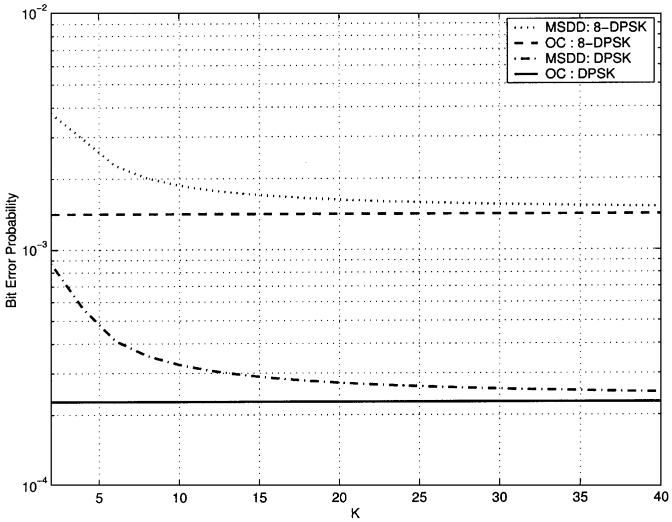


Fig. 4. BEP versus the number of symbols in the observation interval K for $L = 4$ branches, SIR $= -6$ dB, SNR $= 10$ dB.

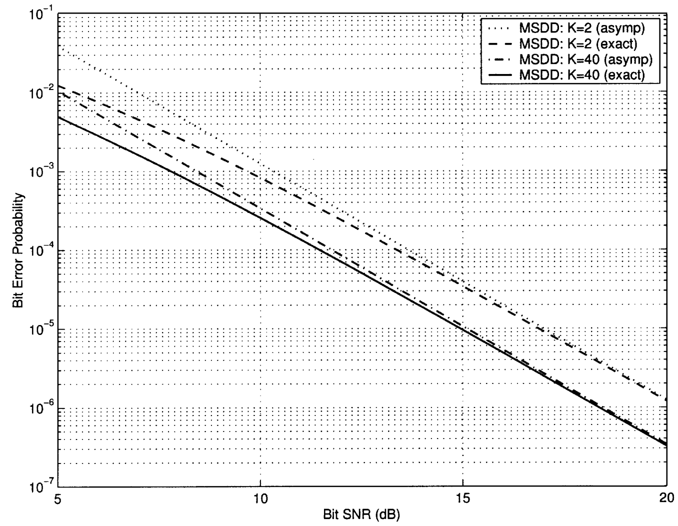


Fig. 6. Comparison of asymptotic results and exact results for $L = 4$ branches, DQPSK modulation, SIR $= -6$ dB.

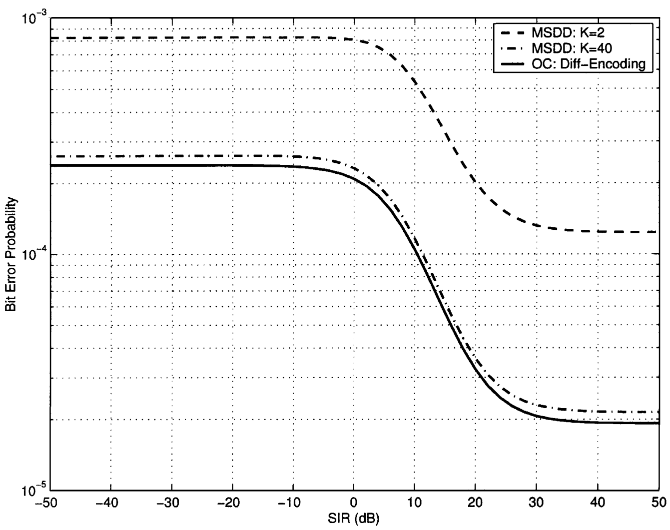


Fig. 5. BEP versus SIR for $L = 4$ branches, DQPSK modulation, bit SNR $=$ 10 dB.

The optimum combining curve was generated by simulation. It can be observed that analysis results are very close to simulation results. It is also observed that the performance of MSDD approaches that of optimum combining with differential encoding as the observation interval K increases. For example, at BEP $= 2 \times 10^{-3}$, when $K = 2$, the SNR difference between MSDD and optimum combining is about 2.2 dB. When $K = 7$, the difference is about 1.0 dB. At $K = 40$, the difference becomes an insignificant 0.2 dB. Numerical results for DQPSK and 8-DPSK (which are not included in this paper) show the same trends as DPSK.

The results shown in Figs. 4–6 are all analytical results. In these figures, BEPs are represented by their approximate upper bounds. The approximate upper bound is computed based on the exact PEP expression in (29), except for Fig. 6. Fig. 4 shows the BEP of MSDD as a function of the number of symbols in the observation interval K. It is evident that for both DPSK (binary modulation) and for 8-DPSK ($M = 8$), the performance of MSDD approaches that of optimum combining as the observation interval increases.

Fig. 5 shows the BEP versus SIR, for bit SNR $= 10$ dB and for the cases of $K = 2$ and $K = 40$ symbols. It is observed that when $K = 40$, MSDD achieves performance close to that of optimum combining with differential encoding regardless of the SIR.

Fig. 6 is intended to verify the asymptotic large SNR approximation to the PEP. The signal modulation is DQPSK. Curves labeled "asymp" represent asymptotic results computed by applying (48) (for $K = 2$) and (49) (for $K = 40$); curves labeled "exact" represent exact results from (46) (for $K = 2$) and (47) (for $K = 40$). It is observed that for most SNR of interest (SNR > 10), the approximate upper bound based on asymptotic PEP is very close to the approximate upper bound based on the exact PEP.

VI. Conclusion

In this paper, we developed and analyzed an MSDD for differential PSK modulations over a Rayleigh fading channel with multiple independent outputs and in the presence of a Gaussian interference. A decision statistic utilizing blocks of observations was derived based on the maximum-likelihood criterion. Closed-form expressions were obtained for the PEP. Simpler approximations to the PEP were developed for the special case of large SNR and small SIR. It was shown that an approximation to the union bound can be used as an approximation to the BEP for binary and multiple-level differential PSK modulation. Moreover, it was shown that the MSDD detector could achieve performance close to that of optimum combining for an increasing observation interval. Theoretical results were demonstrated by comparison with Monte Carlo simulations. A suboptimal iterative MSDD algorithm was presented to facilitate the Monte Carlo simulation for long observation intervals.

APPENDIX
CHARACTERISTIC FUNCTION OF MSDD

In this Appendix, we derive the expression (25) for the characteristic function $\Phi_D(j\omega)$ of the MSDD test statistic D. To that end, using (17) and (23), we express the test statistic D in quadratic form

$$D = \sum_{\ell=1}^{L} \frac{1}{\lambda_\ell (KP_s + \lambda_\ell)} \left(|y_\ell(\mathbf{s}_k)|^2 - |y_\ell(\mathbf{s}'_k)|^2 \right) = \sum_{\ell=1}^{L} d_\ell \tag{53}$$

where

$$d_\ell = b_\ell^2 \left(|y_\ell(\mathbf{s}_k)|^2 - |y_\ell(\mathbf{s}'_k)|^2 \right) \tag{54}$$

$$b_\ell = \sqrt{\frac{1}{\lambda_\ell (KP_s + \lambda_\ell)}}. \tag{55}$$

Define vector

$$\begin{aligned}\mathbf{y}(\mathbf{s}_k) &= [y_1(\mathbf{s}_k), y_2(\mathbf{s}_k), \ldots, y_L(\mathbf{s}_k)]^T \\ &= \left(\sum_{i=0}^{K-1} \mathbf{x}_{k-i} s_{k-i}^* \right) \\ &= \mathbf{U}_z^H \left(\sum_{i=0}^{K-1} \mathbf{r}_{k-i} s_{k-i}^* \right).\end{aligned} \tag{56}$$

From the signal model in Section II, $E[\mathbf{y}(\mathbf{s}_k)] = E[\mathbf{y}(\mathbf{s}'_k)] = 0$. After some algebra, the covariance matrix of $\mathbf{y}(\mathbf{s}_k)$ can be evaluated as

$$\begin{aligned}E\left[\mathbf{y}(\mathbf{s}_k)\mathbf{y}(\mathbf{s}_k)^H\right] &= E\left[\mathbf{U}_z^H \left(\sum_{i=0}^{K-1} \mathbf{r}_{k-i} s_{k-i}^* \right) \cdot \left(\sum_{\ell=0}^{K-1} \mathbf{r}_{k-\ell} s_{k-\ell}^* \right)^H \mathbf{U}_z \right] \\ &= K^2 P_s \mathbf{I}_L + K \mathbf{\Lambda}_z\end{aligned} \tag{57}$$

where $\mathbf{\Lambda}_z$ was defined in Section III. Similarly, we have the following results:

$$E\left[\mathbf{y}(\mathbf{s}'_k)\mathbf{y}(\mathbf{s}'_k)^H\right] = |v(\mathbf{s}_k, \mathbf{s}'_k)|^2 P_s \mathbf{I}_L + K\mathbf{\Lambda}_z \tag{58}$$

$$E\left[\mathbf{y}(\mathbf{s}_k)\mathbf{y}(\mathbf{s}'_k)^H\right] = v^*(\mathbf{s}_k, \mathbf{s}'_k)(KP_s\mathbf{I}_L + \mathbf{\Lambda}_z) \tag{59}$$

$$E\left[\mathbf{y}(\mathbf{s}'_k)\mathbf{y}(\mathbf{s}_k)^H\right] = v(\mathbf{s}_k, \mathbf{s}'_k)(KP_s\mathbf{I}_L + \mathbf{\Lambda}_z) \tag{60}$$

where $v(\mathbf{s}_k, \mathbf{s}'_k) = \mathbf{s}'^H_k \mathbf{s}_k$.

To use the results in [18, App. B], we identify the following quantities using the notation in the reference: $X_\ell = y_\ell(\mathbf{s}_k)$, $Y_\ell = y_\ell(\mathbf{s}'_k)$. Then using (57) to (60) and [18, App. B, B-5, B-6], after some straightforward manipulations, we get the characteristic function of d_ℓ as

$$\phi_{d_\ell}(j\omega) = \frac{\theta_{1,\ell}\theta_{2,\ell}}{(\omega + j\theta_{1,\ell})(\omega - j\theta_{2,\ell})}. \tag{61}$$

where

$$\theta_{1,\ell} = \frac{1}{2(KP_s + \lambda_\ell)\lambda_\ell \zeta b_\ell^2} \cdot \left[\sqrt{\zeta^2 P_s^2 + 4(KP_s + \lambda_\ell)\lambda_\ell \zeta} - \zeta P_s \right] \tag{62}$$

$$\theta_{2,\ell} = \frac{1}{2(KP_s + \lambda_\ell)\lambda_\ell \zeta b_\ell^2} \cdot \left[\sqrt{\zeta^2 P_s^2 + 4(KP_s + \lambda_\ell)\lambda_\ell \zeta} + \zeta P_s \right] \tag{63}$$

$$\zeta = K^2 - |v(\mathbf{s}_k, \mathbf{s}'_k)|^2. \tag{64}$$

It follows that the characteristic function of D is

$$\Phi_D(j\omega) = \prod_{\ell=1}^{L} \phi_{d_\ell}(j\omega) = \prod_{\ell=1}^{L} \frac{1}{\left(1 - j\omega \frac{1}{\theta_{1,\ell}}\right)\left(1 + j\omega \frac{1}{\theta_{2,\ell}}\right)}. \tag{65}$$

For a system with a single interference source, the eigenvalues of the interference-plus-noise covariance matrix are $\lambda_\ell = \sigma^2$ for $\ell = 2, 3, \ldots, L$. It follows that $\theta_{1,\ell} = \theta_{1,2}$ and $\theta_{2,\ell} = \theta_{2,2}$ for $\ell = 2, 3, \ldots, L$. Hence, the characteristic function can be expressed as

$$\Phi_D(j\omega) = \frac{1}{(1 - j\mu_1\omega)} \frac{1}{(1 - j\mu_2\omega)} \cdot \frac{1}{(1 - j\mu_3\omega)^{L-1}} \frac{1}{(1 - j\mu_4\omega)^{L-1}} \tag{66}$$

where $\mu_1 = 1/\theta_{1,1}$, $\mu_2 = -1/\theta_{2,1}$, $\mu_3 = 1/\theta_{1,2}$, and $\mu_4 = -1/\theta_{2,2}$.

REFERENCES

[1] D. Divsalar and M. K. Simon, “Multiple symbol differential detection of MPSK,” *IEEE Trans. Commun.*, vol. 38, pp. 300–308, Mar. 1990.

[2] ——, “The performance of trellis-coded MDPSK with multiple symbol detection,” *IEEE Trans. Commun.*, vol. 38, pp. 1391–1404, Sept. 1990.

[3] ——, “Maximum-likelihood differential detection of uncoded and trellis-coded amplitude phase modulation over AWGN and fading channels—Metrics and performance,” *IEEE Trans. Commun.*, vol. 42, pp. 76–89, Jan. 1994.

[4] P. Ho and D. Fung, “Error performance of multiple-symbol differential detection of PSK signals transmitted over correlated Rayleigh fading channels,” *IEEE Trans. Commun.*, vol. 40, pp. 1566–1569, Oct. 1992.

[5] P. Ho, J. H. Kim, and E. B. Kim, “Performance of multiuser receivers for asynchronous CDMA with multiple symbol differential detection and pilot-aided coherent detection,” in *Proc. 1999 IEEE Int. Conf. Communications*, vol. 2, 1999, pp. 902–906.

[6] Q. Wang, M. Zeng, H. Yashima, and J. Suzuki, “Multiple-symbol detection of MPSK in narrowband interference and AWGN,” *IEEE Trans. Commun.*, vol. 46, pp. 460–463, Apr. 1998.

[7] M. K. Simon and M.-S. Alouini, “Multiple symbol differential detection with diversity reception,” *IEEE Trans. Commun.*, vol. 49, pp. 1312–1319, Aug. 2001.

[8] P. Fan, “Multiple-symbol detection for transmit diversity with differential encoding scheme,” *IEEE Trans. Consumer Electron.*, vol. 47, pp. 96–100, Feb. 2001.

[9] D. Makrakis and K. Feher, “Optimal noncoherent detection of PSK signals,” *Electron. Lett.*, vol. 26, pp. 398–400, Mar. 1990.

[10] G. Colavolpe and R. Raheli, “Noncoherent sequence detection,” *IEEE Trans. Commun.*, vol. 47, pp. 1376–1385, Sept. 1999.

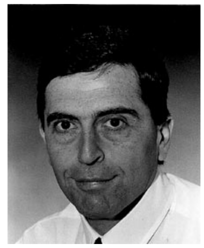
[11] V. Bogachev and I. Kiselev, “Optimum combining of signals in space-diversity reception,” *Telecommun. Radio Eng.*, vol. 34/35, pp. 83–85, Oct. 1980.

[12] A. Shah, A. M. Haimovich, M. K. Simon, and M.-S. Alouini, “Exact bit-error probability for optimum combining with a Rayleigh fading Gaussian cochannel interference,” *IEEE Trans. Commun.*, vol. 48, pp. 908–912, June 2000.

[13] M. K. Simon and M.-S Alouini, *Digital Communication Over Fading Channel: A Unified Approach to Performance Analysis*. New York: Wiley, 2000.
[14] K. M. Mackenthum, Jr., “A fast algorithm for multiple-symbol differential detection of MPSK,” *IEEE Trans. Commun.*, vol. 42, pp. 1471–1474, Feb.-Apr. 1994.
[15] R. Schober, W. H. Gerstacker, and J. B. Huber, “Decision-feedback differential detection of MDPSK for flat Rayleigh fading channels,” *IEEE Trans. Commun.*, vol. 47, pp. 1025–1035, July 1999.
[16] F. Edbauer, “Bit-error rate of binary and quaternary DPSK signals with multiple differential feedback detection,” *IEEE Trans. Commun.*, vol. 40, pp. 457–460, Mar. 1992.
[17] Y. Yu and Z. Chen, “Fast computation algorithm for multiple symbol detection of uncoded MPSK sequences over an AWGN channel,” *Electron. Lett.*, vol. 35, pp. 2086–2087, Nov. 1999.
[18] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.
[19] R. V. Churchill and J. W. Brown, *Complex Variables and Applications*. New York: McGraw-Hill, 1984.
[20] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*. San Diego, CA: Academic, 1994.
[21] D. Lao, “Performance evaluation for communication systems with reception diversity and interference,” Ph.D. dissertation proposal, NJ Inst. Technol., Newark, 2001.

Debang Lao (S’99) received the B.S. and M.S. degrees in electrical engineering from the University of Science and Technology of China, Hefei, China, in 1988 and 1993, respectively.

From 1993 to 1998 he was an Electrical Engineer at Beijing Astronomical Observatory, Beijing, China. Currently he is a Research Assistant at New Jersey Institute of Technology, Newark, where he is working toward the Ph.D. degree. His research interests include multiple symbol differential detection, space-time coding, turbo space-time coding, MIMO systems, and WCDMA standard.



Alexander M. Haimovich (S’82–M’87–SM’97) received the Ph.D. degree in systems from the University of Pennsylvania, Philadelphia, in 1989, the M.Sc. degree from Drexel University, Philadelphia, PA, in 1983, and the B.Sc. degree from the Technion, Haifa, Israel in 1977, both in electrical engineering.

He is a Professor of Electrical and Computer Engineering at the New Jersey Institute of Technology (NJIT), Newark. He serves as the Director of the New Jersey Center for Wireless Telecommunications, a state-funded consortium consisting of NJIT, Princeton University, Rutgers University, and Stevens Institute of Technology. His research interests include MIMO systems, array processing for wireless, turbo coding, space-time coding, and ultra-wideband systems.