On Multiple Symbol Detection for Diagonal DUSTM over Ricean Channels

Tao Cui and Chintha Tellambura, *Senior Member, IEEE*

***Abstract*— This letter considers multiple symbol differential detection for multiple-antenna systems over flat Ricean-fading channels when partial channel state information (CSI) is available at the transmitter. Using the maximum likelihood (ML) principle, and assuming perfect knowledge of the channel mean, we derive the optimal multiple symbol detection (MSD) rule for diagonal differential unitary space-time modulation (DUSTM). This rule is used to develop a sphere decoding bound intersection detector (SD-BID) with low complexity. A suboptimal MSD based decision feedback DD (DF-DD) algorithm is also derived. The simulation results show that our proposed MSD algorithms reduce the error floor of conventional differential detection and that the computational complexity of these new algorithms is reasonably low.**

***Index Terms*— Multiple symbol detection, sphere decoding, DUSTM, Ricean channel, MIMO.**

I. Introduction

DIFFERENTIAL detection operates without channel state information (CSI) and is robust against the carrier phase ambiguity. It has been used in practical standards such as Bluetooth 2.0 and IEEE 802.11b (or Wi-Fi). However, since conventional differential detection performs worse than coherent detection, in single-input and single output (SISO) systems, multiple-symbol detection (MSD) for M-ary differential phase-shift keying (MDPSK) has been developed [1]. MSD jointly detects N data symbols using N consecutive received samples. The computational complexity of MSD hence grows exponentially with N. The sphere decoder (SD) [2] has also been used to further reduce the complexity of MSD [3]. Alternatively, decision feedback differential detection (DF-DD) offers reasonable performance while ensuring low complexity [4].

In [5], noncoherent differential unitary space-time modulation (DUSTM) receivers based on MSD and DF-DD are derived. A fast, suboptimal DUSTM detector is derived in [6]. We have recently derived, for MSD of DUSTM over quasi-static fading channels, an efficient MSD bound intersection detector (BID) in [7], [8]. Our BID is optimal and can be more efficient than that in [6] in high signal-to-noise ratio (SNR). Various implementations of a tree search based MSD for DUSTM by using the suboptimal detector in [6] has been proposed in [9]. However, the tree search detectors in [9] are still suboptimal.

The Ricean distribution is a widely used fading model. It encompasses, as special cases, Rayleigh fading and additive white Gaussian noise (AWGN) channels. For single-antenna systems, an MSD rule has been given in [1], but no efficient detection algorithm is proposed. A DF-DD scheme for flat Ricean-fading channels based on linear prediction is proposed in [10]. A MSD-based DF-DD rule for Ricean fading is also given in [10]. Conventional differential detection uses the previous received symbol as the reference symbol, whereas the prediction based DF-DD [10] uses several previous received signals to predict the reference symbol to mitigate the noise and channel varying effects. For multiple-antenna systems over Ricean channels, the derivation of the maximum likelihood (ML) metric for the single symbol detection of DUSTM with a known direct fading component for Ricean fading is given in [11]. To the best of our knowledge, the optimal MSD rule for DUSTM transmitted over multiple-antenna Ricean channels has not been derived in the literature.

In this paper, we first derive the optimal MSD metric for diagonal DUSTM over flat fast Ricean-fading channels, which includes the MSD metric for MDPSK as a special case. The derivation assumes that the channel direct component is known, as in [11], [12]. The decision rule for diagonal DUSTM reduces to the one in [11] when $N = 1$. In order to significantly reduce the detection complexity, we combine the branch and bound (BnB) principle and BID [7], [8] and derive a sphere decoding bound intersection detector (SD-BID), with the performance identical to that of maximum-likelihood exhaustive search. We also generalize the Schnorr-Euchner search strategy to SD-BID. We also propose a suboptimal MSD-based DF-DD using BID [7], [8]. Although this scheme does not achieve ML performance, it performs substantially better than differential detection and its complexity is linear in N. Surprisingly, in high SNR, the complexity of SD-BID is even lower than that of the DF-DD scheme. This finding agrees with the observation in [7], [9] for the Rayleigh fading channel.

Notation: $E\{\cdot\}$, $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^\dagger$ denote expectation, complex conjugation, transpose, conjugate transpose and Moore-Penrose pseudo-inverse. The imaginary unit is $\jmath = \sqrt{-1}$. The trace, determinant and the Frobenius norm of matrix $\mathbf{A}$ are $\mathrm{tr}(\mathbf{A})$, $\det(\mathbf{A})$ and $\|\mathbf{A}\|_F^2 = \mathrm{tr}(\mathbf{A}\mathbf{A}^H)$. $\lfloor x \rceil$ denotes the closest integer to x. A circularly complex Gaussian variable with mean m and variance σ^2 is denoted by $z \sim \mathcal{CN}(m, \sigma^2)$. The sets of real numbers and integers are

Manuscript received October 26, 2006; revised January 18, 2007; accepted September 3, 2007. The associate editor coordinating the review of this paper and approving it for publication was S. Kishore. This work has been supported in part by the Natural Sciences and Engineering Research Council of Canada, Informatics Circle of Research Excellence and Alberta Ingenuity Fund. This paper was presented in part at the IEEE International Conference on Communications, June 2006, Istanbul, Turkey.

T. Cui is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA (e-mail: taocui@caltech.edu).

C. Tellambura is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: chintha@ece.ualberta.ca).

Digital Object Identifier 10.1109/TWC.2008.060872.

$\mathbb{R}$ and $\mathbb{Z}$, and the $N \times N$ identity matrix is $\mathbf{I}_N$. $\delta_{i,j}$ is the Kronecker delta, for $i, j \in \mathbb{Z}$, $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ if $i \neq j$. $\mathbf{A} = \text{diag}\{\mathbf{A}_1, \mho, \mathbf{A}_N\}$ denotes a block diagonal matrix with $\mathbf{A}_1, \mho, \mathbf{A}_N$ in its diagonal blocks.

II. System Model

We consider a multiple-antenna system with N_t transmit and N_r receive antennas. For simplicity, all signals are represented in the equivalent complex-valued low-pass domain. Ideal symbol synchronization is assumed. Each time slot occupies an interval T_s in seconds, and each block consists of T time slots. The block interval is thus $T_B = T_s T$ in seconds. All the transmitted symbols over the nth block are arranged in the $T \times N_t$ matrix $\mathbf{S}[n] = [s_{t,i}[n]]$, $t = 1, 2, \ldots, T$ and $i = 1, 2, \ldots, N_t$, where $s_{t,i}[n]$ is transmitted from the ith antenna in the $t + (n-1)T$ time slot.

We consider a flat Ricean-fading multiple-antenna channel arising from a rich scattering environment. We assume that the channel does not change significantly during one slot interval T_s. The complex base-band received signal at the jth receive antenna, at time slot t in the nth block can be written as

$$r_{t,j}[n] = \sum_{i=1}^{N_t} h_{i,j}[t + (n-1)T] s_{t,i}[n] + w_{t,j}[n], \quad j = 1, \ldots, N_r,\ t = 1, \ldots, T, \tag{1}$$

where $h_{i,j}[t]$ denotes the channel gain from the ith transmit antenna to the jth receive antenna in the tth time slot, and $w_{t,j}[n]$ is the complex additive white Gaussian noise (AWGN) at the jth receive antenna. The additive Gaussian noise processes at different receive antennas are independent and have equal variance σ_n^2. The channel gain $h_{i,j}[t]$ is a complex Gaussian random process and can be expressed as

$$h_{i,j}[t] = (h_d)_{i,j}[t] + (h_s)_{i,j}[t], \tag{2}$$

where $(h_d)_{i,j}[t]$ is the direct component $(h_d)_{i,j}[t] = E\{h_{i,j}[t]\}$, and $(h_s)_{i,j}[t]$ is the scattered component, which is a zero mean Gaussian process. We assume that all path gains are statistically independent ($E\{(h_s)_{i,j}[n](h_s)^*_{i',j'}[n]\} = 0$, $\forall (i,j) \neq (i',j')$) and have the same autocorrelation function $\varphi_h[k] = E\{(h[t+k] - h_d[t+k])(h[t] - h_d[t])^*\}$.

In DUSTM, the signals are modulated by choosing a matrix from a finite group $\mathcal{V} = \{\mathbf{V}_l, l = 0, 1, \ldots, L-1\}$, where $\mathbf{V}_l$ is a $T \times N_t$ unitary matrix ($\mathbf{V}_l \mathbf{V}_l^H = \mathbf{I}_T$), and $L = 2^{N_t R}$, and R denotes the normalized data rate [13]. To realize DUSTM, we assume $T = N_t$ and $\mathbf{V}_0 = \mathbf{I}_{N_t}$. The $N_t R$ binary information bits are first converted to an integer l within $[0, L-1]$, and $\mathbf{V}[n] = \mathbf{V}_l$ is chosen from $\mathcal{V}$. The transmitted block at the nth block is encoded as $\mathbf{S}[n] = \mathbf{V}[n]\mathbf{S}[n-1]$. The first block is $\mathbf{S}[0] = \mathbf{V}_0$. The internal composition property of a group ensures that $\mathbf{S}[n] \in \mathcal{V}$ is unitary for any positive n. In this paper, we consider diagonal constellations, a special kind of DUSTM. For diagonal constellations, the unitary matrices $\mathbf{V}_l$ are chosen as

$$\mathbf{V}_l = \text{diag}\left\{e^{\jmath 2\pi u_1 l/L}, e^{\jmath 2\pi u_2 l/L}, \ldots, e^{\jmath 2\pi u_{N_t} l/L}\right\}, \tag{3}$$

where u_i, $i = 1, 2, \ldots, N_t$ are obtained from [13]. From [13], u_i is relatively prime to L. We thus have $s_{t,i}[n] \neq 0$ $t = i$, and $s_{t,i}[n] = 0$ for $t \neq i$. Therefore, (1) can be rewritten as

$$r_{t,j}[n] = h_{t,j}[t + (n-1)T] s_{t,t}[n] + w_{t,j}[n], \quad j = 1, \ldots, N_r,\ t = 1, \ldots, N_t. \tag{4}$$

The matrix form of (4) is thus

$$\mathbf{R}[n] = \mathbf{S}[n]\mathbf{H}[n] + \mathbf{W}[n] = \mathbf{S}[n](\mathbf{H}_d[n] + \mathbf{H}_s[n]) + \mathbf{W}[n], \tag{5}$$

where $\mathbf{R}[n] = [r_{t,j}[n]]$ is the $N_t \times N_r$ receive matrix, $\mathbf{W}[n] = [w_{t,j}[n]]$ is the $N_t \times N_r$ noise matrix, $\mathbf{H}[n]$ is the $N_t \times N_r$ channel matrix, and the (i,j)th entry of $\mathbf{H}[n]$ is $h_{i,j}[i + (n-1)T]$. The second equality comes from (2), where the (i,j)th entries of $\mathbf{H}_d[n]$ and $\mathbf{H}_s[n]$ are the direct and scattering components of $h_{i,j}[i + (n-1)T]$, respectively. From [14, pp. 34, (3.2)], we have

$$\mathbf{H}_d[t] = h_m e^{\jmath(2\pi f_D \cos(\theta_0) T_s t + \phi)} \mathbf{a}(\theta_t)\mathbf{a}(\theta_r)^T, \tag{6}$$

where h_m and ϕ denote the amplitude and the phase of the direct component, θ_0 is the angle between the line-of-sight and the mobile direction, θ_r and θ_t as the angle of arrival and the angle of departure, respectively, and $\mathbf{a}(\theta_t)$, $\mathbf{a}(\theta_r)$ are the direct array responses at the transmitter and receiver. The array response corresponding to an N-element ($N \in \{N_t, N_r\}$) linear array is given by $\mathbf{a}(\theta) = [1, e^{\jmath 2\pi d \cos(\theta)}, \ldots, e^{\jmath 2\pi d(N-1)\cos(\theta)}]^T$, where θ is the angle of arrival or departure of the direct component and d is the antenna spacing in wavelengths. In (6), we assume that all the direct components have the same amplitude and that the transmit antennas are far from the receive antennas such that θ_0 is common to every transmit and receive antenna pair. The Rice factor K is defined as $K = \frac{h_m^2}{\sigma_h^2}$ [14].

III. Multiple-symbol differential detection with partial channel state information

By following the same approach as in [5], [7], we consider the received signal from $n = k+1$ to $n = k+N$. Let $\bar{\mathbf{R}}[k] = [\mathbf{R}^H[k+1], \ldots, \mathbf{R}^H[k+N]]^H$ (a $NN_t \times N_r$ matrix) and $\bar{\mathbf{H}}[k] = [\mathbf{H}^H[k+1], \ldots, \mathbf{H}^H[k+N]]^H$ (a $NN_t \times N_r$ matrix). The matrix input-output relationship is

$$\bar{\mathbf{R}}[k] = \bar{\mathbf{S}}_D[k]\bar{\mathbf{H}}[k] + \bar{\mathbf{W}}[k] = \bar{\mathbf{S}}_D[k](\bar{\mathbf{H}}_d[k] + \bar{\mathbf{H}}_s[k]) + \bar{\mathbf{W}}[k], \tag{7}$$

where $\bar{\mathbf{S}}_D[k] = \text{diag}\{\mathbf{S}[k+1], \ldots, \mathbf{S}[k+N]\}$ is an $NN_t \times NN_t$ block diagonal matrix, and $\bar{\mathbf{H}}_d[k] = [\mathbf{H}_d^H[k+1], \ldots, \mathbf{H}_d^H[k+N]]^H$, $\bar{\mathbf{H}}_s[k] = [\mathbf{H}_s^H[k+1], \ldots, \mathbf{H}_s^H[k+N]]^H$, and $\bar{\mathbf{W}}[k] = [\mathbf{W}^H[k+1], \ldots, \mathbf{W}^H[k+N]]^H$ are all $NN_t \times N_r$ matrices. We drop the time index k in the following for simplicity. As is the case with single-antenna systems, $\text{vec}(\bar{\mathbf{R}})$ is a complex Gaussian vector, and the conditional pdf given $\bar{\mathbf{S}}_D[k]$ is

$$f(\bar{\mathbf{R}} | \bar{\mathbf{S}}_D) = \frac{1}{(\pi^{NN_t} \det(\mathbf{C}_R))^{N_r}} \times \exp\left\{-\text{tr}\left((\bar{\mathbf{R}} - \bar{\mathbf{S}}_D\bar{\mathbf{H}}_d)^H \mathbf{C}_R^{-1} (\bar{\mathbf{R}} - \bar{\mathbf{S}}_D\bar{\mathbf{H}}_d)\right)\right\}, \tag{8}$$

where we have assumed that the channel direct component is given; i.e., $\bar{\mathbf{H}}_d$ is known. The autocovariance matrix $\mathbf{C}_R$ is given by

$$\begin{aligned}\mathbf{C}_R &= E\{(\bar{\mathbf{R}} - \bar{\mathbf{S}}_D\bar{\mathbf{H}}_d)(\bar{\mathbf{R}} - \bar{\mathbf{S}}_D\bar{\mathbf{H}}_d)^H\} \\ &= \bar{\mathbf{S}}_D \mathbf{C}_H \bar{\mathbf{S}}_D^H + N_r \sigma_n^2 \mathbf{I}_{N_t N} = N_r(\mathbf{C}_h \otimes \mathbf{I}_{N_t}),\end{aligned} \tag{9}$$

where $\mathbf{C}_H$ is the covariance matrix of $\bar{\mathbf{H}}$, $\otimes$ denotes the Kronecker product [15], and

$$\mathbf{C}_h = \begin{bmatrix} \varphi_h[0] & \varphi_h[T] & \cdots & \varphi_h[(N-1)T] \\ \varphi_h[-T] & \varphi_h[0] & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_h[-(N-1)T] & \cdots & \cdots & \varphi_h[0] \end{bmatrix}. \tag{10}$$

Note that (9) is true because the (i,j)th entries of $\mathbf{H}[n]$ and $\mathbf{H}[m]$ are $h_{i,j}[i+(n-1)T]$, $h_{i,j}[i+(m-1)T]$. Thus, $E\{h_{i,j}[i+(n-1)T]h^*_{i,j}[i+(m-1)T]\} = \varphi_h[(n-m)T]$. Since $\mathbf{S}[n]$ $(n=k,k+1,\ldots,k+N)$ are unitary matrices, we have $\bar{\mathbf{S}}_D\bar{\mathbf{S}}_D^H = \mathbf{I}_{N_tN}$ and $\mathbf{C}_R = N_r\bar{\mathbf{S}}_D\left(\mathbf{C}\otimes\mathbf{I}_{N_t}\right)\bar{\mathbf{S}}_D^H$, where $\mathbf{C} = \mathbf{C}_h + \sigma_n^2\mathbf{I}_N$ and it follows the distributivity property of the Kronecker product [15]. We also have $\det(\mathbf{C}_R) = \det\left(N_r\left(\mathbf{C}\otimes\mathbf{I}_{N_t}\right)\bar{\mathbf{S}}_D^H\bar{\mathbf{S}}_D\right) = \det\left(N_r\left(\mathbf{C}\otimes\mathbf{I}_{N_t}\right)\right)$. Therefore, maximizing (8) is equivalent to minimizing

$$g(\bar{\mathbf{S}}_D) = \mathrm{tr}\left((\bar{\mathbf{R}} - \bar{\mathbf{S}}_D\bar{\mathbf{H}}_d)^H\mathbf{C}_R^{-1}(\bar{\mathbf{R}} - \bar{\mathbf{S}}_D\bar{\mathbf{H}}_d)\right). \tag{11}$$

Note that

$$\mathbf{C}_R^{-1} = \frac{1}{N_r}\bar{\mathbf{S}}_D\left(\mathbf{C}\otimes\mathbf{I}_{N_t}\right)^{-1}\bar{\mathbf{S}}_D^H = \frac{1}{N_r}\bar{\mathbf{S}}_D\left(\mathbf{C}^{-1}\otimes\mathbf{I}_{N_t}\right)\bar{\mathbf{S}}_D^H, \tag{12}$$

where the first equality derives from $\bar{\mathbf{S}}_D\bar{\mathbf{S}}_D^H = \mathbf{I}_{N_tN}$, and the second equality derives from the Kronecker product property $(\mathbf{A}\otimes\mathbf{B})^{-1} = \mathbf{A}^{-1}\otimes\mathbf{B}^{-1}$ ($\mathbf{A}$ and $\mathbf{B}$ are square nonsingular matrices) [15]. Let the Cholesky factorization of $\mathbf{C}^{-1}$ be $\mathbf{C}^{-1} = \mathbf{U}^H\mathbf{U}$, where $\mathbf{U}$ is upper triangular. By using the Kronecker product property $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = \mathbf{AC}\otimes\mathbf{BD}$, $\mathbf{C}^{-1}\otimes\mathbf{I}_{N_t} = (\mathbf{U}\otimes\mathbf{I}_{N_t})^H(\mathbf{U}\otimes\mathbf{I}_{N_t}) = \bar{\mathbf{U}}^H\bar{\mathbf{U}}$, and $\bar{\mathbf{U}}$ is also upper triangular. After several manipulations, we simplify (11) as

$$g(\bar{\mathbf{S}}_D) = \left\|\bar{\mathbf{U}}\bar{\mathbf{H}}_d - \bar{\mathbf{U}}\bar{\mathbf{S}}_D^H\bar{\mathbf{R}}\right\|_F^2 = \left\|\bar{\mathbf{Y}} - \bar{\mathbf{U}}\bar{\mathbf{S}}_D^H\bar{\mathbf{R}}\right\|_F^2, \tag{13}$$

where $\bar{\mathbf{Y}} = \bar{\mathbf{U}}\bar{\mathbf{H}}_d = [\mathbf{Y}^H[k+1],\mathbf{Y}^H[k+2],\ldots,\mathbf{Y}^H[k+N]]^H$, and $\mathbf{Y}[n]$ is an $N_t\times N_r$ matrix. The MSD rule for DUSTM over multiple-antenna Ricean channels is given by

$$\left\{\hat{\mathbf{S}}[k+1],\ldots,\hat{\mathbf{S}}[k+N]\right\} = \underset{\mathbf{S}[k+1],\ldots,\mathbf{S}[k+N]\in\mathcal{V}}{\arg\min}\left\|\bar{\mathbf{Y}} - \bar{\mathbf{U}}\bar{\mathbf{S}}_D^H\bar{\mathbf{R}}\right\|_F^2. \tag{14}$$

The transmitted signals can be differentially detected as

$$\hat{\mathbf{V}}[n] = \hat{\mathbf{S}}[n+1]\hat{\mathbf{S}}^H[n]. \tag{15}$$

IV. Efficient Multiple Symbol Differential Detection

A. SD-BID based MSD

We now present our sphere decoding bound intersection detector (SD-BID) based on the MSD rule (14). As with the sphere decoder, we examine only the candidates that satisfy

$$\|\bar{\mathbf{Y}} - \bar{\mathbf{U}}\bar{\mathbf{S}}_D^H\bar{\mathbf{R}}\|_F^2 \le R^2. \tag{16}$$

Instead of explicitly enumerating all the candidates, we combine the branch-and-bound (SD) and the divide-and-conquer approach of BID. Let the entries of $\mathbf{U}$ be denoted by $u_{i,j}$, $i\le j$. By taking the upper triangular and Kronecker product structure of $\bar{\mathbf{U}}$ into account, (16) can be written as

$$\sum_{i=1}^{N}\left\|\mathbf{Y}[k+i] - \sum_{j=i}^{N}u_{i,j}\mathbf{S}^H[k+j]\mathbf{R}[k+j]\right\|_F^2 \le R^2. \tag{17}$$

Thus a necessary condition for (17) is

$$\sum_{i=n}^{N}\left\|\mathbf{Y}[k+i] - \sum_{j=i}^{N}u_{i,j}\mathbf{S}^H[k+j]\mathbf{R}[k+j]\right\|_F^2 \le R^2, \quad n=N,\ldots,1. \tag{18}$$

Condition (18) can be checked component by component. To proceed, we start from $\mathbf{S}[k+N]$. Using BID, we can obtain its candidate set

$$\mathcal{I}_N = \left\{\mathbf{V}_l \,\middle|\, \left\|\mathbf{Y}[k+N] - u_{N,N}\mathbf{V}_l^H\mathbf{R}[k+N]\right\|_F^2 \le R^2, \; l\in\{0,1,\ldots,L-1\}\right\}, \tag{19}$$

where $\mathbf{V}_l$ is defined in (3). After choosing $\hat{\mathbf{S}}[k+j]$ for $\mathbf{S}[k+j]$ from their candidate set, $i+1\le j\le N$, we define

$$d_i^2 = \left\|\mathbf{Z}[k+i] - u_{i,i}\mathbf{S}^H[k+i]\mathbf{R}[k+i]\right\|_F^2, \tag{20}$$

$$R_N^2 = R^2,\; R_i^2 = R_{i+1}^2 - d_{i+1}^2,\; 1\le i\le N-1, \tag{21}$$

where $\mathbf{Z}[k+i] = \mathbf{Y}[k+i] - \sum_{j=i+1}^{N}u_{i,j}\hat{\mathbf{S}}^H[k+j]\mathbf{R}[k+j]$.

By following the same arguments as in [7], [8], d_i^2 can be further simplified as

$$d_i^2 = \sum_{p=1}^{N_t}A_p - B_p\cos\left((u_pl+\phi_p)\frac{2\pi}{L}\right), \tag{22}$$

where

$$\begin{aligned} A_p &= \sum_{q=1}^{N_r}|z_{p,q}[k+i]|^2 + |u_{i,j}r_{p,q}[k+i]|^2, \\ B_p &= 2\left|\sum_{q=1}^{N_r}z_{p,q}^*[k+i]u_{i,j}r_{p,q}[k+i]\right|, \\ \phi_p &= \arg\left(\sum_{q=1}^{N_r}z_{p,q}^*[k+i]u_{i,j}^*r_{p,q}^*[k+i]\right)L/2\pi. \end{aligned} \tag{23}$$

The candidate set for $\mathbf{S}[k+i]$ can be obtained as

$$\mathcal{I}_i = \left\{\mathbf{V}_l \,\middle|\, \sum_{p=1}^{N_t}A_p - B_p\cos\left((u_pl+\phi_p)\frac{2\pi}{L}\right) \le R_k^2, \; l\in\{0,1,\ldots,L-1\}\right\}. \tag{24}$$

To find $\mathcal{I}_i$, we need first to find

$$\mathcal{L}_p = \left\{l \,\middle|\, A_p - B_p\cos\left((u_pl+\phi_p)\frac{2\pi}{L}\right) \le R_k^2, \; l\in\{0,1,\ldots,L-1\}\right\},\; p=1,\ldots,N_t. \tag{25}$$

$\mathcal{L}_p$ can be found by using the algorithm in [7], [8]. $\mathcal{I}_i$ can be obtained as

$$\mathcal{I}_i = \left\{\mathbf{V}_1^l \,\middle|\, l\in\bigcap_{p=1}^{N_t}\mathcal{L}_p\right\}. \tag{26}$$

Due to space limitations, we cannot outline the details of the BID algorithm. For the whole BID algorithm and efficient implementations, the reader is referred to [7], [8].

When all $\hat{\mathbf{S}}[k+i]$ are found, all the R_i's are updated according to

$$R_N^2 = \|\bar{\mathbf{Y}} - \bar{\mathbf{U}}\bar{\mathbf{S}}_D^H\bar{\mathbf{R}}\|_F^2,\ R_i^2 = R_{i+1}^2 - d_{i+1}^2, \quad i = N-1,\ldots,1. \tag{27}$$

The process continues until all the candidates satisfying (16) have been checked. The best candidate is the ML solution. If $N_t = N_r = 1$, the SD-BID reduces to the conventional SD.

The initial radius R can also be obtained according to the statistic of $g(\bar{\mathbf{S}}_D)$ in (11). If $\bar{\mathbf{S}}_D$ is the true solution, $\mathbf{X} = \bar{\mathbf{S}}_D^H[k]\bar{\mathbf{R}}[k] - \bar{\mathbf{H}}_d[k] = \bar{\mathbf{H}}_s[k] + \mathbf{S}_D^H[k]\bar{\mathbf{W}}[k]$ is zero mean complex Gaussian with autocovariance matrix $\mathbf{C}_X = \mathbf{C}_h + \sigma_n^2\mathbf{I}_N$. Therefore, $e = \mathrm{tr}\{\mathbf{X}^H(\mathbf{C}_h+\sigma_n^2\mathbf{I}_N)^{-1}\mathbf{X}\}$ is a chi-square random variable with $2NN_rN_t$ degrees of freedom. By using the radius-selection strategy of [16], R^2 can be chosen to make the probability that e is less than R^2 very high:

$$\int_0^{R^2} \frac{x^{NN_rN_t-1}e^{-x/2}}{\Gamma(NN_rN_t)2^{NN_rN_t}}\mathrm{d}x = 1-\epsilon, \tag{28}$$

where ϵ can be reduced to ensure with high probability that the ML solution is contained within the initial hypersphere. Clearly, the initial radius does not depend on the noise variance.

The Schnorr and Euchner [17] strategy can also be generalized to SD-BID. In each step, we not only find the candidate set $\mathcal{I}_i$ but also compute the corresponding d_i^2 by using (22) and store it in $\mathcal{D}_i$. $\mathcal{I}_i$ is sorted according to $\mathcal{D}_i$. The candidate with minimum d_i^2 is searched first. Note that each test in (22) needs $4N_t$ flops only. However, if d_i^2 is calculated by using (20) directly, it needs $15N_tN_r$ flops. Therefore, using (22) to compute d_i^2, sorting and R_i^2 updating offer significant complexity saving. To avoid computing d_i^2 for each candidate in $\mathcal{I}_i$, we suggest sorting $\mathcal{I}_i$ by $A_1 - B_1\cos((u_1 l - \phi_1)2\pi/L)$ instead of sorting $\mathcal{I}_i$ according to d_i^2. Thus, we should sort $\mathcal{I}_i$ by $\mathrm{mod}(|l-\phi_1|, L)$, $\mathbf{V}_1^l \in \mathcal{I}_i$. This sorting can be done in a similarly to that in the Schnorr and Euchner strategy by choosing l according to the distance from ϕ_1.

B. Reduced-state DD

Assuming that the previous decisions $\hat{\mathbf{S}}[k+1],\ldots,\hat{\mathbf{S}}[k+N-1]$ have been correct, the MSD for DUSTM (14) can be readily modified to MSD based DF-DD by replacing $\mathbf{S}[k+1],\ldots,\mathbf{S}[k+N-1]$ in (14) with $\hat{\mathbf{S}}[k+1],\ldots,\hat{\mathbf{S}}[k+N-1]$. Our BID can be used to solve the DF-DD. We also note that the decision feedback sequence estimator is a special case of the reduced-state sequence estimator (RSSE) [18]. Similarly, a reduced-state differential detector (RS-DD) can be used to solve (14) as a generalization of the DF-DD. Instead of assuming $N-1$ correct feedbacks in (14), RS-DD uses only M $(0 \le M \le N-1)$ decision feedbacks. $\mathbf{S}[k+1],\ldots,\mathbf{S}[k+M]$ in (14) are replaced with $\hat{\mathbf{S}}[k+1],\ldots,\hat{\mathbf{S}}[k+M]$, and SD-BID is used for the resulting $N-M$ dimensional problem. If $M=0$, RS-DD reduces to SD-BID and DF-DD when $M = N-1$. Thus, both the performance and complexity of RS-DD are between those of SD-BID and DF-DD.

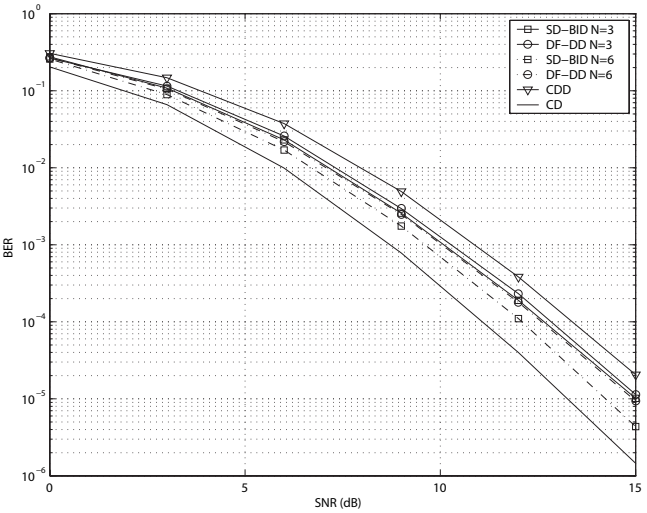

Fig. 1. The performance comparison between SD-BID, MSD based DF-DD, differential detection and coherent detection with $N = 3, 6$ for DUSTM ($N_t = 4$, $N_r = 1$ and $R = 1$) over flat Ricean channels ($f_DT = 0.0075$ and $K = 5$dB).

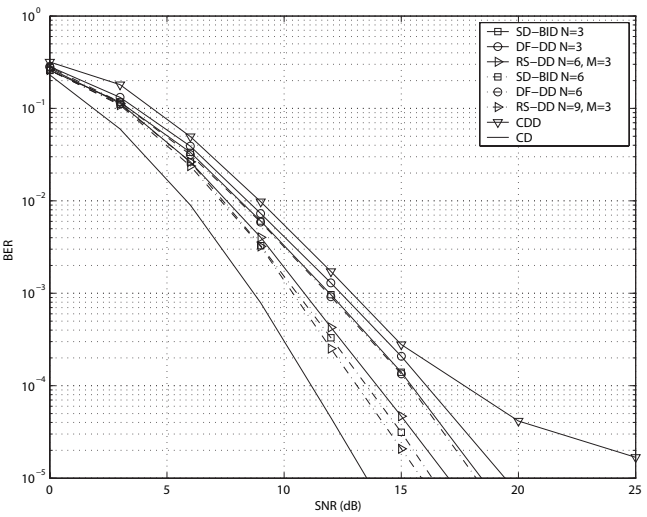

Fig. 2. The performance comparison between SD-BID, MSD based DF-DD, differential detection and coherent detection with $N = 3, 6$ for DUSTM ($N_t = 4$, $N_r = 1$ and $R = 1$) over flat Ricean channels ($f_DT_B = 0.03$ and $K = 5$dB).

V. Simulation Results

The $N_t = 4$, $N_r = 1$ and rate $R = 1$ DUSTM is used. The code parameters are taken from [13]. The Jakes' model is assumed for each channel. The direct channel matrix is assumed to be $\mathbf{H}_d[n] = \sqrt{K/(K+1)}\mathbf{1}_{N_t\times N_r}$ [19], where $\mathbf{1}_{N_t\times N_r}$ is an all one matrix. We assume that the receiver has perfect knowledge of K, $\mathbf{C}_h$ and σ_n^2.

Fig. 1 shows the BER versus SNR for SD-BID, MSD based DF-DD (DF-DD), with $N = 3, 6$, $f_DT = 0.0075$, and Rice factor $K = 5$ dB. They are compared with conventional differential detection (CDD) stated in [1] and coherent detection (CD) with perfect CSI. Compared with DF-DD, the SD-BID has a 0.1 dB gain ($N = 3$) and a 0.4 dB gain ($N = 6$) at a BER of 10^{-4}, respectively. Both DF-DD and SD-BID can reduce the performance gap between differential detection and coherent detection. The performance loss of SD-

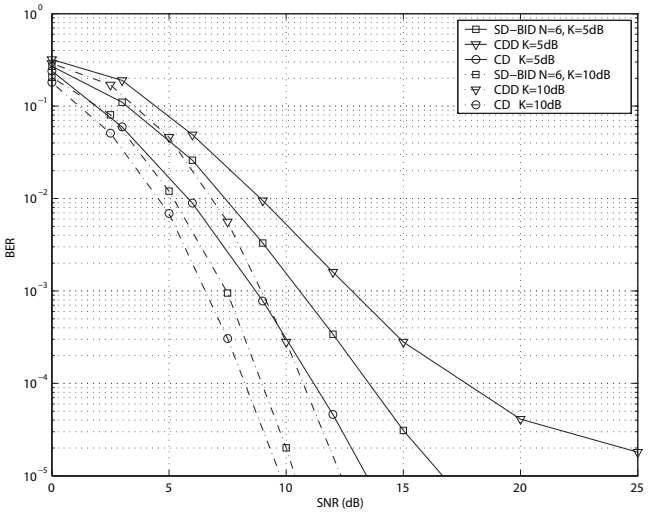

Fig. 3. The performance comparison between SD-BID ($N = 6$), differential detection and coherent detection for DUSTM ($N_t = 4$, $N_r = 1$ and $R = 1$) over flat Ricean channels ($f_D T_B = 0.03$) with different Rice factor K.

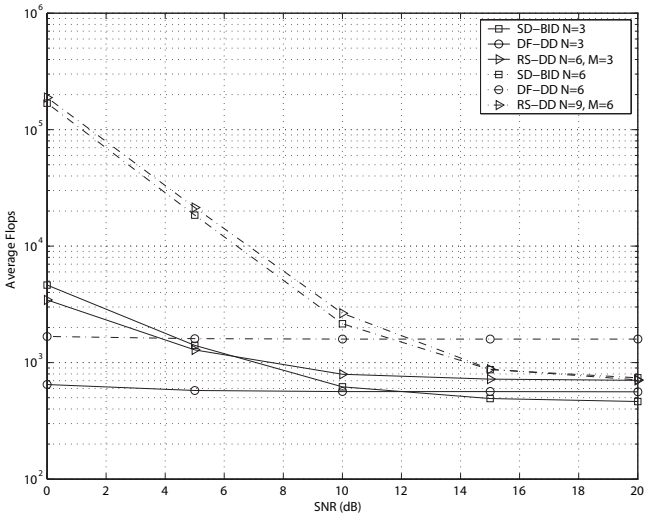

Fig. 4. The average number of flops comparison between SD-BID, RS-DD, MSD based DF-DD for DUSTM ($N_t = 4$, $N_r = 1$ and $R = 1$) over flat Ricean channels ($f_D T_B = 0.03$ and $K = 5$dB).

BID over coherent detection reduces as N increases. When the Doppler rate $f_D T_B$ increases to 0.03, the gap between SD-BID and coherent detection increases from 1 dB to 2 dB (Fig. 2). At a BER of 5×10^{-4}, the DF-DD scheme performs 0.6 dB and 1.2 dB worse than SD-BID. This finding agrees with what is known for single-antenna systems; that is, the gap between SD-BID and DF-DD increases with the increase of N. We also show the performance of RS-DD in Fig. 2. When $N = 6$, $M = 3$, RS-DD has about a 0.6 dB gain over SD-BID with $N = 3$ at BER$= 10^{-4}$, when both use a 3-dimensional exhaustive search. RS-DD outperforms SD-BID by 0.2 dB when $N = 9$, $M = 3$. RS-DD thus performs well and maintains low complexity.

Fig. 3 presents the performance of SD-BID ($N = 6$), differential detection and coherent detection for several Rice factors. All the detectors perform better as the Rice factor increases. The gaps between SD-BID and coherent detection

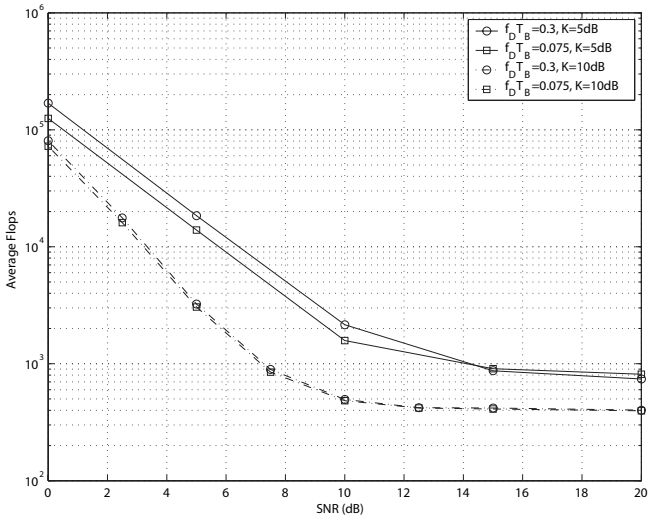

Fig. 5. The average number of flops of SD-BID for DUSTM ($N_t = 4$, $N_r = 1$ and $R = 1$) over flat Ricean channels for different $f_D T_B$ and Rice factor K.

and between SD-BID and differential detection reduce as the Rice factor increases.

The complexities of different detectors are compared in Fig. 4. SD-BID not only has a decrease in complexity as SNR increases, but has lower complexity than DF-DD in high SNR, where SD-BID has both complexity and performance gains. Interestingly, the RS-DD complexity is less than that of SD-BID. In RS-DD, the matrix $\mathbf{U}$ after deleting the corresponding columns to the feedback signals is different from the $\mathbf{U}$ in SD-BID with the same size. The diagonal terms of the matrix in RS-DD are larger than those in pure SD-BID. Fig. 5 shows the complexity of SD-BID as a function of SNR for different Rice factors and Doppler rates. When the Rice factor increases, the complexity of SD-BID reduces significantly as SD-BID becomes coherent detection when $K \to \infty$. Fig. 5 also shows that a smaller Doppler rate results in less complexity. As with RS-DD, the complexity reductions increase because of the change on the structure of matrix $\mathbf{U}$.

VI. Conclusion

In this paper, we have derived the optimal decision metrics of multiple symbol differential detection for diagonal DUSTM over Ricean fading channels with partial channel state information. A BID with a modified Schnorr and Euchner strategy was generalized to the MIMO case, leading to an efficient SD-BID algorithm for MSD of DUSTM. Several efficient implementation issues have also been addressed. The simulation results confirm the relationship between Ricean fading, Rayleigh fading and perfect CSI cases. Compared with DF-DD, the sphere decoder and SD-BID perform better and maintain a reasonably low complexity.

References

[1] D. Divsalar and M. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 300–308, Mar. 1990.

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

[3] L. Lampe, R. Schober, V. Pauli, and C. Windpassinger, "Multiple-symbol differential sphere decoding," *IEEE Trans. Commun.*, vol. 53, no. 12, pp. 1981–1985, Dec. 2005.

[4] 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, no. 7, pp. 1025–1035, July 1999.

[5] R. Schober and L. Lampe, "Noncoherent receivers for differential space-time modulation," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 768–777, May 2002.

[6] K. L. Clarkson, W. Sweldens, and A. Zheng, "Fast multiple-antenna differential decoding," *IEEE Trans. Commun.*, vol. 49, no. 2, pp. 253–261, Feb. 2001.

[7] T. Cui and C. Tellambura, "Multiple-symbol differential unitary space-time demodulation with reduced-complexity," in *Proc. IEEE Int. Conf. Commun.*, May 2005, pp. 16–20.

[8] ——, "Bound intersection detection for multiple-symbol differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 53, no. 12, pp. 2114–2123, Dec. 2005.

[9] V. Pauli and L. Lampe, "Tree-search multiple-symbol differential decoding for unitary space-time modulation," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1567–1576, Aug. 2007.

[10] R. Schober and W. H. Gerstacker, "Decision-feedback differential detection based on linear prediction for MDPSK signals transmitted over Ricean fading channels," *IEEE J. Select. Areas Commun.*, vol. 18, no. 3, pp. 391–402, Mar. 2000.

[11] D. P. Liu, Q. T. Zhang, and Q. Chen, "Structures and performance of noncoherent receivers for unitary space-time modulation on correlated fast-fading channels," *IEEE Trans. Veh. Technol.*, vol. 53, no. 4, pp. 1116–1125, July 2004.

[12] D. Makrakis, P. Mathiopoulos, and D. Bouras, "Optimal decoding of coded PSK and QAM signals in correlated fast fading channels and AWGN: a combined envelope, multiple differential and coherent detection approach," *IEEE Trans. Commun.*, vol. 42, no. 1, pp. 63–75, Jan. 1994.

[13] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2041–2052, Dec. 2000.

[14] M. Pätzold, *Mobile Fading Channels*. New York: Wiley, 2002.

[15] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge: Cambridge University Press, 1985.

[16] B. Hassibi and H. Vikalo, "On the sphere-decoding algorithm I. expected complexity," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 2806–2818, Aug. 2005.

[17] C. P. Schnorr and M. Euchner, "Lattice basis reduction: improved practical algorithms and solving subset sum problems," *Math. Programming*, vol. 66, pp. 181–191, 1994.

[18] M. V. Eyuboglu and S. U. H. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Trans. Commun.*, vol. 36, no. 1, pp. 13–20, Jan. 1988.

[19] R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Outage properties of space-time block codes in correlated Rayleigh or Rician fading environments," in *Proc. Int. Conf. Acoust., Speech Signal Process*, May 2002, pp. 13–17.