Multiple-Symbol Detection of Differential Unitary Space-Time Modulation in Fast-Fading Channels with Known Correlation

Bijoy Bhukania and Phil Schniter [1]
Department of Electrical Engineering
The Ohio State University
2015 Neil Ave, Columbus, OH 43210
E-mail: {bhukanib, schniter}@ee.eng.ohio-state.edu

Abstract — **With a fading channel, standard ML detection of differential Unitary Space-Time modulation (DUST) leads to an error floor in the BER-versus-SNR curve since it is derived under the assumption that the channel remains constant during every consecutive pair of matrix-symbols. Assuming knowledge of the channel fading correlations, we design multiple-symbol ML differential detectors which drastically reduce this error floor, especially for fast-fading channels. Multiple-symbol differential detection has the additional benefit of reducing the 3dB loss in SNR (relative to coherent detection) that characterizes standard one-symbol differential detection.**

To derive these new detectors, we make the simplifying assumption that the channel changes once per matrix-symbol (i.e., block fading) rather than once per channel use (i.e., continuous fading). However, the block fading assumption is not required when diagonal codes are used. The simulation results show that, in spite of the block-fading assumption, the new receivers far outperform standard single-symbol detection under continuous fading as well.

I. Introduction

The capacity of wireless communication systems over fading channels is significantly enhanced by the use of multiple antennas at the transmitter and/or at the receiver. Space-time coding is a bandwidth and power efficient method of communication over fading channels that realizes benefits of multiple transmit antennas [1]. Although the majority of space-time codes proposed so far assume knowledge of channel state information (CSI) at the receiver [1], the class of unitary space-time codes [2] and differential unitary space-time codes [3, 4] have been proposed for Rayleigh flat-fading channels when neither the transmitter nor the receiver knows the fading coefficients.

Differential unitary space-time modulation (DUST) can be seen as an extension of differential phase-shift keying (DPSK), commonly used in single antenna systems when the channel is unknown to the receiver as well as to the transmitter. The information is encoded in the phase difference between two consecutive symbols, so that the information can be decoded from phase difference between two consecutive observations. In DUST, the transmitted symbols can be considered $M \times M$ unitary matrices, where M is the number of transmit antennas. The previously transmitted matrix-symbol is pre-multiplied by the current information matrix-symbol to form the current transmitted matrix-symbol. The standard differential receiver detects one information matrix-symbol from each pair of consecutive received matrices.

[1] Corresponding author

Ideally, we desire that the channel experienced by every pair of consecutively transmitted matrices is the same. However, this would require that the channel be forever constant—a requirement which is never met in practice. The performance of DUST with standard single-symbol maximum-likelihood (ML) decoding in continuously-fading channel has been evaluated in [5], where it has been shown that an error floor is achieved when the error due to channel variation dominates that due to additive noise.

In this paper, we present DUST detectors for use in fast fading channels—channels for which standard single-symbol detection is unsuccessful. In deriving the new detectors, we make the simplifying assumption that the channel changes once per matrix-symbol (i.e., M channel uses), and that the receiver knows the fading correlation. We will see, however, that the block fading assumption is not necessary when diagonal space-time codes [3] are used. Under these assumptions, single- and multiple-symbol ML differential detectors are derived. Simulation results show that, in spite of the block-fading assumption, the new detectors exhibit significantly improved performance compared to the standard single-symbol detector in fast continuously-fading channels.

The notations used in this paper are as follows: Matrices will be denoted by capital letters (e.g., X and $\mathbf{X}$) and vectors by lower case bold (e.g., $\mathbf{x}$). $\mathbf{I}_N$ will denote identity matrix of size $N \times N$. The operator vec$(\cdot)$, e.g., $\mathbf{x}_n = \text{vec}(X_n)$, denotes stacking of the columns of matrix X_n in column vector $\mathbf{x}_n$. $(\cdot)^*$ denotes conjugate transposition, $\otimes$ denotes the Kronecker product, tr$(\cdot)$ denotes the trace operator, and det$(\cdot)$ the determinant.

II. Background

The system model in a continuously-fading channel is

$$X_n = \sqrt{\frac{\rho}{M}} \begin{bmatrix} \mathbf{s}_{0,n}^* & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{s}_{1,n}^* & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{s}_{M-1,n}^* \end{bmatrix} \begin{bmatrix} H_{0,n} \\ H_{1,n} \\ \vdots \\ H_{M-1,n} \end{bmatrix} + W_n \tag{1}$$

where X_n is the $M \times N$ received matrix during the n^{th} matrix-symbol interval, and where M and N are the number of transmit and receive antennas, respectively. $H_{k,n}$ is the $M \times N$ MIMO channel response matrix at the k^{th} time instant in the n^{th} matrix-symbol interval, i.e., at the $(nM+k)^{th}$ channel use. $S_n = [\mathbf{s}_{0,n}\ \mathbf{s}_{1,n}\ \dots\ \mathbf{s}_{M-1,n}]^*$ is the n^{th} $M \times M$ transmitted matrix-symbol, encoded as $S_n = V_{z_n} S_{n-1}$. $z_n \in \mathcal{L} = \{0, 1, \dots, 2^{RM}-1\}$ is the time-n integer index into matrix alphabet $\mathcal{A}$ of size 2^{RM}, so that $V_{z_n} \in \mathcal{A}$. Thus R is the number of bits per channel use. W_n is a matrix of i.i.d. unit variance Gaussian entries and ρ is the average SNR.

Under the assumption that the channel remains constant for one matrix-symbol interval, we have $H_n = H_{0,n} = \cdots = H_{M-1,n}$, which changes (1) to

$$X_n = \sqrt{\frac{\rho}{M}} S_n H_n + W_n \tag{2}$$

Even without this assumption, the use of diagonal space time codes S_n implies that the continuous fading model (1) simplifies to

$$X_n = \sqrt{\frac{\rho}{M}} S_n H_n^{(c)} + W_n \tag{3}$$

where the k^{th} row of the "equivalent continuous fading channel matrix" $H_n^{(c)}$ is the k^{th} row of $H_{k,n}$, for $k = 0, \ldots, M-1$. If the MIMO fading process $H_{k,n}$ is independent between antennas, then the equivalent fading process $H_n^{(c)}$ is also independent between antennas; the process $H_n^{(c)}$, however, is M-fold "faster" than $H_{k,n}$. Keeping this in mind, $H_n^{(c)}$ shall be denoted as H_n from here onwards, and (2) will be used as the system model. Note that (2) is an approximate model when non-diagonal codes are used in continuous fading channel.

Since $X_{n-1} = \sqrt{\frac{\rho}{M}} S_{n-1} H_{n-1} + W_{n-1}$, on realizing that $S_n = V_{z_n} S_{n-1}$, and assuming $H_n = H_{n-1}$, we can write

$$X_n = V_n X_{n-1} + \underbrace{W_n - V_{z_n} W_{n-1}}_{\tilde{W}_n} \tag{4}$$

Since V_{z_n} is unitary, $\tilde{W}_n$ contains i.i.d. Gaussian entries with variance twice of those in W_n. From (4) it is straight-forward to show that the ML detection rule for z_n is [3]

$$\hat{z}_n = \arg\max_{z_n \in \mathcal{L}} \Re[\mathrm{tr}\{X_n^* V_{z_n} X_{n-1}\}] \tag{5}$$

Due to increased noise variance in $\tilde{W}_n$, this detector loses 3dB in performance compared to coherent detection. Under the assumption of $H_n = H_{n-1}$, the Chernoff upper bound on P_e for the detector in (5) converges to zero as $\rho \to \infty$ [3]. However if the channel changes from one matrix-symbol interval to the next, i.e., if $H_n = H_{n-1} + \Delta H$, then (4) becomes

$$X_n = V_{z_n} X_{n-1} + \underbrace{\sqrt{\frac{\rho}{M}} S_n \Delta H}_{\tilde{X}_n} + W_n - V_{z_n} W_{n-1}$$

where the term $\tilde{X}_n$ creates additional "noise" that induces an error floor in BER curve. Note that the detector in (5) ignores channel variation and is therefore suboptimal in a fading environment. In following sections we derive detection rules that exploit knowledge of the autocorrelation of the time-varying channel coefficients.

III. One-Symbol ML Differential Detection

Here we derive the ML detector for z_n given X_n and X_{n-1} that exploits knowledge of the correlation between H_n and H_{n-1} (rather than assuming $H_n = H_{n-1}$ as in standard detector [3]). Denoting $\mathbf{h}_n = \mathrm{vec}(H_n), \mathbf{x}_n = \mathrm{vec}(X_n)$, and $\mathbf{w}_n = \mathrm{vec}(W_n)$,

$$\begin{aligned} \underbrace{\begin{bmatrix} \mathbf{x}_n \\ \mathbf{x}_{n-1} \end{bmatrix}}_{\underline{\mathbf{x}}_n} &= \sqrt{\frac{\rho}{M}} \begin{bmatrix} \mathbf{I}_N \otimes S_n & 0 \\ 0 & \mathbf{I}_N \otimes S_{n-1} \end{bmatrix} \begin{bmatrix} \mathbf{h}_n \\ \mathbf{h}_{n-1} \end{bmatrix} \\ &\quad + \begin{bmatrix} \mathbf{w}_n \\ \mathbf{w}_{n-1} \end{bmatrix} \end{aligned} \tag{6}$$

Assuming $\mathbf{h}_n$ and $\mathbf{w}_n$ contain zero-mean unit-variance i.i.d. Gaussian random variables, we have $E[\mathbf{h}_n \mathbf{h}_n^*] = \mathbf{I}_{MN}$. Furthermore, we assume that $E[\mathbf{h}_n \mathbf{h}_{n-k}^*] = \zeta_k \mathbf{I}_{MN}$, $k > 0$. Conditioned S_n and S_{n-1}, $\underline{\mathbf{x}}_n$ is a zero-mean Gaussian vector with autocorrelation matrix

$$R^{(1)} = \frac{\rho}{M} \begin{bmatrix} \mathbf{I}_{MN} & \zeta_1 \mathbf{I}_N \otimes V_{z_n} \\ \zeta_1^* \mathbf{I}_N \otimes V_{z_n}^* & \mathbf{I}_{MN} \end{bmatrix} + \mathbf{I}_{2MN} \tag{7}$$

where we have used the facts that S_n is unitary and $S_n = V_{z_n} S_{n-1}$. Note that the distribution of $\underline{\mathbf{x}}_n$ depends on V_{z_n} rather than S_n and S_{n-1}. Equation (7) can be rewritten

$$\begin{aligned} R^{(1)} &= \left(\frac{\rho}{M}(1 - |\zeta_1|) + 1\right) \mathbf{I}_{2MN} \\ &\quad + \frac{\rho}{M} \begin{bmatrix} \sqrt{|\zeta_1|} \\ \frac{\zeta_1^*}{\sqrt{|\zeta_1|}} \mathbf{I}_N \otimes V_{z_n}^* \end{bmatrix} \begin{bmatrix} \sqrt{|\zeta_1|} & \frac{\zeta_1}{\sqrt{|\zeta_1|}} \mathbf{I}_N \otimes V_{z_n} \end{bmatrix} \end{aligned}$$

From the identity $\det(I + AB) = \det(I + BA)$ it follows that $\det(R^{(1)})$ is independent of V_{z_n}. Note also that

$$\begin{aligned} R^{(1)\,-1} &= \frac{1}{d_1} \begin{bmatrix} \left(1 + \frac{\rho}{M}\right) \mathbf{I}_{MN} & -\frac{\rho}{M} \zeta_1 \mathbf{I}_N \otimes V_{z_n} \\ -\frac{\rho}{M} \zeta_1^* \mathbf{I}_N \otimes V_{z_n}^* & \left(1 + \frac{\rho}{M}\right) \mathbf{I}_{MN} \end{bmatrix} \\ d_1 &= \left(1 + \tfrac{\rho}{M}\right)^2 - \left(\tfrac{\rho}{M}\right)^2 |\zeta_1|^2 > 0 \end{aligned} \tag{8}$$

Thus the ML detector for z_n given X_n and X_{n-1} (equivalently given $\underline{\mathbf{x}}_n$) becomes

$$\begin{aligned} \hat{z}_n &= \arg\max_{z_n \in \mathcal{L}} p(\underline{\mathbf{x}}_n | V_{z_n}) \\ &= \arg\max_{z_n \in \mathcal{L}} e^{-\underline{\mathbf{x}}_n^* R^{(1)\,-1} \underline{\mathbf{x}}_n} \\ &= \arg\min_{z_n \in \mathcal{L}} \mathrm{tr}\Big[\tfrac{-\rho \zeta_1}{M d_1} X_n^* V_{z_n} X_{n-1} \\ &\qquad + \tfrac{-\rho \zeta_1^*}{M d_1} X_{n-1}^* V_{z_n}^* X_n \Big] \\ &= \arg\max_{z_n \in \mathcal{L}} \Re\{\mathrm{tr}[\zeta_1 X_n^* V_{z_n} X_{n-1}]\} \end{aligned} \tag{9}$$

While detector (9) is similar to detector (5), it exploits the known fading correlation ζ_1. When ζ_1 is real and positive, however, it does not affect the decision rule and can be removed, making (9) identical to (5). For reasons that will become clear later, we denote $a_{0,1}^{(1)} = \zeta_1$. In order to further improve performance we consider joint differential detection of multiple symbols in the next section.

IV. Multiple-symbol ML differential detection

Multiple-symbol differential detection of M-PSK has been proposed as an effective way to reduce the 3dB SNR loss incurred by one-symbol differential detection [6], as well as to enhance performance in correlated Rayleigh fading channels [7]. We extend this idea to DUST and incorporate knowledge of fading correlation, assuming the block-fading channel described in the previous section. First we derive the 2- and 3-symbol joint detectors and later generalize to an arbitrary number of symbols.

A. Two-Symbol ML Detection

We now collect X_{n-2}, X_{n-1}, and X_n, with the intention of

detecting z_{n-1} and z_n. Straightforward extension of (6) yields

$$\underbrace{\begin{bmatrix}\mathbf{x}_n\\ \mathbf{x}_{n-1}\\ \mathbf{x}_{n-2}\end{bmatrix}}_{\underline{\mathbf{x}}_n} = \begin{bmatrix}\mathbf{w}_n\\ \mathbf{w}_{n-1}\\ \mathbf{w}_{n-2}\end{bmatrix} + \sqrt{\tfrac{\rho}{M}}\begin{bmatrix}\mathbf{I}_N\otimes S_n & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_N\otimes S_{n-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I}_N\otimes S_{n-2}\end{bmatrix}\begin{bmatrix}\mathbf{h}_n\\ \mathbf{h}_{n-1}\\ \mathbf{h}_{n-2}\end{bmatrix}$$

Under the channel and noise assumptions of Section III, the received vector $\underline{\mathbf{x}}_n$ conditioned on S_n, S_{n-1} and S_{n-2} is zero-mean Gaussian with an autocorrelation matrix $R^{(2)}$ that depends only on $V_{z_{n-1}}, V_{z_n}, \zeta_1, \zeta_2$ and $\frac{\rho}{M}$. As before, $\det(R^{(2)})$ can be shown to be independent of V_{z_n} and $V_{z_{n-1}}$. In this case,

$$R^{(2)^{-1}} = \frac{1}{d_2}\begin{bmatrix} a_{0,0}^{(2)}\mathbf{I}_{MN} & -a_{0,1}^{(2)}\mathbf{I}_N\otimes V_{z_{n-1}} & -a_{0,2}^{(2)}\mathbf{I}_N\otimes V_{z_n}V_{z_{n-1}}\\ -a_{0,1}^{(2)^*}\mathbf{I}_N\otimes V_{z_{n-1}}^* & a_{1,1}^{(2)}\mathbf{I}_{MN} & -a_{1,2}^{(2)}\mathbf{I}_N\otimes V_{z_n}\\ -a_{0,2}^{(2)^*}\mathbf{I}_N\otimes V_{z_{n-1}}^*V_{z_n}^* & -a_{1,2}^{(2)^*}\mathbf{I}_N\otimes V_{z_n}^* & a_{2,2}^{(2)}\mathbf{I}_{MN}\end{bmatrix} \quad (10)$$

where

$$\begin{aligned}
d_2 &= \left(\tfrac{\rho}{M}+1\right)^3 - \left(\tfrac{\rho}{M}\right)^2\left(\tfrac{\rho}{M}+1\right)(2|\zeta_1|^2+|\zeta_2|^2) + \left(\tfrac{\rho}{M}\right)^3(\zeta_1^{*^2}\zeta_2+\zeta_1^2\zeta_2^*)\\
a_{0,0}^{(2)} &= a_{2,2}^{(2)} = \left(1+\tfrac{\rho}{M}\right)^2 - \left(\tfrac{\rho}{M}\right)^2|\zeta_1|^2\\
a_{1,1}^{(2)} &= \left(1+\tfrac{\rho}{M}\right)^2 - \left(\tfrac{\rho}{M}\right)^2|\zeta_2|^2\\
a_{0,1}^{(2)} &= a_{1,2}^{(2)} = \tfrac{\rho}{M}\left(\tfrac{\rho}{M}+1\right)\zeta_1 - \left(\tfrac{\rho}{M}\right)^2\zeta_1^*\zeta_2 \quad (11)\\
a_{0,2}^{(2)} &= \tfrac{\rho}{M}\left(\tfrac{\rho}{M}+1\right)\zeta_2 - \left(\tfrac{\rho}{M}\right)^2\zeta_1^2 \quad (12)
\end{aligned}$$

Hence the detection rule is given by

$$\begin{aligned}
\{\hat{z}_n, \hat{z}_{n-1}\} &= \arg\max_{z_{n-1},z_n\in\mathcal{L}} -\underline{\mathbf{x}}_n^* R^{(2)^{-1}}\underline{\mathbf{x}}_n\\
&= \arg\max_{z_{n-1},z_n\in\mathcal{L}} \Re\big[\mathrm{tr}\{a_{0,1}^{(2)}X_n^*V_{z_n}X_{n-1}\\
&\quad + a_{1,2}^{(2)}X_{n-1}^*V_{z_{n-1}}X_{n-2}\\
&\quad + a_{0,2}^{(2)}X_n^*V_{z_n}V_{z_{n-1}}X_{n-2}\}\big] \quad (13)
\end{aligned}$$

The main difference between detector (13) and one-symbol detector (9) is that (13) makes use of additional channel parameters. Thus we hope for improved performance in fast-fading channels. It is worth noticing that when the channels in subsequent blocks are independent (i.e., $\zeta_1 = \zeta_2 = 0$), then $a_{0,1}^{(2)} = a_{0,2}^{(2)} = a_{1,2}^{(2)} = 0$, implying that differential encoding cannot be used—an intuitively satisfying observation.

B. Three-symbol ML detection

We now give the structure of the 3-symbol differential detector as a prelude to the general m-symbol case and to demonstrate that closed-form expressions for the parameters $a_{k,l}^{(m)}$ become quite lengthy when $m \geq 3$. As in previous sections, we compute the inverse of the autocorrelation matrix of $\underline{\mathbf{x}}_n = [\mathbf{x}_n^*\ \mathbf{x}_{n-1}^*\ \mathbf{x}_{n-2}^*\ \mathbf{x}_{n-3}^*]^*$, conditioned on $\{S_n, S_{n-1}, S_{n-2}, S_{n-3}\}$, and find that its determinant is independent of $\{V_{z_n}, V_{z_{n-1}}, V_{z_{n-2}}\}$. This leads to the ML detector

$$\begin{aligned}
\{\hat{z}_n, \hat{z}_{n-1}, \hat{z}_{n-2}\} = \arg\max_{z_n,z_{n-1},z_{n-2}\in\mathcal{L}} \Re[\mathrm{tr}\{ \quad (14)\\
a_{2,3}^{(3)}X_{n-2}^*V_{z_{n-2}}X_{n-3} + a_{1,2}^{(3)}X_{n-1}^*V_{z_{n-1}}X_{n-2} +\\
a_{0,1}^{(3)}X_n^*V_{z_n}X_{n-1} + a_{1,3}^{(3)}X_{n-1}^*V_{z_{n-1}}V_{z_{n-2}}X_{n-3} +\\
a_{0,2}^{(3)}X_n^*V_{z_n}V_{z_{n-1}}X_{n-2} + a_{0,3}^{(3)}X_n^*V_{z_n}V_{z_{n-1}}V_{z_{n-2}}X_{n-3}\}]
\end{aligned}$$

where

$$\begin{aligned}
a_{2,3}^{(3)} &= a_{0,1}^{(3)} = \left(\left(1+\tfrac{\rho}{M}\right)^2 - \left(\tfrac{\rho}{M}\right)^2|\zeta_1|^2\right)b_1\\
a_{1,2}^{(3)} &= b_5\left(\left(\tfrac{\rho}{M}\right)^2\zeta_1^*\zeta_2 - \tfrac{\rho}{M}\left(1+\tfrac{\rho}{M}\right)\zeta_1\right) + b_1^*\left(-\left(\tfrac{\rho}{M}\right)^2\zeta_1^*\zeta_3 + \tfrac{\rho}{M}\left(1+\tfrac{\rho}{M}\right)\zeta_2\right)\\
a_{1,3}^{(3)} &= a_{0,2}^{(3)} = b_5\left(\left(\tfrac{\rho}{M}\right)^2\zeta_1^2 - \tfrac{\rho}{M}\left(1+\tfrac{\rho}{M}\right)\zeta_2\right) + b_1^*\left(\tfrac{\rho}{M}\left(1+\tfrac{\rho}{M}\right)\zeta_3 - \left(\tfrac{\rho}{M}\right)^2\zeta_1\zeta_2\right)\\
a_{0,3}^{(3)} &= b_1\left(\tfrac{\rho}{M}\left(1+\tfrac{\rho}{M}\right)\zeta_2 - \left(\tfrac{\rho}{M}\right)^2\zeta_1^2\right) - b_6\left(\tfrac{\rho}{M}\left(1+\tfrac{\rho}{M}\right)\zeta_3 - \left(\tfrac{\rho}{M}\right)^2\zeta_1\zeta_2\right)
\end{aligned}$$

and

$$\begin{aligned}
b_1 &= \left(\tfrac{\rho}{M}\right)^2(\zeta_1^{*^2}\zeta_3 + \zeta_1|\zeta_2|^2 - \zeta_1|\zeta_1|^2) - \left(\tfrac{\rho}{M}\right)^2\left(1+\tfrac{\rho}{M}\right)(\zeta_1^*\zeta_2 + \zeta_2^*\zeta_3) + \tfrac{\rho}{M}\left(1+\tfrac{\rho}{M}\right)^2\zeta_1\\
b_5 &= \left(1+\tfrac{\rho}{M}\right)^3 + 2\left(\tfrac{\rho}{M}\right)^3\Re\left(\zeta_1^*\zeta_2^*\zeta_3\right) - \left(\tfrac{\rho}{M}\right)^2\left(1+\tfrac{\rho}{M}\right)\left(|\zeta_1|^2+|\zeta_2|^2+|\zeta_3|^2\right)\\
b_6 &= \left(1+\tfrac{\rho}{M}\right)^3 + 2\left(\tfrac{\rho}{M}\right)^3\Re\left(\zeta_1^2\zeta_2^*\right)
\end{aligned}$$

Note that (14) is a straightforward extension of (13), except that the coefficients $\{a_{k,l}^{(3)}\}$ are different from $\{a_{k,l}^{(2)}\}$. As before, if $\zeta_1 = \zeta_2 = \zeta_3 = 0$, then $a_{k,l}^{(3)} = 0\ \forall k \neq l$, implying that differentially-encoded symbols cannot be detected. When $\zeta_k = 1\ \forall k$, i.e., the channel is fixed, the detection rules (13) and (14) do not depend on SNR, since in that case $a_{0,1}^{(2)} = a_{0,2}^{(2)} > 0$ and $a_{0,1}^{(3)} = a_{1,2}^{(3)} = a_{0,2}^{(3)} = a_{0,3}^{(3)} > 0$.

Though we have not derived ML detection rules for non-diagonal DUST codes in *continuously*-fading channels, the multiple-symbol detection rules (13) and (14), derived for the block-fading channel, should still show improvement over the standard detector in a continuous-fading environment. The coefficients $\{a_{i,j}^{(m)}\}$ used in this case would be recomputed with ζ_k defined such that $E[\mathbf{h}_{0,n}\mathbf{h}_{0,n-k}^*] = \zeta_k\mathbf{I}_{MN}$ for $\mathbf{h}_{0,n} = \mathrm{vec}(H_{0,n})$. The simulation results in Section VI confirm that detectors (13) and (14) significantly outperform the standard single-symbol detector (5) under fast continuous fading.

C. m-Symbol ML Detection

The ML joint detection rule for m symbols can be shown to be

$$\begin{aligned}
\{\hat{z}_n, \hat{z}_{n-1}, \ldots, \hat{z}_{n-m+1}\} = \arg\max_{z_n,z_{n-1},\ldots,z_{n-m+1}\in\mathcal{L}} \quad (15)\\
\Re\left[\mathrm{tr}\left\{\sum_{k=0}^{m-1}\sum_{i=0}^{m-k-1} a_{i,i+k+1}^{(m)}X_{n-i}^*\left(\prod_{j=i}^{i+k}V_{z_{n-j}}\right)X_{n-i-k-1}\right\}\right]
\end{aligned}$$

where $\prod_{j=0}^{m-1}(A_j) = A_0A_1\cdots A_{m-1}$. From the previous two subsections, we see that deriving closed-form expressions for

the coefficients $\{a_{k,l}^{(m)}\}$ is quite difficult when $m > 3$. These coefficients can, however, be computed numerically by first constructing an $(m+1)MN \times (m+1)MN$ autocorrelation matrix as an extension of (7), but with the additional simplification that $V_{z_{n-j}} = \mathbf{I}_M \ \forall j$, then computing its inverse. In the case of a fixed channel it can be shown that $a_{i,j}^{(m)} = a_{k,l}^{(m)} > 0 \ \forall i,j,k,l \in \{0,\ldots,m\}$ & $i \neq j, k \neq l$. As a result, the detector (15) becomes independent of coefficients $a_{i,j}^{(m)}$.

It is important to note that there does not appear to exist a Viterbi-like algorithm for m-symbol ML detection that has complexity linear in m. Observe that it is not possible to write the detection metric as a sum of terms that contain strict subsets of the symbol set $\{V_{z_n},\ldots,V_{z_{n-m+1}}\}$; there is always one term with the form $\mathrm{tr}\{X_n^*(\prod_{j=0}^{m-1} V_{z_{n-j}})X_{n-m}\}$. Thus, maximization of the quantity in (15) can only accomplished using a brute force search over all symbol combinations, yielding a complexity exponential in m.

V. Suboptimal Sequence Detection

Practical applications require the detection of $N \gg 3$ symbols. Yet, as we have seen, the joint ML differential detector for N symbols has a complexity that is exponential in N. Thus, we are motivated to consider suboptimal N-symbol detection using some combination of m-symbol ML detectors for, say, $m \leq 3$.

It is instructive to note the difference between the N-symbol ML detector and any suboptimal N-symbol detector constructed from m-symbol ML detectors ($m < N$). From (15), we see that the suboptimal sequence detector uses detection metrics that linearly combine terms based on subsets of $\{V_{z_n},\ldots,V_{z_{n-m+1}}\}$ for $n \in \{m-1,\ldots,N-1\}$. The combining coefficients are taken from the set $\{a_{k,l}^{(m)}\}$. In contrast, the optimal N-symbol detection metric linearly combines these terms with additional terms based on subsets of $\{V_{z_{N-1}},\ldots,V_{z_0}\}$ that are ignored by the suboptimal detector. In addition, the N-symbol combining coefficients are $\{a_{k,l}^{(N)}\}$, which are, in general, different from $\{a_{k,l}^{(m)}\}$. Thus, the task of constructing a "good" N-symbol detector that employs (at most) m-symbol optimal detections ($m < N$) can be considered equivalent to the approximation of $\{a_{k,l}^{(N)}\}$ by a sparse coefficient set.

For the simulations in Section VI, we divide the $(N+1)$-matrix observation sequence into consecutive subsequences of length $m+1$, each of which overlaps its neighbor by one matrix-observation. Then, m-symbol detection is performed on each subsequence. This is equivalent to replacing the coefficients $\{a_{k,l}^{(N)}\}$ with $\{a_{k,l}^{(m)}\}$ where defined, and replacing the rest with zeros. Schemes in which neighboring observation subsequences overlap by more than one matrix-observation lead to different approximations of the coefficient set $\{a_{k,l}^{(N)}\}$ and are currently under investigation.

VI. Simulations

We evaluate the performance of the detectors in two types of channel: the "block fading channel" (2) and the "continuous fading channel" (3). The correlation between fading coefficients k symbols apart is given by $J_0(2\pi f_D T_s k)$ [8] in continuous fading channels, where $f_D T_s$ is the normalized Doppler frequency. In block fading channel, correlation between channel coefficients m matrix-symbols apart is given by $J_0(2\pi f_D T_s M m)$.

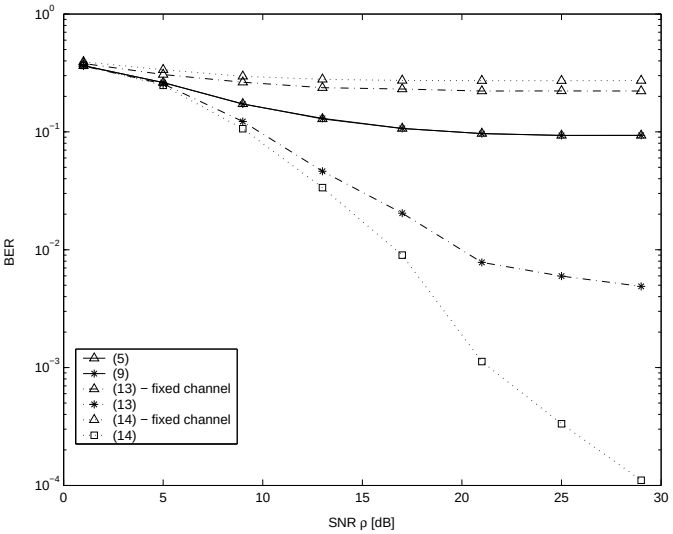

Fig. 1: Diagonal codes in continuous fading channel with $f_D T_s = 0.1$.

As shown in Section II, the use of diagonal codes in continuous-fading channels yield the same model as general codes in block-fading channels, and hence detector performance with diagonal codes is identical for these two channel types. Therefore, we first present the performance of the detectors (5) (9) (13) and (14) in continuous-fading channels using diagonal codes. The simulations assume two transmit and two receive antennas with $R = 1$ and the constellation specified in [3].

As described in Section V, detection is accomplished by detecting non-overlapping portions of the transmitted symbol sequence using 1-, 2-, and 3-symbol ML differential detectors (9), (13), and (14). Results will be presented for detectors with exact knowledge of fading correlations, as well as for detectors which assume that the channel is fixed (marked by "fixed" in the figures). In the figures shown here, solid, dashed, and dotted lines correspond to 1-, 2- and 3- symbol detection, respectively.

Fig. 1, where $f_D T_s = 0.1$, clearly illustrates the advantage of detectors which jointly detect multiple symbols *and* incorporate channel fading parameters. Note that detector (5), which ignores the fading correlation, succumbs to a very high error floor. Detector (9), designed to incorporate fading correlation into single-symbol detection, performs same as (5) does, since Rayleigh flat-fading model used in our simulations leads to $\zeta_1 > 0$. Thus, both forms of single-symbol detection perform very sub-optimally in the fading environment. The 2-symbol detector (13) which incorporates fading correlation exhibits considerably improved performance, although still succumbing to an error floor. Meanwhile, the 2-symbol detector which ignores fading correlation performs even worse than the one-symbol detector. Performance increases dramatically with the 3-symbol detector that incorporates fading, and decreases with the 3-symbol detector that ignores fading. As hinted by the plot, even the good 3-symbol detector will succumb to an error floor at high-enough SNR. The important point, however, is that the error floor has been pushed outside of the expected operating range.

Figures 2 and 3, corresponding to normalized Doppler fre-

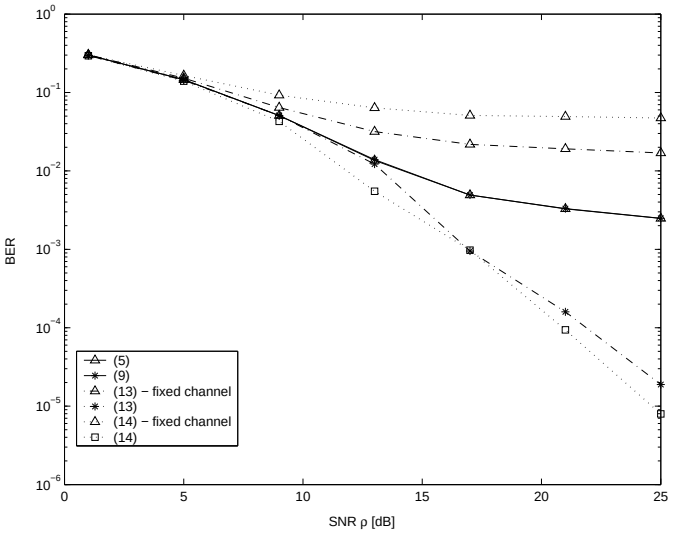

Fig. 2: Diagonal codes in continuous fading channel with $f_D T_s = 0.05$.

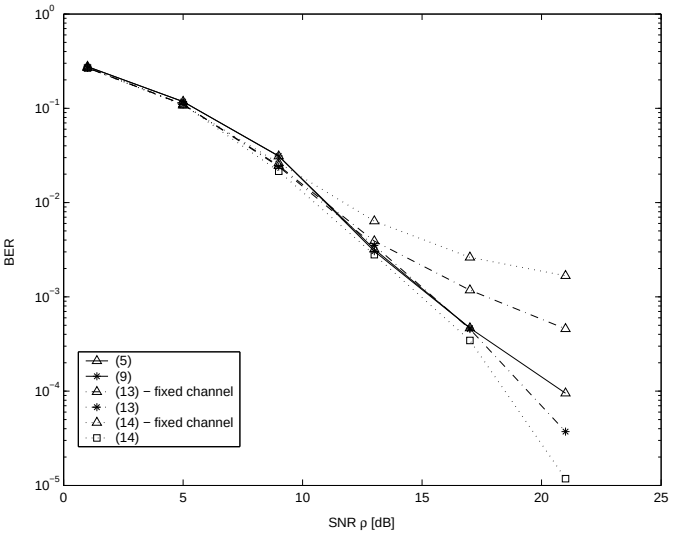

Fig. 3: Diagonal codes in continuous fading channel with $f_D T_s = 0.025$.

quencies of $f_D T_s = 0.05$ and 0.025, respectively, mimic the results of Fig. 1 but in a less pronounced fashion. Again we see the improvement associated with multiple-symbol ML differential detectors that incorporate fading correlation.

Next, the performance of the detectors has been evaluated in continuously-fading channels with non-diagonal codes to illustrate the performance loss due to approximation in the system model (2). The non-diagonal codes are generated by right multiplying the diagonal codes by a fixed non-diagonal unitary matrix. Because such an operation does not change the product distance of the constellation [3], the comparison of diagonal to non-diagonal codes is fair.

Figures 4 and 5 illustrate the performance of the detectors with non-diagonal codes in continuously-fading channel with $f_D T_s = 0.05$ and 0.025, respectively, and compares them with performance of 3-symbol detector with diagonal codes. Although the performance loss due to neglecting the channel variation within the matrix-symbol interval is significant when $f_D T_s = 0.05$ and negligible when $f_D T_s = 0.025$, in both cases 3- and 2-symbol detectors that incorporate knowledge of fading correlations perform better than the single symbol detector in terms of reducing the error floor.

The above observations lead us to important conclusions about the detectors described in this paper. First, multiple-symbol detection is essential to combat fading channels since the detectors (5) and (9) are often equivalent. Increasing the number of symbols being jointly detected improves the performance in terms of SNR loss and increases the robustness against fading channels when knowledge of fading correlations is properly incorporated into the detection rule. Generally, the faster the fading, the more symbols are required in joint detection to push the error floor out of the operating range.

VII. Conclusions

In this paper, we have demonstrated the efficacy of multiple-symbol ML differential detection that incorporates channel fading parameters in combating the error floor exhibited by standard one-symbol ML detection of DUST. In fact, we have shown that multiple-symbol (versus single-symbol) detection is essential to performance enhancement. Our multiple-symbol detection rules, which assume the channels to be block-fading when non-diagonal codes are used, have been shown to improve performance in continuously-fading channels as well. We are currently investigating the robustness of these detectors to imperfect knowledge of fading correlation and SNR as well as reduced complexity implementation of the multiple-symbol detectors for large m. In addition, we are working to derive theoretical error bounds for these detectors.

References

[1] A. Naguib, N. Seshadri, and A.R. Calderbank, "Increasing data rate over wireless channels", *IEEE Signal Processing Magazine*, vol. 17, pp. 76-92, May 2000.

[2] B.M. Hochwald and T.L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading", *IEEE Trans. on Information Theory*, vol. 46, no. 2, pp. 543-564, Mar. 2000.

[3] B.M. Hochwald and W. Sweldens, "Differential unitary space-time modulation", *IEEE trans. on Communications*, vol. 48, no. 12, pp. 2041-2052, Dec. 2000.

[4] B.L. Hughes, "Differential space-time modulation", *IEEE Trans. on Information Theory*, vol. 46, no. 7, pp. 2567-2578, Nov. 2000.

[5] C. Peel and A. Swindlehurst, "Performance of unitary space-time modulation in continuously changing channel", in *IEEE Internat. Conf. on Acoustics, Speech, and Signal Processing*, 2001.

[6] D. Divsalar and M.K. Simon, "Multiple-symbol differential detection of MPSK", *IEEE trans. on Communications*, vol. 38, no. 3, pp. 300-308, March 1990.

[7] P. Ho and D. Fung, "Error performance of multiple symbol differential detection of PSK signals transmitted over correlated Rayleigh fading channels", *IEEE Intern. Conf. on Communications*, vol. 2, pp. 568-574, 1991.

[8] W.C. Jakes, *Microwave Mobile Communications*, IEEE press, Piscataway, NJ, 1993.

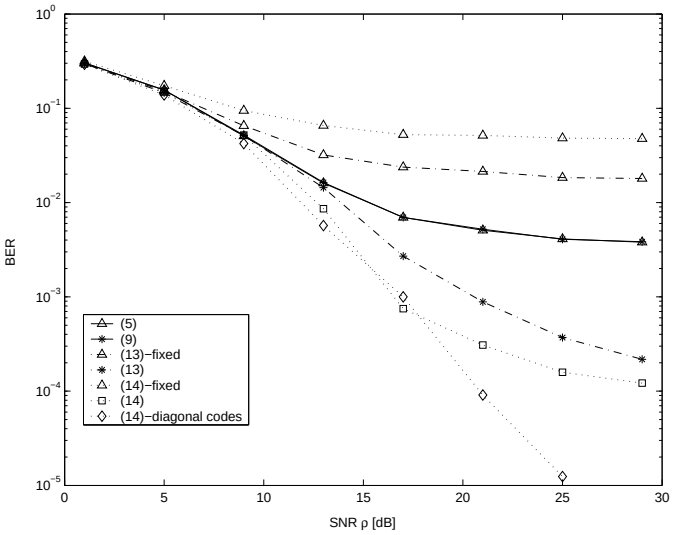

Fig. 4: Non-diagonal codes in continuously-fading channel with $f_D T_s = 0.05$.

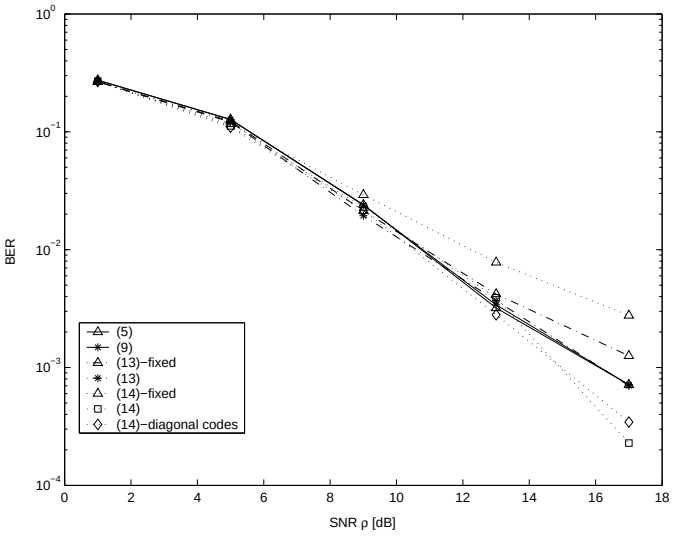

Fig. 5: Non-diagonal codes in continuously-fading channel with $f_D T_s = 0.025$.