Maximum Likelihood Detection and Optimal Code Design for Differential Unitary Space-Time Modulation with Carrier Frequency Offset

Tao Cui
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA 91125, USA
Email: taocui@caltech.edu

Feifei Gao and Arumugam Nallanathan
Department of Electrical
and Computer Engineering
National University of Singapore
Singapore 119260
Email: {feifeigao, elena}@nus.edu.sg

Chintha Tellambura
Department of Electrical
and Computer Engineering
University of Alberta
Edmonton, AB T6G 2V4, Canada
Email: chintha@ece.ualberta.ca

Abstract— In this paper, we answer the question that "Can conventional differential unitary space time modulation (DUSTM) be applied when there is an unknown carrier frequency offset (CFO)?" and present a maximum likelihood (ML) detection rule for this scenario. We analyze the asymptotical performance of our ML detection and provide the code design criterion by using the modified diversity product. The analysis also brings the insight that our proposed decision rule is a new differential modulation scheme in both temporal and spatial domains. Various simulations are conducted, and the proposed algorithm is shown to be more robust to the CFO drifting than the existing double temporal differential modulation.

I. INTRODUCTION

Wireless multiple-input multiple-output (MIMO) systems with space time coding (STC) [1] have received much attention due to the potential of increasing the transmission rate and combating fading. Hochwald and Marzetta [2] investigate non-coherent detection for space time modulation on MIMO systems which release the requirement of channel state information (CSI). Following this pioneering work, Hochwald [3] and Huges [4] have developed a general framework for differential unitary space time modulation (DUSTM) that is viewed as a generalization of the classical single antenna differential phase shift keying (DPSK).

In wireless communications, carrier frequency offset (CFO) always arises due to the transceiver oscillator mismatch or the Doppler shift caused by the relative motion between the transmitter and the receiver. The performance of DUSTM degrades greatly if the non-zero CFO is not compensated at the receiver before symbol detection. Recently, Liu [5] and Ma [6] proposed a new modulation scheme for MIMO and SISO-OFDM systems that could deal with both unknown CSI and CFO. This method applies a double differential modulation (DDUSTM) in time domain, where the effect of the unknown CFO is removed by double temporal differentiation.

Although DDUSTM provides an effective way to deal with CFO, a natural question arises: *Can conventional DUSTM be applied with an unknown CFO?* In this paper, we give an affirmative answer to this question. We find that by sacrificing one degree of freedom in transmit antennas, the ML single-symbol detector (SSD), using only two consecutive blocks, could be derived for DUSTM even with the unknown CFO. We name the proposed modulation scheme as modified DUSTM (MDUSTM) throughout the rest of this paper. The asymptotical performance of MDUSTM is analyzed and an optimal MDUSTM code design criterion is derived. Since MDUSTM only requires constant CFO and channel for two consecutive blocks, simulation results show that our MDUSTM is more robust to the CFO drifting than DDUSTM that requires constant CFO and channel for three consecutive blocks.

II. SYSTEM MODEL

Consider a MIMO system with N_t transmit and N_r receive antennas that operates in a Rayleigh flat fading environment. Each time slot occupies an interval T_s in seconds, and each block is consisted of T time slots. The block interval is thus $T_s T$ in seconds. The transmitted symbols during the kth block is denoted by the $T \times N_t$ matrix $\mathbf{S}[k] = [s_{t,i}[k]]$, $t = 1, \ldots, T$ and $i = 1, \ldots, N_t$, where $s_{t,i}[k]$ is transmitted from the ith antenna in the $t + (k-1)T$ time slot. 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 kth block can be written as

$$x_{t,j}[k] = \sqrt{\rho} \sum_{i=1}^{N_t} h_{i,j}[t + (k-1)T]s_{t,i}[k] + w_{t,j}[k], \quad (1)$$
$$j = 1, \ldots, N_r, \ t = 1, \ldots, T,$$

where $h_{i,j}[t] \in \mathcal{CN}(0,1)$ denotes the channel gain from the i-th transmit antenna to the j-th receive antenna in the t-th time slot, and $w_{t,j}[k] \in \mathcal{CN}(0,1)$ is the complex additive white Gaussian noise (AWGN) at the j-th receive antenna which is statistically independent across the receiver and time index. We assume that all path gains are statistically independent ($E\{h_{i,j}[k]h_{i',j'}^*[k]\} = \delta(i-i')\delta(j-j')$) and have the same autocorrelation function $\varphi_h[n] = E\{(h_{i,j}[t+n]h_{i,j}[t]^*\}$.

Typically, when Jakes' model [7] is used, $\varphi_h[n]$ is given by

$$\varphi_h[n] = J_0(2\pi f_D T_s n), \qquad (2)$$

where f_D is the Doppler spread due to user mobility, and $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind. The power of transmitted symbols is normalized such that

$$\mathrm{E}\left\{\sum_{i=1}^{N_t} |s_{t,i}[n]|^2\right\} = 1. \qquad (3)$$

Therefore, ρ is the expected SNR at each receive antenna.

We consider DUSTM in this paper. In [3], 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 $(\mathbf{V}_l \mathbf{V}_l^H = \mathbf{I}_T)$, $L = 2^{N_t R}$, and R denotes the data rate. To make DUSTM feasible, we take $T = N_t$ 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 $\mathbf{V}[k] = \mathbf{V}_l$ is chosen from $\mathcal{V}$. The transmitted symbol at the nth block is encoded as

$$\mathbf{S}[k] = \mathbf{V}[k]\mathbf{S}[k-1]. \qquad (4)$$

In the first block, $\mathbf{S}[0] = \mathbf{V}_0$ is sent. The internal composition property of a group ensures that $\mathbf{S}[k] \in \mathcal{V}$ and is unitary for any positive k. In this paper, we consider diagonal constellations of the cyclic group, for which the unitary matrices $\mathbf{V}_l$ are chosen as

$$\mathbf{V}_l = \mathrm{diag}\left\{e^{j2\pi u_1 l/L}, e^{j2\pi u_2 l/L}, \ldots, e^{j2\pi u_{N_t} l/L}\right\}, \qquad (5)$$

where u_i for $i = 1, 2, \ldots, N_t$ are optimized to achieve the maximum diversity product in [3]. We thus have $s_{t,i}[k] \neq 0$ $t = i$ and $s_{t,i}[k] = 0$ for $t \neq i$. Therefore, (1) can be rewritten as

$$\mathbf{X}[k] = \mathbf{S}[k]\mathbf{H}[k] + \mathbf{W}[k], \qquad (6)$$

where $\mathbf{X}[k] = [x_{i,j}[k]]$ is the $N_t \times N_r$ receive matrix, $\mathbf{W}[k] = [w_{i,j}[k]]$ is the $N_t \times N_r$ noise matrix, $\mathbf{H}[k]$ is the $N_t \times N_r$ channel matrix, and the (i,j)-th entry of $\mathbf{H}[k]$ is $h_{i,j}[i + (k-1)N_t]$.

Suppose the CFO is constant in one block but may vary from block to block due to the CFO drifting [6]. Let ε_k denote the normalized CFO by the sampling period T_s during the kth block. Then, the phase distortion on the $i + (k-1)N_t$ time slot is

$$\theta_{i,k} = 2\pi\varepsilon_k, \quad i = 1, \ldots, N_t, k = 0, 1, \ldots. \qquad (7)$$

Therefore, the overall phase rotation before the $i + (k-1)N_t$ time slot, denoted as $\vartheta_{i,k}$, is that accumulated from all the previous time slots:

$$\vartheta_{i,k} = \sum_{\tau < i} \sum_{c < k} \theta_{\tau,c} = 2\pi N_t \sum_{c=0}^{k-1} \varepsilon_c + (i-1)\varepsilon_k. \qquad (8)$$

If CFO does not vary, the accumulated phase $\vartheta_{i,k}$ becomes $2\pi(N_t k + i - 1)\varepsilon_0$, as usually the case. The receive signal block $\mathbf{X}[k]$ with the unknown CFO is then modeled as

$$\mathbf{X}[k] = \sqrt{\rho}e^{j\vartheta_{1,k}}\mathbf{\Gamma}(\varepsilon_k)\mathbf{S}[k]\mathbf{H}[k] + \mathbf{W}[k], \qquad (9)$$

where $\mathbf{\Gamma}(\varepsilon_k)$ is the diagonal matrix with the form

$$\mathbf{\Gamma}(\varepsilon_k) = \mathrm{diag}\{1, e^{j2\pi\varepsilon_k}, \ldots, e^{j2\pi(N_t-1)\varepsilon_k}\}. \qquad (10)$$

III. ML DETECTION OF DUSTM WITH CFO

In this paper, we only consider single symbol detection (SSD) while multiple symbol detection (MSD) is presented in [8]. For SSD, we assume that both the CFO and the channel do not vary across two blocks, say the $(k-1)$th block and the kth block, for ease of exposition[1]. We relax this channel constraint in Section III. B. Hence, the indices $k-1$ and k are dropped in both the channel matrix and the CFO. The received signal blocks may be rewritten as

$$\mathbf{X}[k-1] = \sqrt{\rho}e^{j\vartheta_{1,k}}\mathbf{\Gamma}(\varepsilon)\mathbf{S}[k-1]\mathbf{H} + \mathbf{W}[k-1], \qquad (11)$$

$$\mathbf{X}[k] = \sqrt{\rho}e^{j(\vartheta_{1,k}+2\pi N_t \varepsilon)}\mathbf{\Gamma}(\varepsilon)\mathbf{S}[k]\mathbf{H} + \mathbf{W}[k] \qquad (12)$$

$$= \left(e^{j2\pi N_t \varepsilon}\mathbf{V}[k]\right)\left(\sqrt{\rho}e^{j\vartheta_{1,k}}\mathbf{\Gamma}(\varepsilon)\mathbf{S}[k-1]\mathbf{H}\right) + \mathbf{W}[k].$$

Let $\bar{\mathbf{S}}[k] = [\mathbf{S}^T[k-1], \mathbf{S}^T[k]]^T$ and $\bar{\mathbf{X}}[k] = [\mathbf{X}^T[k-1], \mathbf{X}^T[k]]^T$. The probability density function (PDF) of $\bar{\mathbf{X}}[k]$ could be expressed as

$$f(\bar{\mathbf{X}}[k]|\mathbf{V}[k], \varepsilon, \vartheta_{1,k}) = \frac{\exp(-tr(\mathbf{\Lambda}^{-1}\bar{\mathbf{X}}[k]\bar{\mathbf{X}}^H[k]))}{\pi^{2N_t N_r}\det^{N_r}\{\mathbf{\Lambda}\}}, \qquad (13)$$

where

$$\mathbf{\Lambda} = \mathbf{I}_{2N_t} + \rho\left[\begin{array}{cc} \mathbf{I}_{N_t} & e^{-j2\pi N_t \varepsilon}\mathbf{V}^H[k] \\ e^{j2\pi N_t \varepsilon}\mathbf{V}[k] & \mathbf{I}_{N_t} \end{array}\right]. \qquad (14)$$

Since $f(\bar{\mathbf{X}}[k]|\mathbf{V}[k], \varepsilon, \vartheta_{1,k})$ does not depend on $\vartheta_{1,k}$, we can rewrite it as $f(\bar{\mathbf{X}}[k]|\mathbf{V}[k], \varepsilon)$. After some algebra, the ML estimate can be obtained as

$$\{\hat{\varepsilon}, \hat{\mathbf{V}}[k]\} = \arg\max_{\varepsilon, \mathbf{V}[k]} f(\bar{\mathbf{X}}[k]|\mathbf{V}[k], \varepsilon)$$

$$= \arg\max_{\varepsilon, \mathbf{V}[k]} K\exp\left[\Re\{e^{j2\pi N_t \varepsilon}tr(\mathbf{V}[k]\mathbf{X}[k-1]\mathbf{X}^H[k])\}\right], (15)$$

where K contains all the factors of $f(\bar{\mathbf{X}}[k]|\mathbf{V}[k], \varepsilon)$ that are independent of $\mathbf{V}[k]$ and ε.

We first fix $\mathbf{V}[k]$. By maximizing (15) over ε, the ML estimate of ε is given by

$$\hat{\varepsilon} = -\frac{1}{2\pi N_t}\angle tr(\mathbf{V}[k]\mathbf{X}[k-1]\mathbf{X}^H[k]) + \frac{n}{N_t}, \qquad (16)$$

where $n \in \{0, \ldots, N_t - 1\}$ is the integer ambiguity related to ε. Substituting (16) back into (15), the ML detection (MLD) for $\mathbf{V}[k]$ is obtained as

$$\hat{\mathbf{V}}[k] = \arg\max_{\mathbf{V}[k]} |tr(\mathbf{V}[k]\mathbf{X}[k-1]\mathbf{X}^H[k])|. \qquad (17)$$

Clearly, the ambiguity n does not affect the MLD of $\mathbf{V}[k]$.

[1]In comparison, DDUSTM assume constant CFO and channel across three blocks.

IV. PERFORMANCE ANALYSIS AND CODE DESIGN

A. Pairwise Error Probability Analysis

Assuming the codeword $\mathbf{V}_l \in \mathcal{V}$ was transmitted, the pairwise error probability (PEP) that the maximum likelihood detector (17) decides $\mathbf{V}_{l'} \in \mathcal{V}$, rather than $\mathbf{V}_l$, when only these two codewords are possible, is given by

$$
P_e\left(\mathbf{V}_l \to \mathbf{V}_{l'}\right) = P\left\{\left|tr\left(\mathbf{V}_l \mathbf{X}[k-1]\mathbf{X}^H[k]\right)\right| \right.
$$
$$
\left. < \left|tr\left(\mathbf{V}_{l'} \mathbf{X}[k-1]\mathbf{X}^H[k]\right)\right|\ \Big|\ \mathbf{V}_l \text{ transmitted}\right\}. (18)
$$

Similar definition holds for $P_e\left(\mathbf{V}_{l'} \to \mathbf{V}_l\right)$. It can be easily known that $P_e\left(\mathbf{V}_{l'} \to \mathbf{V}_l\right) = P_e\left(\mathbf{V}_l \to \mathbf{V}_{l'}\right)$. Therefore we will denote the PEP as P_e for brevity, whose exact form is proved in [8] to be

$$
P_e = \mathrm{E}_{\mathbf{H}}\left\{P_{e|\mathbf{H}}\right\}
$$
$$
= \mathrm{E}_{\mathbf{H}}\left\{\frac{1}{2}[1 + Q(\sqrt{a},\sqrt{b}) - Q(\sqrt{b},\sqrt{a})]\right\}, \quad (19)
$$

where

$$
Q(\alpha,\beta) = \int_\beta^\infty \exp\left(-\frac{\alpha^2 + x^2}{2}\right) I_0(\alpha x) x\, dx \quad (20)
$$

is the Marcum Q function, and $a, b, \mathbf{Q}$ are defined as

$$
\left\{\begin{array}{c} a \\ b \end{array}\right\} = \frac{\rho\|\mathbf{H}\|_F^2}{4}\left(1 \mp \sqrt{1 - \frac{|tr(\mathbf{H}^H\mathbf{Q}\mathbf{H})|^2}{\|\mathbf{H}\|_F^4}}\right), \quad (21)
$$
$$
\mathbf{Q} \triangleq \mathbf{V}_l^H \mathbf{V}_{l'}. \quad (22)
$$

This expression of P_e is in fact related to the fourth order Gaussian statistics and is currently intractable.

In fact, Marcum Q function can be approximated in several ways, for example, as those in [9] and [10]. However, we will adopt a more general way by utilizing the following inequalities. From [11], we know $Q(\alpha,\beta) < \exp\left(-\frac{(\beta-\alpha)^2}{2}\right)$ when $\beta > \alpha \geq 0$, and $Q(\alpha,\beta) > 1 - \frac{1}{2}\left[\exp\left(-\frac{(\alpha-\beta)^2}{2}\right) - \exp\left(-\frac{(\alpha+\beta)^2}{2}\right)\right]$ when $\alpha > \beta \geq 0$. Then

$$
P_{e|\mathbf{H}} \leq \frac{1}{4}\left[3\exp\left(-\frac{(\sqrt{b}-\sqrt{a})^2}{2}\right) - \exp\left(-\frac{(\sqrt{a}+\sqrt{b})^2}{2}\right)\right]
$$
$$
\leq \frac{3}{4}\exp\left(-\frac{(\sqrt{b}-\sqrt{a})^2}{2}\right). \quad (23)
$$

Substituting the specific forms of a, b into (23) results in

$$
P_{e|\mathbf{H}} \leq \frac{3}{4}\exp\left(-\frac{\rho}{4}\left(\|\mathbf{H}\|_F^2 - \left|tr\left(\mathbf{H}^H\mathbf{Q}\mathbf{H}\right)\right|\right)\right). \quad (24)
$$

The unfriendly factor $|tr(\mathbf{H}^H\mathbf{Q}\mathbf{H})|$ forbids the further derivation on the conditioned PEP bound (24). To simplify the analysis, we only consider the asymptotic performance as with [12], [13]. We consider the following two cases:

1) $N_r \to \infty$ and N_t is finite: Similar to [12], we assume that the variance of h_{ij} is $1/N_r$, because the normalizing factor $1/N_r$ ensures the total power received by N_r antennas from each transmit antenna remains constant as N_r approaches ∞. By the strong law of large numbers, we have $\mathbf{H}\mathbf{H}^H \approx \mathbf{I}_{N_t}$ and $|tr(\mathbf{H}^H\mathbf{Q}\mathbf{H})| \approx |tr(\mathbf{Q})|$. Then

$$
P_{e|\mathbf{H}} \leq \frac{3}{4}\exp\left(-\frac{\rho}{4}(N_t - |tr(\mathbf{Q})|)\right). \quad (25)
$$

Now, $P_{e|\mathbf{H}}$ is independent from $\mathbf{H}$, then the upper bound on P_e is

$$
P_e \leq \frac{3}{4}\exp\left(-\frac{\rho}{4}(N_t - |tr(\mathbf{Q})|)\right). \quad (26)
$$

2) $N_t \to \infty$ and N_r is finite: First of all, it is noted that

$$
\|\mathbf{H}\|_F^2 - |tr(\mathbf{H}^H\mathbf{Q}\mathbf{H})| = \sum_{j=1}^{N_r}\sum_{i=1}^{N_t}|h_{ij}|^2 - \left|\sum_{j=1}^{N_r}\sum_{i=1}^{N_t}q_i|h_{ij}|^2\right|
$$
$$
\geq \sum_{j=1}^{N_r}\left(\sum_{i=1}^{N_t}|h_{ij}|^2 - \left|\sum_{i=1}^{N_t}q_i|h_{ij}|^2\right|\right). \quad (27)
$$

Since h_{ij} are independent and identically distributed random variables over (i,j), the upperbound on P_e could be expressed as

$$
P_e = \mathrm{E}\{P_{e|\mathbf{H}}\}
$$
$$
\leq \frac{3}{4}\mathrm{E}^{N_r}\left\{\exp\left(-\frac{\rho}{4}\left(\sum_{i=1}^{N_t}|h_{i1}|^2 - \left|\sum_{i=1}^{N_t}q_i|h_{i1}|^2\right|\right)\right)\right\} (28)
$$

From the free probability theory [12, Fact A.2 or Eq. (24)], we know

$$
\sum_{i=1}^{N_t}q_i|h_{i1}|^2 \to \frac{1}{N_t}tr(\mathbf{Q})\sum_{i=1}^{N_t}|h_{i1}|^2. \quad (29)
$$

Then, P_e can be asymptotically bounded by

$$
P_e \leq \frac{3}{4}\mathrm{E}^{N_r}\left\{\exp\left(-\frac{\rho}{4}\left(\sum_{i=1}^{N_t}\left(1 - \frac{1}{N_t}|tr(\mathbf{Q})|\right)|h_{i1}|^2\right)\right)\right\}
$$
$$
= \frac{3}{4}\mathrm{E}^{N_t N_r}\left\{\exp\left(-\frac{\rho}{4}\left(1 - \frac{1}{N_t}|tr(\mathbf{Q})|\right)|h_{11}|^2\right)\right\}
$$
$$
= \frac{3}{4}\left(\frac{1}{1 + \frac{\rho}{4}\left(1 - \frac{1}{N_t}|tr(\mathbf{Q})|\right)}\right)^{N_t N_r}. \quad (30)
$$

B. Code Design

To minimize the bound of P_e in both cases, it is clear from (26) and (30) that we should minimize $|tr(\mathbf{Q})|$. As opposed to the definition in [3], we define the *modified diversity product* ζ as

$$
\zeta = \max_{0 \leq l < l' \leq L-1}\left|tr\left(\mathbf{V}_l^H\mathbf{V}_{l'}\right)\right|. \quad (31)
$$

Therefore, the designing criterion is to minimize the modified diversity product.

For cyclic group codes as in [3], the unitary matrices $\mathbf{V}_l$ are chosen as in (5), where u_i for $i = 1, \ldots, N_t$ need to

N_t	R	L	$[u_1, u_2, ..., u_{N_t}]$	N_t	R	L	$[u_1, u_2, ..., u_{N_t}]$
2	1	4	[0, 3]	2	1.5	8	[0, 5]
3	1	8	[0, 1, 3]	4	1.5	64	[0, 11, 55, 59]
4	1	16	[0, 2, 11, 15]	2	2	16	[0, 13]
5	1	32	[0, 15, 25, 26, 28]	4	2	256	[0, 131, 135, 186]

be optimized to achieve minimum modified diversity product (31). In this case, $\mathbf{Q}$ can be expressed as

$$\mathbf{Q} = \text{diag} \left\{ e^{j2\pi u_1(l-l')/L}, ..., e^{j2\pi u_{N_t}(l-l')/L} \right\}. \quad (32)$$

Denote $l - l'$ as Δl. The modified diversity product ζ can be simplified as

$$\zeta = \max_{0 \leq \Delta l \leq L-1} \left| \sum_{i=1}^{N_t} e^{j2\pi u_i \Delta l/L} \right|. \quad (33)$$

We resort to exhaustive computer search for the best set of $u_1, \ldots, u_{N_t} \in \{0, ..., L-1\}$. The search space can be reduced using the following rules:

1) Note that $|tr(\mathbf{Q})| = |e^{-j2\pi u_1 \Delta l/L} tr(\mathbf{Q})|$. Therefore, we can always set $u_1 = 0$.
2) The order of u_i does not affect the metric, so we always assume that $u_{i_2} > u_{i_1}$, if $i_2 > i_1$.
3) In (33), Δl and $L - \Delta l$ give the same ζ. Hence we only need to search Δl from 1 to $\left\lfloor \frac{L}{2} \right\rfloor$.

Table I shows the results of our search for parameters that minimize ζ.

Remarks:

- When $\mathbf{Q} = \gamma \mathbf{I}_{N_t}$ and γ is a complex scalar with unit norm, it can be readily obtained that $|tr(\mathbf{Q})| = N_t$, which leads to the worst PEP. It is interesting to see that the traditional designed cyclic group codes [3] are among those who provide the worst PEP. For example, if we choose $l = l_1$ and $l' = l_1 + L/2$, $\Delta l = l' - l = L/2$ and $\left| tr\left(\mathbf{V}_l^H \mathbf{V}_{l'}\right) \right| = \left| \sum_{i=1}^{N_t} e^{j2\pi u_i \Delta l/L} \right| = \left| \sum_{i=1}^{N_t} e^{j\pi u_i} \right|$. From [3], all u_i are odd integers [3]. Therefore, we have $\left| tr\left(\mathbf{V}_l^H \mathbf{V}_{l'}\right) \right| = N_t$, and $\zeta = N_t$.
- In fact, our MLD can be considered as employing differential modulation across the transmit antennas, which is a new differential modulation scheme. Our scheme applies differential modulation in both spatial and temporal domains as opposed to temporal domain only in DDUSTM [5]. DDUSTM needs that the CFO remains constant over three blocks to detect the transmitted signals, while we only need it be constant in two blocks. This suggests that our method is more robust to CFO drifting than that in [5].

V. SIMULATION RESULTS

In this section, we simulate our proposed MDUSTM scheme using symbol error rate (SER) as our figure of merit under various scenarios. The signal transmitted in the first block is chosen as $\mathbf{S}[0] = \mathbf{I}_{N_t}$.
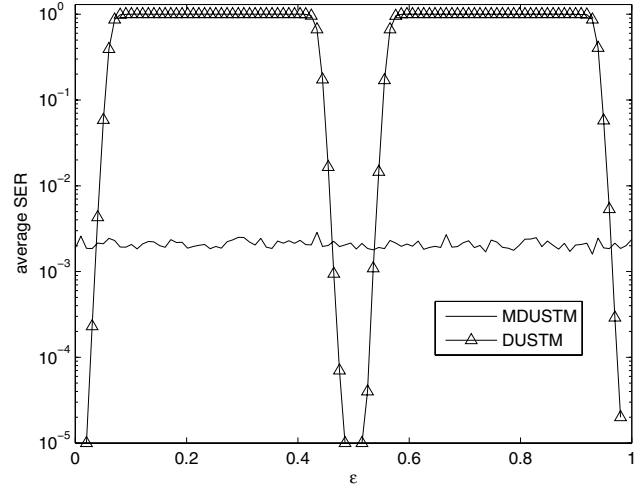

Fig. 1. Average SER versus CFO for $N_t = N_r = 2$ (SNR=12dB).

1) Performance comparison with DUSTM under frequency errors. In this example, we consider a 2×2 MIMO system and investigate the performance of both the MDUSTM scheme and the DUSTM without compensating the frequency errors. The maximum range of ε, $[0, 1]$, is examined. For both schemes, their respective optimal codes with $R = 1$ ($\mathbf{u} = [0, 3]$, and $\mathbf{u} = [1, 1]$) are employed. The simulation results are shown in Fig. 1 for SNR=12 dB. We note that, when CFO is small, DUSTM without CFO compensation performs better. However, for a general range of CFO, the DUSTM totally fails while MDUSTM is still effective. Similar phenomenon has been observed both in [5] and [6].

2) Performance comparison of the designed optimal codes with all other codes. In this example, we choose $N_t = 3$, $N_r = 4$ and adopt the cyclic group codes for demonstration. The performance line of our designed code $\mathbf{u} = [0, 1, 3]$, as well as those of all other possible $\mathbf{u}$ are drawn in Fig. 2. Interestingly, all codes could be divided into several groups and within each group the codes provide similar performances. However, the performances from group to group are quite different. It can be seen that the designed asymptotically optimal codes lies in the bunch that gives the best performance. In fact, only one choice of $\mathbf{u} = [0, 1, 6]$, in this example, outperforms the designed code. This is not unexpected since the optimal code for regular number of antennas may not necessarily be the asymptotic codes.
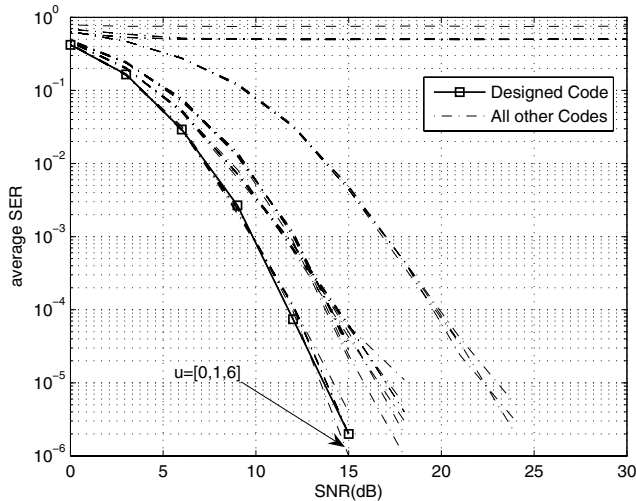
Fig. 2. Demonstration of the optimality of the designed code under regular transceiver antennas.

Fig. 3. Performance comparison between SSD and DDUSTM under various CFO drifting, $N_t = 3$, $N_r = 4$.

3) Performance comparison with normal DDUSTM under frequency drifting. We apply the optimal cyclic group code for both MDUSTM and DDUSTM in this example. Since MDUSTM sacrifices one degree of freedom in transmit antenna to cope with CFO, it is expected that the DDUSTM would outperform MDUSTM at high SNR due to its higher diversity order.

Nevertheless, it can be found that the proposed scheme is highly robust for frequency drifting than DDUSTM. Similar to [6], the CFO drifting is modeled as a random process from block to block but the drifting is zero inside each block. Suppose the CFO at the kth block is ε_k. For the $(k + 1)$st block, the CFO becomes

$$\varepsilon_{k+1} = \varepsilon_k + \Delta\varepsilon_k, \qquad (34)$$

where $\Delta\varepsilon_k$ is drawn uniformly from $[-0.5\Delta\varepsilon_{\max}, 0.5\Delta\varepsilon_{\max}]$. Four different values of $\Delta\varepsilon_{\max}$ are chosen and the performance comparison between MDUSTM and DDUSTM are plotted in Fig. 3. The parameters are selected as $N_t = 3$, $N_r = 4$, $\varepsilon_0 = 0.4/N_t$. Both schemes employ their respective optimal code design ($\mathbf{u} = [0, 1, 3]$ and $\mathbf{u} = [1, 1, 3]$). It can be seen that with a smaller CFO drifting, both schemes are affected slightly. However, when $\Delta\varepsilon_{\max}$ goes beyond $0.05\varepsilon_0$, the performance of DDUSTM degrades significantly. Especially for very large drifting $\Delta\varepsilon_{\max} = 0.1\Delta\varepsilon_0$, the performance of DDUSTM is almost unacceptable.

VI. CONCLUSIONS

In this paper, we developed ML detection methods for DUSTM under the existence of unknown CFO. The proposed MDUSTM removes the effect of unknown CFO by the spatial differentiation as opposed to the traditional way that operates through the temporal domain. We have derived the asymptotic pairwise symbol error rate analysis and, accordingly, designed the asymptotically optimal code for large number of antennas,
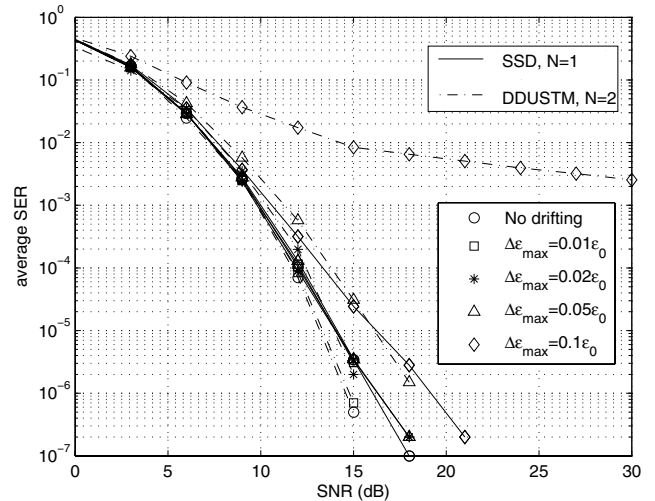
whereas these asymptotic codes also work well when the number of transceiver antennas is small. Since the proposed MDUSTM only requires a slowly varying channel and CFO during only two blocks, it is more robust to CFO drifting than DDUSTM.

REFERENCES

[1] N. Al-Dhahir, N., C. Fragouli, A. Stamoulis, W. Younis, and R. Calderbank, "Space-time processing for broadband wireless access Communications," *IEEE Commun. Mag.*, vol. 40, pp. 136-142, Sept. 2002.

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

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

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

[5] Z. Liu, G.B. Giannakis, and B.L. Hughes, "Double differential space-time block coding for time selective fading channels," *IEEE Trans. Commun.*, vol. 49, pp. 1529-1539, Sept. 2001.

[6] X. Ma, "Low complexity block double differential design for OFDM with carrier frequency offset," *IEEE Trans. Commun.*, vol. 53, pp. 2129-2138, Dec. 2005.

[7] G. L. Stuber, *Principles of Mobile Communication*, 2nd ed. Norwell, MA: Kluwer Academic, 2001.

[8] F. Gao, T. Cui, A. Nallanathan, and C. Tellambura, "Maximum Likelihood Detection for Differential Unitary Space-Time Modulation with Carrier Frequency Offset," submitted to *IEEE Trans. Commun.*.

[9] S. Stein, "Unified analysis of certain coherent and noncoherent binary communications systems," *IEEE Trans. Inform. Theory*, pp. 43-51, Jan. 1964.

[10] D. Divsalar, and M.K. Simon, "Multiple symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 38, pp. 300-308, Mar. 1990.

[11] M.K. Simon, and M.S. Alouini, "Exponential-type bounds on the generalized Marcum Q-function with application to error probability analysis over fading channels," *IEEE Trans. Commun.*, vol. 48, pp. 359-366, Mar. 2000.

[12] E. Biglieri, G. Taricco, and A. Tulino, "Performance of space-time codes for a large number of antennas," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1794-1803, July, 2002.

[13] J. Wang, M.P. Fitz, and K. Yao, "Differential unitary space-time modulation for a large number of receive antennas," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Nov. 2002, pp. 565-569.