Effects of Non-Identical Rayleigh Fading on Differential Unitary Space-Time Modulation

Meixia Tao, *Member, IEEE*

***Abstract*—Non-identical fading distribution in a multiple-input multiple-output (MIMO) channel, including unequal average channel gains and fade rates, often occurs when antennas are not co-located. In this paper, we present an analytical study of the effects of non-identical Rayleigh fading on the error performance of differential unitary space-time modulation (DUSTM). The fading processes for different transmit-receive antenna pairs are assumed to be independent and time-variant. We find that the maximum-likelihood (ML) differential detector of DUSTM over such channels is involved except for differential cyclic group codes. The conventional detector is proved to be asymptotically optimal in the limit of high signal-to-noise ratio (SNR) over static fading channels. Applying the distribution of quadratic forms of Gaussian vectors, we then derive closed-form expressions for the exact error probabilities of two specific unitary classes, namely, cyclic group codes and orthogonal codes. Simple and useful asymptotic bounds on error probabilities are also obtained. Our analysis leads to the following general findings: (1) equal power allocation is asymptotically optimal, and (2) non-identical channel gain distribution degrades the error performance. Finally, we also introduce a water-filling based power allocation to exploit the transmit non-identical fading statistics.**

***Index Terms*—Differential detector, error probability analysis, independent and non-identical channels, Rayleigh fading, space-time modulation.**

I. Introduction

THE use of multiple antenna elements promises considerable diversity and multiplexing gains in wireless communication systems. This motivated enormous development of multiple-input multiple-output (MIMO) techniques in the context of space-time (ST) coding and modulation in the last decade. Existing ST techniques can be broadly divided into coherent schemes and non-coherent schemes, based on whether or not instantaneous channel knowledge is needed by the receiver. As channel estimation is waived, non-coherent schemes can not only reduce receiver complexity but also lower transmission overhead required for sending pilot symbols. Among the non-coherent schemes, differential unitary space-time modulation (DUSTM) [1], [2] is known for its good error performance and high spectral efficiency. DUSTM is often viewed as a multiple-antenna counterpart of differential phase-shift-keying (DPSK) modulation, where the signal constellation is a set of unitary matrices spread over both time and space. A number of unitary ST signal sets have been designed, including orthogonal codes [3]–[5], cyclic group codes [6], and Cayley differential codes [7].

It is commonly assumed in the design and performance analysis of space-time coding that the channels on different transmit-receive antenna pairs are statistically identical. The assumption typically holds when antennas in the system are co-located and hence the channel path loss, as well as potential shadowing, experienced by each signaling branch is the same to each other. There are many occasions, however, that the antennas are not necessarily co-located. For instance, in distributed antenna systems [8], [9], the antennas are geographically distributed at different radio ports and are connected together through high-speed cables. It is natural to expect different path loss as well as fade rates on different links. Similarly, in aeronautical telemetry communications [10], multiple antennas can be placed at different parts of the air vehicle and hence they experience different attenuation during maneuvers. Cooperative communications among mobile nodes in a network is another important scenario. After knowing each other's data to be sent, the cooperating nodes can form a virtual multiple-antenna system and employ space-time coding in a distributed manner [11], [12]. Clearly, the different signaling branches in the cooperation phase can have unequal fading statistics. In all the aforementioned MIMO (or virtual MIMO) communication settings, the resulting channels can be modeled as independent and non-identically distributed (i.n.i.d) fading.

The goal of this paper is to study the effects of non-identical fading distribution on the performance of existing ST codes, in particular differential unitary space-time modulation. There are two issues to be addressed. First, whereas uniform power allocation in the spatial domain for both coherent and non-coherent ST codes is capacity-achieving in traditional independent and identically distributed (i.i.d) fading, it may not be so in i.n.i.d fading. Therefore, it is of interest to investigate the optimal power allocation among the distributed antennas (or cooperating nodes). Second, the conventional differential detector for DUSTM over i.i.d channels may no longer be optimal in the maximum-likelihood (ML) sense. Hence, optimal differential detector is to be discussed.

Attempts have been made recently to study the effects of non-identical channels in MIMO systems from different aspects. The outage probability of mutual information and power control over distributed multiple-input single-output (MISO) channels with independent Rayleigh fading are studied in [13]. The bit error probabilities (BEP) of coherent orthogonal space-time block codes (OSTBC) over i.n.i.d Rayleigh/Riciean and

Paper approved by G. M. Vitetta, the Editor for Equalization and Fading Channels of the IEEE Communications Society. Manuscript received October 15, 2007; revised February 8, 2008.

This work was presented in part at the IEEE International Conference on Communications, Beijing, China, May 2008. This work is supported in part by the Doctoral Fund of the Ministry of Education of China (No. 20082481002).

The author is with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, P. R. China (e-mail: mxtao@sjtu.edu.cn).

Digital Object Identifier 10.1109/TCOMM.2009.05.070534

Nakagami fading channels are analyzed in [14] and [15], respectively. In [16], the authors derived the BEP of differential OSTBC, i.e., the orthogonal-design based DUSTM [3], [5], over independent and semi-identically distributed (i.s.i.d) Rayleigh channels, where the non-identical fading occurs at the receiver side only. The study in [16] shows that in i.s.i.d channels the ML differential detector (DD) for differential OSTBC is still on a per symbol basis but should weight the output from each receive antenna according to its fading statistics. Moreover, the ML detector significantly outperforms the conventional one at high signal-to-noise ratio (SNR) region when the channel fluctuates rapidly over time.

In this paper, we extend the previous work in [16] to a general framework of DUSTM over i.n.i.d time-varying Rayleigh fading channels. We first show that for a general unitary space-time constellation the ML differential detector needs to perform joint optimization of the current data matrix and the previously transmitted signal matrix. However, for cyclic group codes, it is independent of the previous signals and differs from the conventional DD only by appropriate weights. The conventional DD is shown to be asymptotically optimal in the limit of high SNR over static fading channels. We then apply the well-established distribution of quadratic forms of Gaussian variables to derive the error performance for two specific unitary classes: orthogonal codes and cyclic group codes. For cyclic group codes, closed-form expressions for the exact pairwise error probabilities (PEP) with both ML and conventional DD at arbitrary channel fluctuation rates are derived. For orthogonal codes, closed-form expressions for the exact BEP with conventional DD in static fading are derived. Furthermore, simple asymptotic bounds on error probabilities for both codes are obtained. These bounds lead to several useful findings applied to any DUSTM design. Lastly, we propose a water-filling based power control to exploit the transmit non-identical fading statistics. This is carried out by minimizing the Chernoff bound of approximate error probabilities under a total power constraint.

The rest of the paper is organized as follows. In Section II we present the system model of DUSTM over i.n.i.d time-varying flat Rayleigh fading channels. The optimal and suboptimal detectors are presented in Section III. The analysis of error probabilities is presented in Section IV, followed by the derivation of transmit power control in Section V. Some numerical examples are illustrated in Section VI. Finally, Section VII offers some concluding remarks.

Notations: $\mathcal{E}[\cdot]$ denotes expectation over the random variables within the brackets. $\mathrm{Tr}(\mathbf{A})$ and $\mathrm{ReTr}(\mathbf{A})$ stand for the trace and the real part of the trace of matrix $\mathbf{A}$, respectively. $\|\cdot\|^2$ represents the squared Frobenius norm. $\mathbf{I}_M$ is the $M\times M$ identity matrix. $\mathrm{diag}(a_1,\ldots,a_M)$ is the diagonal matrix with element a_m on the m-th diagonal. Superscripts $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ denote transpose, conjugate, and conjugate transpose, respectively. Notation $\odot$ and $\otimes$, respectively, represent the Hadamard product and Kronecker product. $\mathrm{Res}[f(x),p]$ denotes the residue of function $f(x)$ at pole $x=p$.

II. System Model

Consider a communication system with M transmit and N receive antennas over a flat Rayleigh fading channel. The antenna elements at both the transmitter side and receiver side are not necessarily co-located. At each time block k, a set of $\log_2 L$ information bits are mapped onto a data matrix $\mathbf{D}[k]\in\mathcal{V}$, where $\mathcal{V}=\{\mathbf{D}_i, 0\le i<L\}$ denotes a unitary space-time signal constellation with cardinality L. Each element of $\mathcal{V}$ is an $M\times M$ unitary matrix, satisfying $\mathbf{D}_i\mathbf{D}_i^H=\mathbf{I}_M$, for $0\le i<L$. For the special case of differential cyclic group codes [2], [6], the constellation set $\mathcal{V}$ forms a group under matrix multiplication and each element of it is a diagonal matrix. In the case of differential OSTBC, each element $\mathbf{D}_i$ is a linear mapping of a set of P M-ary PSK modulated information symbols, denote as $\{s_p=e^{j\theta_p}\}_{p=1}^P$, and is given by $\mathbf{D}[k]=\frac{1}{\sqrt{P}}\sum_{p=1}^P(\mathbf{\Phi}_p\cos\theta_p+j\mathbf{\Psi}_p\sin\theta_p)$. Here the set of encoding matrices $\{\mathbf{\Phi}_p,\mathbf{\Psi}_p\}_{p=1}^P$ are chosen subject to certain orthogonality constraints [17].

Let $\mathbf{S}[k-1]$ denote the $M\times M$ dimensional code matrix at the $(k-1)$-th time block. The data matrix $\mathbf{D}[k]$ is then differentially encoded as

$$\mathbf{S}[k]=\mathbf{D}[k]\mathbf{S}[k-1],$$

where the initial code matrix $\mathbf{S}[0]$ is an arbitrary unitary matrix. The actual signal matrix to be transmitted at time block k over M antennas is given by

$$\mathbf{X}[k]=\sqrt{E_s}\mathbf{S}[k]\mathbf{\Sigma}^{1/2}, \quad (1)$$

where E_s is the total transmit power, and $\mathbf{\Sigma}^{1/2}=\mathrm{diag}\{\sqrt{\varepsilon_1},\ldots,\sqrt{\varepsilon_M}\}$ is the diagonal power allocation matrix. The power allocation coefficients ε_m's are subject to the constraint $\sum_{m=1}^M\varepsilon_m=M$ and to be optimized.

Since the transmission is on a per block basis, we assume the channel is block-wise time-varying with each block containing M symbol intervals. Let $\mathbf{H}[k]$ denote the $M\times N$ channel matrix of the k-th transmission block, where the (m,n)-th entry $h_{mn}[k]$ represents the fading coefficient from the m-th transmit antenna to the n-th receive antenna. Each $\{h_{mn}[k]\}_k$ is modeled as a complex Gaussian wide-sense stationary random process with zero mean and autocorrelation function $2R_{mn}[l]=\mathcal{E}[h_{mn}[k]h_{mn}^*[k-l]]$, and is independent for different m and n. The channel variance and block correlation coefficient are denoted as $\sigma_{mn}^2=2R_{mn}[0]$ and $\rho_{mn}=R_{mn}[1]/R_{mn}[0]$, respectively. The parameters $\{\sigma_{mn}^2\}$ and $\{\rho_{mn}\}$ represent the unequal average channel gains and unequal channel fluctuation rates, respectively.

Let $\mathbf{Y}[k]$ denote the $M\times N$ received signal matrix from N antennas at the k-th transmission block. It is modeled as

$$\mathbf{Y}[k] = \mathbf{X}[k]\mathbf{H}[k]+\mathbf{W}[k], \quad (2)$$

where $\mathbf{W}[k]$ is the complex-valued additive white Gaussian noise matrix whose entries are i.i.d with zero mean and variance N_0.

III. Differential Detection

Detection techniques of DUSTM over time-varying fading channels have evolved from traditional one-shot differential detection based on two consecutive blocks of received signals

[2] to more advanced sequence detection which jointly processes multiple blocks, e.g., [18]–[22]. While these multiple-symbol based sequence detectors are able to attain coherent-like performance, they are difficult to analyze. To quantitatively study the effects of non-identical fading statistics, we focus on the one-shot differential detection for the convenience of analytical tractability. In this section, we first discuss a general structure of the one-shot optimal differential detector of DUSTM in the ML sense over the considered channel model. Simplified detectors under certain constraints are then discussed. For notation brevity, we define $\gamma_0 = E_s/N_0$ as the total transmit SNR, and define $\gamma_{mn} = \varepsilon_m \sigma_{mn}^2 \gamma_0$ as the SNR on the branch between transmit antenna m and receive antenna n, for $1 \le m \le M$ and $1 \le n \le N$. In our high SNR assumption, all γ_{mn}'s approach infinity as $\gamma_0 \to \infty$, but the ratios between one another are kept constant and finite. We also omit the time index k in the data matrix $\mathbf{D}[k]$ and rewrite $\mathbf{S}[k-1]$ as $\mathbf{S}_{-1}$ hereafter as only one data matrix is processed at one time.

A. ML Detection of a General Constellation

Define $\mathbf{Y} = \left[\mathbf{Y}^T[k-1],\ \mathbf{Y}^T[k]\right]^T$. Given $\mathbf{D}$ and $\mathbf{S}_{-1}$, it can be shown easily that the column vectors of the sufficient statistics $\mathbf{Y}$, denoted as $\mathbf{y}_n$ for $n = 1, \ldots, N$, are mutually independent Gaussian vectors with zero mean and covariance (see (3) at top of next page). Here, $\mathbf{K}_{ni}$'s, for $i = 0, 1$ and $n = 1, \ldots, N$, are $M \times M$ diagonal matrices defined as $\mathbf{K}_{n0} = \mathrm{diag}\{\gamma_{1n}, \ldots, \gamma_{Mn}\}$ and $\mathbf{K}_{n1} = \mathrm{diag}\{\rho_{1n}\gamma_{1n}, \ldots, \rho_{Mn}\gamma_{Mn}\}$. Applying the formula for the determinant of a partitioned matrix [23], we can show that the determinant of the covariance matrix $\mathbf{\Lambda}_n$ is independent of the data matrix $\mathbf{D}$ and the previous code matrix $\mathbf{S}_{-1}$. The inverse of $\mathbf{\Lambda}_n$, however, in general depends not only on $\mathbf{D}$, but also on $\mathbf{S}_{-1}$. Using the Sherman-Morrison-Woodbury formula for the inverse of the matrix of the form $\mathbf{A} + \mathbf{BCD}$ [23] and utilizing the diagonal structure of $\mathbf{K}_{ni}$, we obtain the inverse of $\mathbf{\Lambda}_n$ as

$$(\mathbf{\Lambda}_n)^{-1} = \tfrac{1}{N_0}\mathbf{I}_{2M} - \tfrac{1}{N_0}\tilde{\mathbf{S}} \underbrace{\begin{bmatrix} \mathbf{C}_{n0} & \mathbf{C}_{n1} \\ \mathbf{C}_{n1} & \mathbf{C}_{n0} \end{bmatrix}}_{\tilde{\mathbf{C}}_n} \tilde{\mathbf{S}}^H, \tag{4}$$

where matrices $\mathbf{C}_{ni}$, for $i = 0, 1$ and $n = 1, \ldots, N$, are also diagonal, whose m-th diagonals, for $m = 1, \ldots, M$, are given by

$$\left[\mathbf{C}_{n0}\right]_m = \frac{\gamma_{mn}[1 + \gamma_{mn}(1 - \rho_{mn}^2)]}{(1+\gamma_{mn})^2 - (\rho_{mn}\gamma_{mn})^2}, \tag{5}$$

$$\left[\mathbf{C}_{n1}\right]_m = \frac{\gamma_{mn}\rho_{mn}}{(1+\gamma_{mn})^2 - (\rho_{mn}\gamma_{mn})^2}. \tag{6}$$

The ML differential detector of $\mathbf{D}$ is to choose the candidate $\hat{\mathbf{D}} \in \mathcal{V}$ that maximizes the joint likelihood function of the received signal matrix $\mathbf{Y}$ over all possible $\mathbf{S}_{-1}$. Note that the dependence of ML detection on $\mathbf{S}_{-1}$ also arises in transmit-correlated channels as mentioned in [24] and [25]. But no explicit ML decision metric is given therein due to the lack of closed-form expression for $(\mathbf{\Lambda}_n)^{-1}$. Applying (4), we have the quadratic-form based ML differential detector of DUSTM over i.n.i.d channels:

$$\hat{\mathbf{D}}_{\mathrm{ML}} = \arg\max_{\mathbf{D}\in\mathcal{V}} \max_{\mathbf{S}_{-1}} \sum_{n=1}^{N} \mathbf{y}_n^H \tilde{\mathbf{S}} \tilde{\mathbf{C}}_n \tilde{\mathbf{S}}^H \mathbf{y}_n. \tag{7}$$

The complexity of the ML detector is proportional to the product of the constellation size L and the total number of all possible previous code matrices. For those constellations with group structure, $\mathbf{S}_{-1}$ also belongs to the signal set $\mathcal{V}$ and hence the complexity is in the order of L^2. For constellations without group structure, such as OSTBC, the total number of possible $\mathbf{S}_{-1}$ can grow rapidly as L increases.

For the semi-identical channel [16] and with equal power allocation, we have $\gamma_{mn} = \gamma_n$ and $\rho_{mn} = \rho_n$, for all m. It follows that both $\mathbf{C}_{n0}$ and $\mathbf{C}_{n1}$ become scaled identity matrices. Therefore, the ML detector no longer depends on $\mathbf{S}_{-1}$ and is simplified to

$$\hat{\mathbf{D}}_{\mathrm{ML,semi}} = \arg\max_{\mathbf{D}\in\mathcal{V}} \sum_{n=1}^{N} w_n \mathbf{y}_n^H \begin{bmatrix} \mathbf{I} \\ \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{D}^H \end{bmatrix} \mathbf{y}_n,$$

where $w_n = \gamma_n\rho_n/[(1+\gamma_n)^2 - (\rho_n\gamma_n)^2]$. This detector differs from the conventional detector for i.i.d channel [2] at the weights w_n's only. The weights exploit the knowledge of fading statistics and total transmit SNR, and are optimal for any unitary constellation including OSTBC [16].

B. ML Detection of Cyclic Group Codes

It is shown in [6] that every full-diverse unitary constellation having a group structure can be made equivalent to a cyclic (also called diagonal in [2]) group for odd M, and either a cyclic group or dicyclic group for even M. Therefore, the cyclic group constellations are of particular interest to us. Because of the diagonal structure inherent in cyclic groups, the code matrix $\mathbf{S}_{-1}$ is always diagonal as long as the initial matrix $\mathbf{S}[0]$ is diagonal, say $\mathbf{I}_M$. Since multiplication commutes for diagonal matrices, we have $\mathbf{S}_{-1}\mathbf{C}_{ni}\mathbf{S}_{-1}^H = \mathbf{C}_{ni}$, for all i and n. Therefore, the ML detector for cyclic group codes reduces to:

$$\hat{\mathbf{D}}_{\mathrm{ML,c}} = \arg\max_{\mathbf{D}\in\mathcal{V}} \sum_{n=1}^{N} \mathbf{y}_n^H \begin{bmatrix} \mathbf{I} \\ \mathbf{C}_{n1}\mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{C}_{n1}\mathbf{D}^H \end{bmatrix} \mathbf{y}_n \tag{8}$$

which can be further expressed as:

$$\hat{\mathbf{D}}_{\mathrm{ML,c}} = \arg\max_{\mathbf{D}\in\mathcal{V}} \mathrm{ReTr}\left\{(\mathbf{Y}[k-1] \odot \mathbf{W})^H \mathbf{D}^H \mathbf{Y}[k]\right\}, \tag{9}$$

where $\mathbf{W}$ is an $M \times N$ matrix with the (m,n)-th entry formed by the m-th diagonal of $\mathbf{C}_{n1}$ and repeated as:

$$w_{mn} = \frac{\gamma_{mn}\rho_{mn}}{(1+\gamma_{mn})^2 - (\rho_{mn}\gamma_{mn})^2}. \tag{10}$$

It is clear from (9) that the ML DD for cyclic group codes resembles the conventional DD but applies a weight w_{mn} to the (m,n)-th element of $\mathbf{Y}[k-1]$. The resulting Hadamard product $\mathbf{Y}[k-1] \odot \mathbf{W}$ behaves as the equivalent channel matrix as in coherent receivers.

$$\begin{aligned}\mathbf{\Lambda}_n &= \mathcal{E}[\mathbf{y}_n\mathbf{y}_n^H] \\ &= N_0 \underbrace{\begin{bmatrix} \mathbf{S}_{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}\mathbf{S}_{-1} \end{bmatrix}}_{\tilde{\mathbf{S}}} \cdot \begin{bmatrix} \mathbf{K}_{n0} & \mathbf{K}_{n1} \\ \mathbf{K}_{n1} & \mathbf{K}_{n0} \end{bmatrix} \cdot \underbrace{\begin{bmatrix} \mathbf{S}_{-1}^H & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{-1}^H\mathbf{D}^H \end{bmatrix}}_{\tilde{\mathbf{S}}^H} + N_0\mathbf{I}_{2M}. \end{aligned} \quad (3)$$

C. Asymptotically Optimal Detection of a General Constellation in Static Channels

In static fading channels the channel coefficients are assumed to remain unchanged over the duration of two transmission blocks. Therefore it has $\rho_{mn} = 1$ for all m and n. In the limit $\gamma_{mn} \to \infty$ for all m and n, the matrices $\mathbf{C}_{n0}$ and $\mathbf{C}_{n1}$ defined in (5) (6) all approach $(1/2)\mathbf{I}_M$. Applying these into (7), we obtain the asymptotically optimal detector:

$$\hat{\mathbf{D}}_{\text{AO}} = \arg\max_{\mathbf{D}\in\mathcal{V}} \sum_{n=1}^{N} \mathbf{y}_n^H \begin{bmatrix} \mathbf{I} \\ \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{D}^H \end{bmatrix} \mathbf{y}_n. \quad (11)$$

This is identical to the conventional DD [2]. Hence we conclude that the conventional DD is suboptimal in i.n.i.d time-varying channels but asymptotically optimal (high SNR) in i.n.i.d static channels.

IV. Error Probability Analysis

In this section, we derive the error performance of DUSTM with two specific constellation designs: cyclic group codes and orthogonal codes. Through the analysis, we obtain several general findings that are applicable to an arbitrary DUSTM design.

A. Pairwise Error Probability for Cyclic Group Codes

Since the exact bit or block error probability of a cyclic group code $\mathcal{V}$ with $L > 2$ elements is usually not computable, we resort to the union bound by summing up pairwise error probabilities. In specific, the block error probability (BkEP) for equiprobable elements is bounded by

$$P_{e,\text{UB}} = \frac{1}{L}\sum_{i=0}^{L-1}\sum_{\substack{j=0\\ j\neq i}}^{L-1} P_{e,ij}, \quad (12)$$

where $P_{e,ij}$ denotes the PEP of deciding in favor of data matrix $\mathbf{D}_j$ given that $\mathbf{D}_i$ is sent. In the following we derive the exact expressions for $P_{e,ij}$ and the asymptotics.

Based on the quadratic form of the ML and suboptimal detectors given in (8) and (11) respectively, the PEP can be expressed as

$$P_{e,ij} = P\left(z_{ij} < 0|\mathbf{D}_i\right), \quad (13)$$

where the pairwise decision variable z_{ij} is defined as

$$z_{ij} = \sum_{n=1}^{N} \mathbf{y}_n^H \mathbf{\Omega}_{ij,n}\mathbf{y}_n, \quad (14)$$

with $\mathbf{\Omega}_{ij,n}$ given by

$$\mathbf{\Omega}_{ij,n} = \begin{bmatrix} \mathbf{0} & \mathbf{C}_{n1}(\mathbf{D}_i - \mathbf{D}_j)^H \\ \mathbf{C}_{n1}(\mathbf{D}_i - \mathbf{D}_j) & \mathbf{0} \end{bmatrix} \quad (15)$$

for ML detection and

$$\mathbf{\Omega}_{ij,n} = \mathbf{\Omega}_{ij} = \begin{bmatrix} \mathbf{0} & (\mathbf{D}_i - \mathbf{D}_j)^H \\ (\mathbf{D}_i - \mathbf{D}_j) & \mathbf{0} \end{bmatrix}, \ \forall n \quad (16)$$

for conventional detection. Since each vector $\mathbf{y}_n$ is independent and zero-mean complex Gaussian distributed, the pairwise decision variable z_{ij} is a quadratic form of Gaussian vectors. Therefore, the evaluation of PEP can be carried out by using the well-established techniques in, e.g., [24, Appendix A]. We summarize the results in the following proposition.

Proposition 1: The exact pairwise error probability $P_{e,ij}$ of differential cyclic group codes over i.n.i.d time-varying Rayleigh fading channels is (17) at the bottom of the page, where

$$a_{mn} = \frac{(\rho_{mn}\gamma_{mn})^2}{(1+\gamma_{mn})^2 - (\rho_{mn}\gamma_{mn})^2} d_{ij,m} \quad (18)$$

with $d_{ij,m}$ being the m-th diagonal entry of the difference matrix $(\mathbf{D}_i - \mathbf{D}_j)(\mathbf{D}_i - \mathbf{D}_j)^H$, and

$$b_{mn} = \begin{cases} 1, & \text{ML detector} \\ w_{mn} \text{ in (10)}, & \text{conv. detector} \end{cases}. \quad (19)$$

In the case of ML differential detection, an alternative expression is given by

$$P_{e,ij} = \frac{1}{\pi}\int_0^{\pi/2} \prod_{n=1}^{N}\prod_{m=1}^{M} \left(1 + \frac{a_{mn}}{4\sin^2\theta}\right)^{-1} d\theta. \quad (20)$$

Proof: See Appendix A. ■

Although the expression for $P_{e,ij}$ is exact and in closed form, it does not offer much insight on the effects of channel parameters. Therefore, useful asymptotic bounds are desirable. We first consider the asymptotic $P_{e,ij}$ over static channels with $\rho_{mn} = 1$, for all m, n. Then we conduct the error floor analysis in time-varying channels. The results are summarized in the following two corollaries.

$$P_{e,ij} = -\sum_{\substack{1\le k\le M\\ 1\le l\le N}} \text{Res}\left\{ \frac{1}{s\prod_{m=1}^{M}\prod_{n=1}^{N} a_{mn}\left[\frac{1}{4} + \frac{1}{a_{mn}} - \left(\frac{s}{b_{mn}} - \frac{1}{2}\right)^2\right]},\ s = b_{kl}\left(\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{a_{kl}}}\right)\right\} \quad (17)$$

Corollary 1: The asymptotic PEPs of differential cyclic group codes at high SNR with ML and conventional differential detection over i.n.i.d static Rayleigh fading channels are the same and given by

$$\lim_{\gamma_0\to\infty} P_{e,ij} = \binom{2MN-1}{MN}\left(\prod_{m=1}^{M} d_{ij,m}\right)^{-N}\left(\frac{\gamma_{\rm gm}}{2}\right)^{-MN}, \quad (21)$$

where $\gamma_{\rm gm}$ is the geometric mean of $\{\gamma_{mn}\}$, given by $\gamma_{\rm gm} = \left(\prod_{n=1}^{N}\prod_{m=1}^{M}\gamma_{mn}\right)^{1/MN}$, and $\binom{2MN-1}{MN}$ denotes the binomial coefficient.

Proof: Assuming static fading channels with $\rho_{mn} = 1$ and taking the limit $\gamma_0 \to \infty$, one finds from (18) and (10) that $a_{mn} \to \gamma_{mn} d_{ij,m}/2$ and $w_{mn} \to 1/2$. As a result, the MN poles, where the residues are evaluated in Proposition 1, all approach the constant 1 for ML DD and the constant $1/2$ for conventional DD. Using the residue equation of a function $f(x)$ at a pole p of multiplicity v

$$\mathrm{Res}\left[f(x), p\right] = \frac{1}{(v-1)!}\lim_{x\to p}\frac{d^{v-1}}{dx^{v-1}}\left[(x-p)^v f(x)\right] \quad (22)$$

and applying the formula

$$\frac{d^m}{dx^m}(x+p)^{-k} = \frac{(-1)^m(m+k-1)!}{(k-1)!}(x+p)^{-(m+k)} \quad (23)$$

in (17), we arrive at the asymptotic results in (21) for both ML and conventional DD. ■

Corollary 1 leads to several insights. First, the result that the asymptotic PEPs of ML and conventional DD are the same is consistent with the finding in Section III-C that the conventional DD is in fact asymptotically optimal for a general constellation without assuming a specific signal structure. Second, the traditional diversity product design criterion[1] for i.i.d channels [1], [2] still applies to i.n.i.d channels. That is, the minimum of $\prod_{m=1}^{M} d_{ij,m}$ over all distinct pairs $(\mathbf{D}_i, \mathbf{D}_j)$ should be maximized. Using the arithmetic-geometric-mean inequality (i.e. $\gamma_{\rm gm} \le \gamma_{\rm am} = 1/MN\sum_{n=1}^{N}\sum_{m=1}^{M}\gamma_{mn}$, where the equality holds if and only if γ_{mn}'s are all the same), we can see that the non-identical channel distribution will degrade the error performance compared with the identical case if the total received SNR is kept the same. Furthermore, after rewriting $\gamma_{\rm gm}$ as $\gamma_0(\prod_{m=1}^{M}\varepsilon_m)^{1/M}(\prod_{n=1}^{N}\prod_{m=1}^{M}\sigma_{mn}^2)^{1/MN}$, where ε_m is the power allocation coefficient defined in (1) subject to the constraint $\sum_{m=1}^{M}\varepsilon_m = M$, it follows that $\gamma_{\rm gm}$ is maximized when $\varepsilon_m = 1$ for all m. Therefore, equal power allocation is asymptotically optimal in static channels.

[1]Assume the full-rank criterion is already satisfied.

Corollary 2: The pairwise error floor of differential cyclic group codes with ML differential detection over i.n.i.d time-varying Rayleigh fading channels is independent of the unequal average channel gains $\{\sigma_{mn}^2\}$ but depends on the fading correlation coefficients $\{\rho_{mn}\}$ only. It is given by (24) at the bottom of the page. In the case where $\rho_{m,n} = \rho$ and $\rho \approx 1$ but $\rho \neq 1$, the pairwise error floor is simplified as

$$\lim_{\substack{\gamma_0\to\infty \\ \rho_{mn}=\rho\approx 1}} P_{e,ij} = \binom{2MN-1}{MN}\left(\prod_{m=1}^{M} d_{ij,m}\right)^{-N}\left(\frac{\rho^2}{1-\rho^2}\right)^{-MN} \quad (25)$$

Proof: The proof of the first equation in the corollary is straightforward by observing that $a_{mn} \to \rho_{mn}^2 d_{ij,m}/(1-\rho_{mn}^2)$ when $\gamma_0 \to \infty$ in (17) (with $b_{mn} = 1$). If $\rho_{m,n} = \rho$ and $\rho \approx 1$ but $\rho \neq 1$, the MN positive poles all approach $1/2$. Using the formulas (22) and (23) again, we prove the second equation in the corollary. ■

Corollary 2 concludes that the irreducible error floors achieved over i.n.i.d channels (with ML detection) and traditional i.i.d channels are the same, as long as their fade rates are the same ($\rho_{m,n} = \rho$, $\forall m, n$). Moreover, the error floor decreases exponentially with MN when ρ is very close to but not equal to one. This condition on ρ typically holds if the normalized Doppler frequency of the channel is much less than one.

B. Bit Error Probability for Orthogonal Codes

In this subsection we derive the error performance of differential OSTBC. Only the conventional DD and static channels are considered. The analysis for time-varying fading or ML detection is so far not tractable. Since the data matrix $\mathbf{D}$ is a linear combination of P information symbols as mentioned in Section II, the differential detector (11) reduces to P independent symbol-by-symbol detectors. The details are given in [5] or [16, eq.(12)]. Hence, instead of PEP, BEP is derived.

As shown in [16, eq.(15)-(17)], the BEP conditioned on symbol s_p is the same for all p, and can be expressed as

$$P_b(\alpha) = P(z_p(\alpha) < 0 | s_p = 1), \quad (26)$$

where the decision phasor $z_p(\alpha)$ is defined as

$$z_p(\alpha) = \sum_{n=1}^{N}\mathbf{y}_n^H\mathbf{\Omega}_p\mathbf{y}_n \quad (27)$$

$$\lim_{\substack{\gamma_0\to\infty \\ \rho_{mn}<1}} P_{e,ij} = -\sum_{\substack{1\le k\le M \\ 1\le l\le N}}\mathrm{Res}\left\{\frac{1}{s\prod_{m=1}^{M}\prod_{n=1}^{N}\frac{\rho_{mn}^2 d_{ij,m}}{1-\rho_{mn}^2}\left[\frac{1}{4}+\frac{1-\rho_{mn}^2}{\rho_{mn}^2 d_{ij,m}}-\left(s-\frac{1}{2}\right)^2\right]},\ s=\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{1-\rho_{kl}^2}{\rho_{kl}^2 d_{ij,k}}}\right\} \quad (24)$$

$$P_b(\alpha) = -\sum_{\substack{1\le k\le M \\ 1\le l\le N}}\mathrm{Res}\left\{\frac{1}{s\prod_{n=1}^{N}\prod_{m=1}^{M}\left[1+2c\gamma_{mn}s-(1+2\gamma_{mn})s^2\right]},\ s=\frac{c\gamma_{kl}}{1+2\gamma_{kl}}\left(1+\sqrt{1+\frac{1+2\gamma_{kl}}{c^2\gamma_{kl}^2}}\right)\right\} \quad (28)$$

with the Hermitian matrix $\mathbf{\Omega}_p$ given by

$$\mathbf{\Omega}_p = \begin{bmatrix} \mathbf{0} & \cos\alpha\mathbf{\Phi}_p^H + j\sin\alpha\mathbf{\Psi}_p^H \\ \cos\alpha\mathbf{\Phi}_p - j\sin\alpha\mathbf{\Psi}_p & \mathbf{0} \end{bmatrix},$$

and α is some angle that depends on the symbol modulation scheme. For BPSK, the exact BEP is obtained by letting $\alpha = 0$, and for QPSK with Gray mapping we have $\alpha = -\pi/4$.

Proposition 2: The exact bit error probability $P_b(\alpha)$ of differential OSTBC over i.n.i.d static Rayleigh fading channels with conventional differential detection is (28) (see the bottom of the previous page), where $c = \cos\alpha/\sqrt{P}$.

Proof: See Appendix B ∎

At high SNR, all the MN positive poles, where the residues are evaluated, approach the constant c. Therefore, we obtain the asymptotics of $P_b(\alpha)$ as follows by applying the definition of residue as in the proof of Corollary 1.

Corollary 3: The asymptotic BEP of differential OSTBC over i.n.i.d static Rayleigh fading channels with conventional differential detection is

$$\lim_{\gamma_0\to\infty} P_b(\alpha) = \binom{2MN-1}{MN}\left(\frac{2\cos^2\alpha}{P}\gamma_{\text{gm}}\right)^{-MN}, \quad (29)$$

where γ_{gm} is the geometric mean of $\{\gamma_{mn}\}$.

The implications in Corollary 3 are the same as those in Corollary 1. Therefore, we readily extend the following remarks to the *general DUSTM*.

Remarks: (1) Non-identical fading degrades the error performance compared with the identical case given the same total received SNR; (2) Equal power allocation is asymptotically optimal in i.n.i.d static fading channels.

V. Transmit Power Control

Given the unequal channel gain distribution among different transmit antennas, it is intuitive to use power control to improve the error performance, especially when the total transmit power is small. To simplify investigation, we consider static channels only in this section. Moreover, as shown in Section VI and [16], the conventional detector performs almost the same as the ML detector in i.n.i.d static fading channels. Hence, we assume conventional DD here.

Both the exact PEP in Proposition 1 for cyclic group codes and the exact BEP in Proposition 2 for orthogonal codes are difficult to minimize directly. We resort to minimizing a simple but useful approximate bound of them. In the following we present the derivation of transmit power control for the two codes separately for the ease of presentation, though the approaches are very similar.

A. Power Control for Cyclic Group Codes

The pairwise decision variable in (14) for the conventional DD can be rewritten as

$$z_{ij} = 2\text{ReTr}\{\mathbf{Y}[k]^H\mathbf{E}\mathbf{Y}[k-1]\} \quad (30)$$

where $\mathbf{E} = \mathbf{D}_i - \mathbf{D}_j$. Given that $\mathbf{D}_i$ is sent, substituting (2) into (30) yields

$$\begin{aligned} z_{ij} = {} & 2\text{ReTr}\{E_s\hat{\mathbf{H}}^H\mathbf{D}_i^H\mathbf{E}\hat{\mathbf{H}}\} \\ & + \underbrace{2\text{ReTr}\{\sqrt{E_s}\hat{\mathbf{H}}^H\mathbf{D}_i^H\mathbf{E}\mathbf{W}[k-1]\}}_{\eta_1} \\ & + \underbrace{2\text{ReTr}\{\sqrt{E_s}\mathbf{W}^H[k]\mathbf{E}\hat{\mathbf{H}}\}}_{\eta_2} \\ & + \underbrace{2\text{ReTr}\{\mathbf{W}^H[k]\mathbf{E}\mathbf{W}[k-1]\}}_{\eta_3}, \end{aligned} \quad (31)$$

where $\hat{\mathbf{H}} = \mathbf{S}_{-1}\mathbf{\Sigma}^{1/2}\mathbf{H}$ with $\mathbf{H}$ being the channel matrix in static fading. It can be easily shown that, conditioned on $\hat{\mathbf{H}}$, the noise terms η_1 and η_2 are independent and real-valued zero-mean Gaussian variables with variance $2N_0E_s\|\mathbf{E}\hat{\mathbf{H}}\|^2$ for each. By neglecting the second-order noise term η_3, which has diminishing effect at high SNR (i.e., $E_s/N_0 \gg 1$), and noting $\mathbf{D}_i^H\mathbf{E} + \mathbf{E}^H\mathbf{D}_i = \mathbf{E}^H\mathbf{E}$, the pairwise decision variable z_{ij} can be approximated as a Gaussian variable with mean $E_s\|\mathbf{E}\hat{\mathbf{H}}\|^2$ and variance $4N_0E_s\|\mathbf{E}\hat{\mathbf{H}}\|^2$. As a result, the conditional probability of $z_{ij} < 0$ can be expressed in the form of standard Q-function [26]. Further, applying the inequality $Q(x) \le \frac{1}{2}\exp(-x^2/2)$, we obtain the Chernoff bound of the approximate PEP as

$$P_e \lesssim \frac{1}{2}\mathcal{E}\left[\exp\left(-\frac{\gamma_0}{8}\|\mathbf{E}\hat{\mathbf{H}}\|^2\right)\right].$$

Obtaining the distribution of $\|\mathbf{E}\hat{\mathbf{H}}\|^2$ is difficult in general if $\mathbf{E}$ is the difference matrix of an arbitrary unitary constellation. Fortunately, by utilizing the diagonal structure of cyclic group codes, it is clear that $\|\mathbf{E}\hat{\mathbf{H}}\|^2$ can be expressed as a weighted sum of absolute squares of MN i.i.d complex Gaussian variables with weights given by $\varepsilon_m\sigma_{mn}^2 d_{ij,m}$. Hence, the above expectation can be evaluated as

$$P_{e,ij} \lesssim \frac{1}{2}\prod_{n=1}^{N}\prod_{m=1}^{M}\left(1 + \frac{\gamma_0\varepsilon_m\sigma_{mn}^2 d_{ij,m}}{8}\right)^{-1}. \quad (32)$$

We now find the optimal power allocation coefficients ε_m's to minimize the bound in (32) for a dominant error pair, which consequently provides a good result in minimizing the overall block error probability. The dominant error pair of a cyclic group code is the data matrix pair that has the smallest $\prod_{m=1}^{M} d_{ij,m}$ [2], denoted as $\zeta = \min_{0\le i<j<L}\prod_{m=1}^{M} d_{ij,m}$. However, there can be multiple pairs in the code that result in the same ζ, and they may differ dramatically in $d_{ij,m}$, for $m = 1,\dots,M$. Take the cyclic group code with $M = 4$ and $L = 16$ for example, given by $\mathcal{V}_{4,16} = \{\mathbf{D}_k = \text{diag}(e^{jk\pi/8}, e^{j3k\pi/8}, e^{j5k\pi/8}, e^{j7k\pi/8}), k = 0,\dots,15\}$ [27, Table I]. The data matrix pair $(\mathbf{D}_0, \mathbf{D}_1)$ is a dominant error pair and it has $d_1 = 0.1522$, $d_2 = 1.2346$, $d_3 = 2.7654$ and $d_4 = 3.8478$. On the other hand, the pair $(\mathbf{D}_0, \mathbf{D}_3)$ is also a dominant error pair but it has $d_1 = 1.2346$, $d_2 = 3.8478$, $d_3 = 0.1522$, and $d_4 = 2.7654$. The sets of power allocation coefficients minimizing the two pairwise error probabilities are obviously different. To overcome this problem, we take the mean of $d_{ij,m}$ over all dominant error pairs for each m, denote it as $\bar{d}_m$, and replace all $d_{ij,m}$ with it in the PEP bound (32).

Then, by using the monotonic property of logarithm function, the power allocation problem can be formulated as

$$\max_{\sum_m^M \varepsilon_m = M} \quad \sum_{n=1}^{N}\sum_{m=1}^{M} \log\left(1+\varepsilon_m \frac{\sigma_{mn}^2 \bar{d}_m \gamma_0}{8}\right). \quad (33)$$

In the case of $N = 1$, we obtain the water-filling based closed-form expression for the optimal power control as [28]

$$\varepsilon_m = \left(\mu - \frac{8}{\sigma_m^2 \bar{d}_m \gamma_0}\right)^+, \quad (34)$$

where $(x)^+ = \max\{0, x\}$, and the Lagrange multiplier μ can be determined by the constraint $\sum_{m=1}^{M} \varepsilon_m = M$.

If there are $N > 1$ number of receive antennas, closed-form expressions of optimal power coefficients are difficult to find. Here we propose a suboptimal approach. Applying the inequality [29, eq.(25)]

$$\prod_{i=1}^{N}(1+x_i) \geq (1+x_{\text{gm}})^N,$$

where $x_{\text{gm}} = (\prod_{i=1}^{N} x_i)^{1/N}$, we can reformulate (33) as

$$\max_{\sum_m^M \varepsilon_m = M} \quad \sum_{m=1}^{M} \log\left(1+\varepsilon_m \frac{\sigma_{m,\text{gm}}^2 \bar{d}_m \gamma_0}{8}\right),$$

where $\sigma_{m,\text{gm}}^2 = (\prod_{n=1}^{N} \sigma_{mn}^2)^{1/N}$. Hence, the solution in (34) still applies after replacing σ_m^2 with $\sigma_{m,\text{gm}}^2$, and is given by

$$\varepsilon_m = \left(\mu - \frac{8}{\sigma_{m,\text{gm}}^2 \bar{d}_m \gamma_0}\right)^+. \quad (35)$$

In summary, the proposed transmit power control aims to minimize the Chernoff bound of an approximate PEP of dominant error pairs in the constellation. It has a water-filling structure, and hence inherits the two distinguishing properties of water-filling principle. First, when the total transmit power is low, the transmit antennas with smaller geometric mean of average channel gains should be turned off. Second, when the total transmit power is high enough, the power tends to be equally distributed among all the antennas. The second property is consistent with the finding from Section IV that equal power allocation is asymptotically optimal.

B. Power Control for Orthogonal Codes

The transmit power allocation for DUSTM with orthogonal codes is similar to that for DUSTM with cyclic group codes. The decision phasor $z_p(\alpha)$ (27) can also be expressed as (31), except that $\mathbf{E}$ should be defined as $\mathbf{E} = \cos\alpha \mathbf{\Phi}_p^H + j\sin\alpha \mathbf{\Psi}_p^H$. Using the orthogonal code structure, we can easily show that the distribution of $z_p(\alpha)$ can be approximated as Gaussian with mean $\frac{2\cos(\alpha)E_s}{\sqrt{P}}\|\mathbf{\Sigma}^{1/2}\mathbf{H}\|^2$ and variance $4E_sN_0\|\mathbf{\Sigma}^{1/2}\mathbf{H}\|^2$. Thus, we obtain the Chernoff bound of the approximate BEP for differential OSTBC as

$$P_b(\alpha) \lesssim \frac{1}{2}\prod_{n=1}^{N}\prod_{m=1}^{M}\left(1+\frac{\cos^2(\alpha)}{2P}\varepsilon_m \sigma_{mn}^2 \gamma_0\right)^{-1}.$$

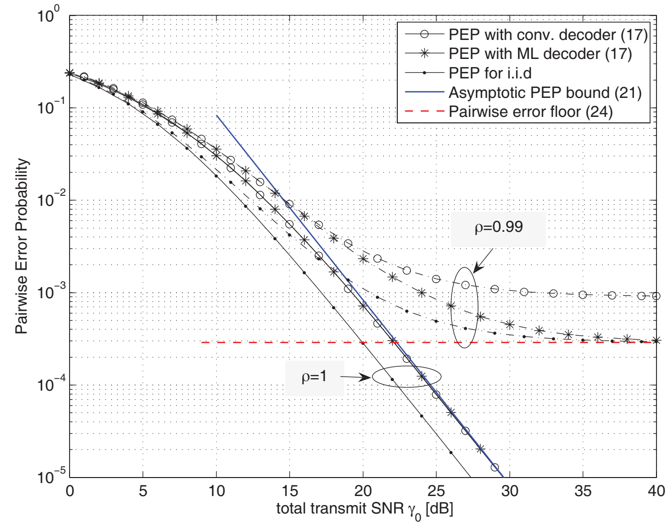


Fig. 1. PEP performance of differential cyclic group code $\mathcal{V}_{2,4}$ over the dominant error pair at $M = 2$ transmit and $N = 1$ receive antenna.

Correspondingly, a water-filling based sub-optimal power control that minimizes the bound is given by

$$\varepsilon_m = \left(\mu - \frac{2P}{\cos^2(\alpha)\sigma_{m,\text{gm}}^2 \gamma_0}\right)^+. \quad (36)$$

VI. Numerical Results

In this section we present some numerical examples to confirm our analytical findings in previous sections. We first verify the error probability analysis using a system with $M = 2$ transmit antennas and $N = 1$ or $N = 2$ receive antennas. Then we demonstrate the performance of the proposed transmit power allocation in a system with $M = 4$ transmit antennas and $N = 1$ receive antenna.

In our first set of examples, we assume equal fade rates on all transmit-receive antenna pairs and illustrate the effects of non-identical channel gain distribution. The unequal average channel gains are generated using the Kronecker model [16]. In specific, the $MN \times MN$ diagonal matrix $\mathbf{\Delta}$ with σ_{mn}^2 on the $[(n-1)M + m]$-th diagonal is decomposed as $\mathbf{\Delta} = \mathbf{\Delta}_T \otimes \mathbf{\Delta}_R$, where $\mathbf{\Delta}_T$ and $\mathbf{\Delta}_R$ are, respectively, the $M \times M$ and $N \times N$ diagonal matrices inducing non-identical fading parameters at the transmitter and receiver. The sum of the average channel gains is normalized so that $\text{Tr}\{\mathbf{\Delta}_T\} = M$ and $\text{Tr}\{\mathbf{\Delta}_R\} = N$. In the system with two transmit antennas, we specify $\mathbf{\Delta}_T = \text{diag}(\frac{1}{5}, \frac{9}{5})$. For one receive antenna, the average channel gains are given by $\sigma_1^2 = 1/5$ and $\sigma_2^2 = 9/5$. For two receive antennas, we let $\mathbf{\Delta}_R = \text{diag}(\frac{1}{5}, \frac{9}{5})$, and the set of average channel gains is given by $\{\sigma_{11}^2 = 1/25, \sigma_{12}^2 = \sigma_{21}^2 = 9/25, \sigma_{22}^2 = 81/25\}$.

In Figs. 1-3, we show the performance of differential cyclic group codes. The exemplary cyclic group code for $M = 2$ and $L = 4$ at rate 1-bit/s/Hz [27, Table I], denoted as $\mathcal{V}_{2,4}$, is chosen. Figs. 1 and 2 show the analytical PEP of the dominant error pair $(\mathbf{D}_0,\ \mathbf{D}_1)$ using one and two receive antennas, respectively. The exact PEP results over i.i.d channels are also plotted for reference, which are obtained using (20) by letting $\sigma_{mn}^2 = 1,\ \forall m, n$. Several useful observations can be made

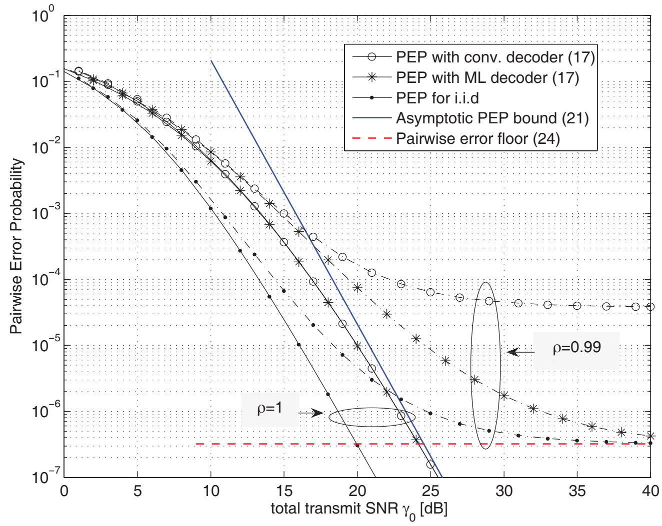


Fig. 2. PEP performance of differential cyclic group code $\mathcal{V}_{2,4}$ over the dominant error pair at $M = 2$ transmit and $N = 2$ receive antennas.

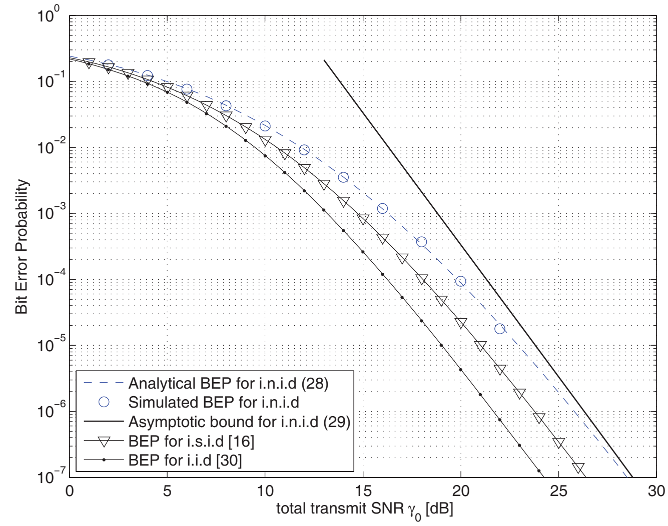


Fig. 4. BEP performance of differential OSTBC with QPSK modulation at $M = 2$ transmit and $N = 2$ receive antennas in static fading.

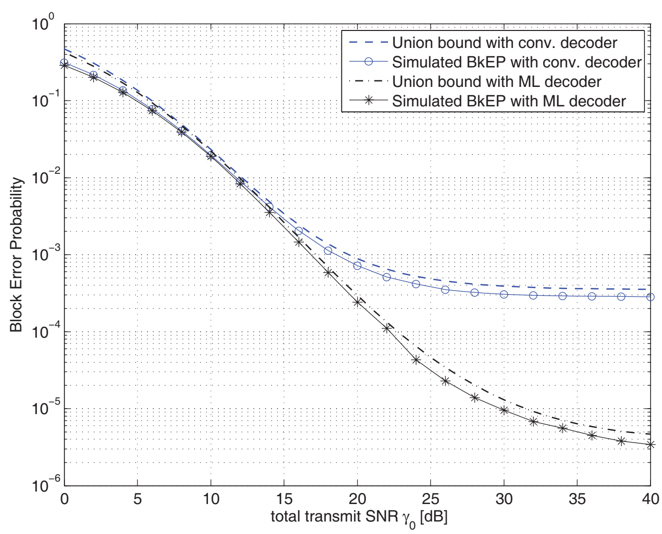


Fig. 3. BkEP performance of differential cyclic group code $\mathcal{V}_{2,4}$ at $M = 2$ transmit and $N = 2$ receive antennas in time-varying fading with $f_d T_s = 0.02$ ($\rho = 0.98427$).

from the two figures. First, the pairwise error performance achieved using the conventional DD is almost the same as that achieved by the ML detector at all SNR when $\rho = 1$ as well as at low SNR in time-varying fading with $\rho = 0.99$. Second, the ML detector considerably reduces the pairwise error floor in fast fading compared with the conventional detector. In particular, the pairwise error floor of the ML detector with two receive antennas is two order of magnitude lower than that of the conventional detector. Moreover, the error floors approach those in i.i.d channels and match very well with the flat lines predicted by the analytical result in (24). This observation confirms our analytical finding from Corollary 2. From the figures we also observe that the simple asymptotic PEP bound (21) is very tight when γ_0 is large enough. Finally, compared with i.i.d channels, the non-identical channel gain distribution degrades the PEP performance. This confirms the analytical finding from Corollary 1.

The overall BkEP performance of this cyclic group code obtained via simulation is shown in Fig. 3 and compared with the BkEP union bound obtained analytically using (12). The block-wise time-varying fading channel is generated using Jakes model with autocorrection function $2R_{mn}[k] = \sigma_{mn}^2 \mathcal{J}_0(2\pi f_d T_s M k)$, where $\mathcal{J}_0(\cdot)$ is the zeroth order Bessel function of the first kind and $f_d T_s$ is the normalized Doppler frequency. In our simulation we set $f_d T_s = 0.02$, which results in $\rho = 0.98427$. It is observed that the analytical BkEP union bound serves as a tight upper bound on the actual BkEP with both conventional and ML detectors. This further validates our theoretical analysis on the exact PEP in Proposition 1.

The BEP results of differential OSTBC over the i.n.i.d channel with two transmit and two receive antennas are depicted in Fig. 4. The orthogonal code for two transmit antennas with $P = 2$ and QPSK modulation at rate 2-bit/s/Hz is used. The analytical BEPs are from (28) and are validated by simulations. For the i.s.i.d channel, the non-identical fading occurs at the receiver side only with $\mathbf{\Delta}_T = \mathbf{I}_2$ and $\mathbf{\Delta}_R = \text{diag}(\frac{1}{5}, \frac{9}{5})$. Its exact BEP curve is obtained from [16, eq.(27)]. The exact BEP for i.i.d channels is from [30]. As expected, the i.n.i.d channel yields the worst performance and the best performance is achieved over i.i.d channels. Note that this conclusion only holds when the sum of average channel gains is the same.

Next, we illustrate the effects of unequal channel fluctuation rates among different signalling branches on the error floors as $\gamma_0 \to \infty$. Fig. 5 shows the irreducible dominant PEP of the differential cyclic group code $\mathcal{V}_{2,4}$ in the system with two transmit antennas and one receive antenna. It is observed that under the same averaged fading correlation coefficient, i.e., $\rho = (\rho_1 + \rho_2)/2$, the error floor reduces as the difference on the fade rates between the two antennas increases.

In all the above figures, equal power allocation is assumed. We now illustrate in Figs. 6 and 7 the performance of the proposed transmit power allocation in a system with four transmit antennas and one receive antenna. An exponentially decaying average channel gain profile is used and character-

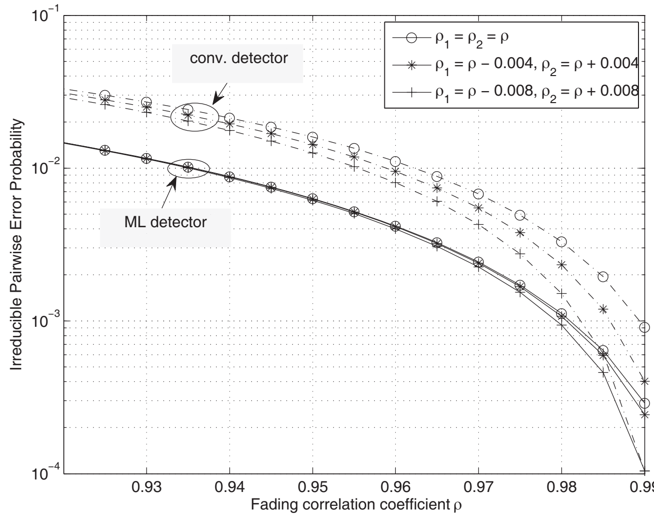


Fig. 5. Dominant pairwise error floor of differential cyclic group code $\mathcal{V}_{2,4}$ at $M = 2$ transmit and $N = 1$ receive antennas

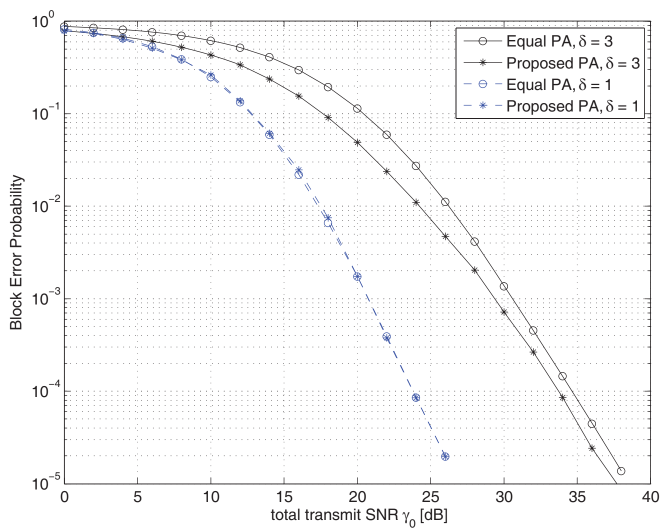


Fig. 6. BkEP performance of differential cyclic group code $\mathcal{V}_{4,16}$ with transmit power control at $M = 4$ transmit and $N = 1$ receive antenna in static fading.

ized by $\sigma_m^2 = e^{-\delta(m-1)}$, for $1 \le m \le 4$, in which $\delta \ge 0$ is the decay factor.

Fig 6 shows the simulated BkEP of the differential cyclic group code $\mathcal{V}_{4,16}$ with $M = 4$ and $L = 16$ at rate 1-bit/s/Hz [27, Table I]. The conventional detector is employed. It is seen that the proposed power allocation (34) cannot outperform equal power allocation when the average channel gains are only slightly unbalanced with $\delta = 1$. On the other hand, for highly unbalanced average channel gains with $\delta = 3$, the water-filling based power allocation can save $2 \sim 3$ dB total transmit power at a given BkEP around 10^{-2}. But the gain diminishes as the target BkEP reduces. This observation confirms our analytical finding in Section IV that equal power allocation is asymptotically optimal.

The BEP performance of the differential orthogonal code with $M = 4$, $P = 3$ and QPSK modulation at rate 1.5-bit/s/Hz based on the analysis (28) is presented in Fig. 7. We see that the gain of the proposed power allocation (36) over equal

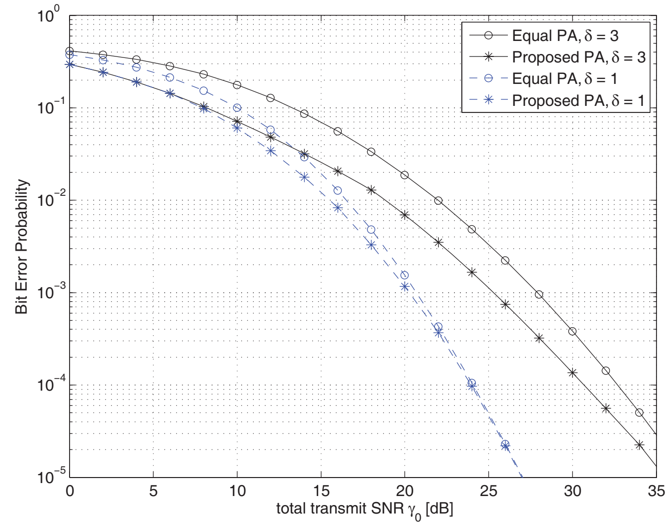


Fig. 7. BEP performance of differential OSTBC with transmit power control at $M = 4$ transmit and $N = 1$ receive antenna in static fading.

power allocation is more significant than in the cyclic group code case. This is because the power allocation for orthogonal codes aims at minimizing the bound of the overall bit error probability directly, whereas the one for cyclic group codes is obtained only through minimizing the bound of the dominant pairwise error probability with certain approximations.

VII. Conclusion

The effects of non-identical fading statistics in MIMO channels on the performance of DUSTM were investigated. Contrary to the detectors for the traditional i.i.d fading model, we found that the ML differential detector of DUSTM generally requires joint optimization of the current data matrix and the previously transmitted signal matrix. However, for DUSTM with cyclic group design, the ML detector is much simplified and is similar to the conventional detector but applies fading statistics-dependent weights. Based on the analysis of exact and asymptotic error probability for both cyclic group codes and orthogonal codes, we obtained several useful findings. Along with numerical results, we conclude that while the ML detector can significantly reduce the error floor over rapidly time-varying fading channels, the conventional detector is near-optimal at all SNR in static fading and low SNR in time-varying fading. In addition, the non-identical channel gain distribution degrades the error performance compared with the identical distribution for a same total received SNR. To exploit the non-identical fading parameters at the transmitter, we also presented a water-filling based transmit power control. It was shown to provide considerable improvement in error probability at low to moderate SNR region when the average channel gains are highly unbalanced. At sufficiently high SNR, equal power allocation is still optimal.

Appendix A
Proof of Proposition 1

By applying the result in [31], the characteristic function (CF) of the pairwise decision variable z_{ij} in the quadratic form of Gaussian vectors (14) can be written as [26]

$$\phi_{z_{ij}}(s) = \mathcal{E}\left[e^{-sz_{ij}}\right] = \frac{1}{\prod_{n=1}^{N}\det(\mathbf{I}+s\mathbf{\Lambda}_n\mathbf{\Omega}_{ij,n})} \quad (37)$$

Substituting (3) and (15) (or (16)) into (37), we obtain

$$\phi_{z_{ij}}(s) = \left(\prod_{n=1}^{N}\prod_{m=1}^{M} a_{mn}\left[\frac{1}{4}+\frac{1}{a_{mn}}-\left(\frac{s}{b_{mn}}-\frac{1}{2}\right)^2\right]\right)^{-1}, \quad (38)$$

where a_{mn} and b_{mn} are given in (18) and (19), respectively. After inverting the Laplace transform, we express $P_{e,ij}$ defined in (13) as [24, Appendix A]

$$P_{e,ij} = \frac{1}{2\pi j}\int_{-j\infty+\eta}^{j\infty+\eta}\frac{\phi_{z_{ij}}(s)}{s}ds, \quad (39)$$

where $\eta > 0$ is within the region of convergence. This integral can be solved using Cauchy's theorem in terms of residues:

$$P_{e,ij} = -\sum_{p_i>0}\text{Res}\left[\frac{\phi_{z_{ij}}(s)}{s},\ s=p_i\right], \quad (40)$$

where p_i's are all the positive poles of $\phi_{z_{ij}}(s)$. Finally, substituting (38) into (40) yields $P_{e,ij}$ expressed more explicitly in (17).

In the case of ML differential detection with $b_{mn} = 1$, we can choose $\eta = 1/2$ for the integration contour in (39). Then, with a change of variables, we obtain

$$P_{e,ij} = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \cdot\frac{ds}{(s+\frac{1}{2})\prod_{m=1}^{M}\prod_{n=1}^{N}a_{mn}\left(\frac{1}{4}+\frac{1}{a_{mn}}-s^2\right)}. \quad (41)$$

Now we let $s = jw$ in (41) and the integration becomes along the real axis. By further letting $w = \tan\theta/2$, an alternative expression of $P_{e,ij}$ in the form of finite integral is obtained in (20).

Appendix B
Proof of Proposition 2

The CF of the quadratic form of Gaussian vectors $z_p(\alpha)$ in (27) is given by

$$\phi_{z_p}(s) = \mathcal{E}\left[e^{-sZ_p(\alpha)}\right] = \frac{1}{\prod_{n=1}^{N}\prod_{i=1}^{K}(1+s\lambda_{n,i})^{u_{n,i}}}.$$

where $\{\lambda_{n,i}\}_{i=1}^{K}$ are the distinct eigenvalues of $\mathbf{\Lambda}_n\mathbf{\Omega}_p$ with multiplicity of $\{u_{n,i}\}_{i=1}^{K}$. Thus, the BEP in (26) can be obtained as [24]

$$P_b(\alpha) = -\sum_{\lambda_{n,i}<0}\text{Res}\left[\frac{\phi_{z_p}(s)}{s},\ s=-\frac{1}{\lambda_{n,i}}\right], \quad (42)$$

where the residues are evaluated at the positive poles of $\phi_{z_p}(s)/s$, that is $-1/\lambda_{n,i}$ with $\lambda_{n,i}$ being negative. Using a similar approach as in the proof [25, Corollary 1], we can show that the eigenvalues of $\mathbf{\Lambda}_n\mathbf{\Omega}_p$ are determined by the p-th information symbol $s_p = e^{j\theta_p}$, and do not rely on the other symbols in the data matrix $\mathbf{D}$, nor the previously transmitted signal matrix $\mathbf{S}_{-1}$. They are:

$$\lambda_{n,i} = \frac{\gamma_{mn}\cos(\alpha+\theta_p)}{\sqrt{P}} \pm\sqrt{\frac{\gamma_{mn}^2\cos^2(\alpha+\theta_p)}{P}+2\gamma_{mn}+1} \quad (43)$$

where $\lambda_{n,i} < 0$ for $1 \leq i \leq M$ and $\lambda_{n,i} > 0$ for $M+1 \leq i \leq 2M$. Substituting (43) into (42) and after some algebra, we obtain the closed-form expression of $P_b(\alpha)$ in Proposition 2.

References

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

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

[3] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Select. Areas Commun.*, vol. 18, no. 7, pp. 1169–1174, 2000.

[4] M. Tao and R. S. Cheng, "Differential space-time block codes," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, 2001.

[5] G. Ganesan and P. Stoica, "Differential detection based on space-time block codes," *Wireless Personal Comm.*, vol. 21, pp. 163–180, 2002.

[6] B. L. Hughes, "Optimal space-time constellations from groups," *IEEE Trans. Inform. Theory*, vol. 49, no. 2, pp. 401–410, Feb. 2003.

[7] B. Hassibi and B. M. Hochwald, "Cayley differential unitary space-time codes," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1485–1503, June 2002.

[8] W. Roh and A. Paulraj, "Outage performance of the distributed antenna systems in a composite fading channels," in *Proc. IEEE VTC'02 Fall*, 2002.

[9] H. Zhang and H. Dai, "On the capacity of distributed MIMO systems," in *Proc. CISS'04*, 2004.

[10] M. A. Jensen, M. D. Rice, and A. L. Anderson, "Unitary space-time coding for multi-antenna aeronautical telemetry transmission," to be published.

[11] J. N. Laneman and G. W. Wornell, "Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[12] S. Yiu, R. Schober, and L. Lampe, "Distributed space-time block coding," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1195–1206, July 2006.

[13] J. Luo, R. S. Blum, L. Cimini, L. Greenstein, and A. Haimovich, "Power allocation in a transmit diversity system with mean channel gain information," *IEEE Commun. Lett.*, vol. 9, no. 7, 2005.

[14] J. He and P. Y. Kam, "On the performance of orthogonal space-time block codes over independent, nonidentical Rayleigh/Ricean fading channels," in *Proc. IEEE GLOBECOM'06*, 2006.

[15] H. Zhao, Y. Gong, Y. L. Guan, C. L. Law, and Y. Tang, "Space-time block codes in Nakagami fading channels with non-identical m-distributions," in *Proc. IEEE WCNC'07*, Mar. 2007.

[16] M. Tao and P. Y. Kam, "Analysis of differential orthogonal space-time block codes over semi-identical MIMO fading channels," *IEEE Trans. Commun.*, vol. 55, no. 2, pp. 282–291, Feb. 2007.

[17] G. Ganesan and P. Stoica, "Space-time block codes: a maximum SNR approach," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1650–1656, May 2001.

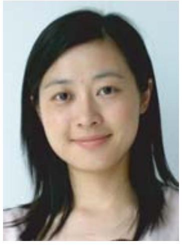
[18] R. Schober and L. Lampe, "Noncoherent receivers for differential space-time modulation," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 768–777, May 2002.

[19] E. Chiavaccini and G. M. Vitetta, "Further results on differential space-time modulation," *IEEE Trans. Commun.*, vol. 51, no. 7, pp. 1093–1101, July 2003.

[20] C. Ling, H. K. Li, and A. C. Kot, "Noncoherent sequence detection of differential space-time modulation," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2727–2734, Oct. 2003.

[21] C. Gao, A. M. Haimovich, and D. Lao, "Multiple-symbol differential detection for MPSK space-time block codes: design metric and performance analysis," *IEEE Trans. Commun.*, vol. 54, no. 8, pp. 1502–1510, Aug. 2006.

[22] P. Pun and P. Ho, "Fano multiple-symbol differential detectors for differential unitary space-time modulation," *IEEE Trans. Comun.*, vol. 55, no. 3, pp. 540–550, Mar. 2007.
[23] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1985.
[24] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receivers in Rayleigh-fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2383–2399, Sept. 2001.
[25] X. Cai and G. B. Giannakis, "Differential space-time modulation with eigen-beamforming for correlated MIMO fading channels," *IEEE Trans. Signal Processing*, vol. 54, pp. 1279–1288, 2006.
[26] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.
[27] A. Shokrollahi, B. Hassibi, B. M. Hochwald, and W. Sweldens, "Representation theory for high-rate multiple-antenna code design," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2335–2367, Sept. 2001.
[28] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
[29] M. K. Byun and B. G. Lee, "New bounds of pairwise error probability for space-time codes in rayleigh fading channels," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1484–1493, Aug. 2007.
[30] T. P. Soh, P. Y. Kam, and C. S. Ng, "Bit error probability for orthogonal space-time block codes with differential detection," *IEEE Trans. Commun.*, vol. 53, no. 11, pp. 1795–1798, Nov. 2005.
[31] G. L. Turin, "The characteristic function of hermitian quadratic forms in complex normal variables," *Biometrika*, vol. 47, pp. 199–201, 1960.



Meixia Tao (S'00-M'04) received the B.S. degree in Electronic Engineering from Fudan University, Shanghai, China, in 1999, and the Ph.D. degree in Electrical and Electronic Engineering from Hong Kong University of Science & Technology in 2003.

She is currently an Associate Professor at the Department of Electronic Engineering, Shanghai Jiao Tong University, China. From Aug. 2003 to Aug. 2004, she was a Member of Professional Staff in the Wireless Access Group at Hong Kong Applied Science & Technology Research Institute Co. Ltd, where she worked on the design of wireless local area networks. From Aug 2004 to Dec. 2007, she was with the Department of Electrical and Computer Engineering at National University of Singapore as an Assistant Professor. Her research interests include MIMO techniques, channel coding and modulation, dynamic resource allocation in wireless networks, and cooperative communications.

Dr. Tao is an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. She served as Track Co-Chair for IEEE ICCCN'07 held in August 2007 at Hawaii, USA, and IEEE ICCCAS'07 held in July 2007 at Fukuoka, Japan. She also served as a Technical Program Committee member for various conferences, including IEEE ICC (2006, 2007, 2008), IEEE WCNC (2007, 2008), IEEE GLOBECOM (2007), and IEEE VTC (2006-Fall, 2008-Spring).