High Frequency Differential MIMO: Basic Theory and Transceiver Architectures

Akbar M. Sayeed and John H. Brady
Department of Electrical and Computer Engineering
University of Wisconsin - Madison
Madison, WI, USA

***Abstract*—We propose a new multiple input multiple output (MIMO) transceiver architecture – Differential MIMO (D-MIMO) – that enables linear interference suppression between multiple spatially multiplexed and differentially encoded data streams. The D-MIMO transceiver architecture is particularly attractive in emerging high-frequency systems, such as millimeter-wave systems, in which the requirement of a phase-coherent local oscillator at the receiver can be challenging. A direct application of conventional linear interference suppression techniques is not possible with differential communication. Thus, we first develop a general model for D-MIMO systems, with a corresponding D-MIMO channel matrix, that forms the basis of the development in this paper. A surprising result is that a quasi-coherent version of the underlying channel matrix can also be estimated from the D-MIMO matrix, making conventional linear interference suppression possible as well. This leads to two D-MIMO transceiver architectures that are developed. Numerical results illustrate the promising and nearly identical performance of the proposed transceivers and the communication breakdown that can occur without interference suppression.**

***Index Terms*—Spatial Multiplexing, Interference Suppression, Differential Signaling, Millimeter-wave, Kronecker Product**

I. Introduction

There is growing interest in exploring higher frequencies ($>$5GHz) for meeting the Gigabit data rates and operational requirements of emerging wireless technologies. In particular, millimeter-wave (mmW) communication systems, ranging from 30GHz-300GHz, are emerging as a promising technology for 5G wireless [1]. In addition to the orders-of-magnitude larger bandwidth available at such high frequencies compared to existing systems, the small wavelengths make high-dimensional MIMO operation very attractive as well. Furthermore, the highly directional and quasi-optical nature of propagation at such high frequencies makes beamspace MIMO techniques and architectures naturally relevant [2]–[4]. However, many technical challenges need to be addressed before the full potential of mmW MIMO can be realized.

One challenging issue at mmW and high frequencies is phase-coherence between the transmitter and the receiver and the associated phase noise [5]. In single channel systems, an attractive solution is differential communication [6]. However, the use of differential communication is challenging in a MIMO system due to the interference between different spatial data streams. Differential space-time block coding schemes (e.g. [7]–[10]) do not support multiple spatial data streams. On the other hand, the differential spatial multiplexing scheme in [11] performs linear interference suppression based on the receive correlation matrix of the coherent MIMO channel. Finally, differential space-time coding schemes that support multiple spatial data streams without knowledge of the coherent MIMO channel (e.g. [12], [13]) require complex non-linear detectors. Linear interference suppression techniques that have been extensively studied for coherent MIMO systems, require knowledge of a coherent estimate of the MIMO channel matrix, and thus cannot be directly used.

In this paper, we propose new differential MIMO transceiver architectures that enable linear MIMO interference suppression within the context of differential communication. We first develop a general model for D-MIMO systems and identify a fundamental system equation, with a corresponding D-MIMO channel matrix, that forms the basis of the development in this paper. A surprising result is that a quasi-coherent version of the underlying channel matrix can also be estimated from the D-MIMO matrix, making conventional linear interference suppression feasible as well. This leads to the development of two D-MIMO transceiver architectures. Numerical results illustrate the promising and near-identical performance of the proposed transceivers compared to idealized systems in which there is no interference, and the communication breakdown that can occur without interference suppression. The results in this paper are based on a sub-system of the general model that suggests new avenues for future research.

II. Differential Signaling and Reception

Differential communication is typically used when a phase-coherent local oscillator is not available at the receiver, resulting in an unknown phase offset between the transmitter and receiver, and possibly even a sufficiently small frequency offset [6]. This problem is even more acute at high frequencies, such as mmW [5]. Consider a constant modulus constellation in which the transmitted symbols are of the form $s = Ae^{j\phi}$ for some given fixed A. Let $A = 1$ for simplicity. In a differential communication system, information is typically encoded in the *phase difference* $\Delta\phi$ between the current transmit symbol $s = s(t)$ and previous transmit symbol $s_\tau = s(t - T)$ where T is the symbol period; that is,

$$s = Ae^{j\phi} = e^{j\Delta\phi} s_\tau \ ; \ s_\tau = Ae^{j\phi_\tau} \ . \qquad (1)$$

We assume that the differential symbols $\Delta\phi$ are chosen randomly from a symmetric constellation, such as QPSK, and are independent across time. It follows that $e^{j\Delta\phi}$ is zero mean and independent of s_τ. Under these assumptions, the following

This work is partly supported by the NSF under grants 1247583 and 1444962, and the Wisconsin Alumni Research Foundation.

can be readily shown:

$$\begin{aligned} E[s_\tau] &= 0\ ;\ E[s] = E[e^{j\phi}]E[s_\tau] = 0 \\ |s|^2 &= |s_\tau|^2 = A^2 = 1 \\ ss_\tau^* &= e^{j\Delta\phi}|s_\tau|^2\ ;\ E[ss_\tau^*] = 0 \end{aligned} \quad (2)$$

which also specifies the second-order statistics of the entire sequence of symbols, under the assumption that the starting symbol, s_0, at time zero satisfies $E[s_0] = 0$ and $E[|s_0|^2] = A^2 = 1$, which can readily satisfied. The received signals and the differential measurements are

$$r = e^{j\phi_o}s + v\ ; r_\tau = e^{j\phi_o}s_\tau + v_\tau \quad (3)$$

$$rr_\tau^* = ss_\tau^* + sv_\tau^* + vs_\tau^* + vv_\tau^* = e^{j\Delta\phi} + w\ , \quad (4)$$

where v, v_τ, and $w = sv_\tau^* + vs_\tau^* + vv_\tau^*$ represent noise. A key assumption is that the unknown phase offset ϕ_o remains constant (or varies sufficiently slowly) over consecutive symbols, thereby enabling the detection of the differentially encoded symbols $\Delta\phi$ from rr_τ^* in (4).

III. Differential MIMO System Model

In this section, we develop a complex baseband model for the differential MIMO system. Consider a general $n \times n$ MIMO system with n transmit and receive antennas. Define the two transmitted signal vectors for the current symbol and the previous symbol corresponding to n differential symbols $\boldsymbol{\Delta\phi} = [\Delta\phi_1, \Delta\phi_2, \cdots, \Delta\phi_n]^T$:

$$\begin{aligned} \boldsymbol{s} &= [s_1, s_2, \cdots, s_n]^T && (5) \\ &= \boldsymbol{s}(t) = [s_1(t), s_2(t), \cdots, s_n(t)]^T && (6) \\ \boldsymbol{s}_\tau &= [s_{1\tau}, s_{2\tau}, \cdots, s_{n\tau}]^T && (7) \\ &= \boldsymbol{s}(t-T) && (8) \\ &= [s_1(t-T), s_2(t-T), \cdots, s_n(t-T)]^T && (9) \end{aligned}$$

The corresponding received signals $\boldsymbol{r}$, $\boldsymbol{r}_\tau$ are defined similarly. Finally, define the composite $2n \times 1$ transmitted and received signal vectors as

$$\boldsymbol{s}_c = \begin{bmatrix} \boldsymbol{s} \\ \boldsymbol{s}_\tau \end{bmatrix}\ ;\ \boldsymbol{r}_c = \begin{bmatrix} \boldsymbol{r} \\ \boldsymbol{r}_\tau \end{bmatrix}\ . \quad (10)$$

The overall MIMO system equation for the two symbol vectors and the composite vector is

$$\boldsymbol{r} = \boldsymbol{H}\boldsymbol{s}\ ,\ \boldsymbol{r}_\tau = \boldsymbol{H}_\tau\boldsymbol{s}_\tau\ ;\ \boldsymbol{r}_c = \boldsymbol{H}_c\boldsymbol{s}_c \quad (11)$$

where $\boldsymbol{H} = \boldsymbol{H}(t)$ and $\boldsymbol{H}_\tau = \boldsymbol{H}(t-T)$ and the $2n \times 2n$ composite channel matrix $\boldsymbol{H}_c$ is given by

$$\boldsymbol{H}_c = \begin{bmatrix} \boldsymbol{H} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H}_\tau \end{bmatrix} \quad (12)$$

A key assumption for differential communication is that $\boldsymbol{H} = \boldsymbol{H}_\tau$; that is, the channel does not change across two symbol durations.

A. A Fundamental Equation

A key observation is that the following differential measurements are possible at the receiver

$$\boldsymbol{R}_c = \boldsymbol{r}_c\boldsymbol{r}_c^H = \begin{bmatrix} \boldsymbol{r}\boldsymbol{r}^H & \boldsymbol{r}\boldsymbol{r}_\tau^H \\ \boldsymbol{r}_\tau\boldsymbol{r}^H & \boldsymbol{r}_\tau\boldsymbol{r}_\tau^H \end{bmatrix} \quad (13)$$

Using (11) the system equation for these differential measurements at the receiver (without noise) is

$$\boldsymbol{R}_c = \boldsymbol{r}_c\boldsymbol{r}_c^H = \boldsymbol{H}_c\boldsymbol{s}_c\boldsymbol{s}_c^H\boldsymbol{H}_c^H = \boldsymbol{H}_c\boldsymbol{Q}_c\boldsymbol{H}_c^H \quad (14)$$

where $\boldsymbol{Q}_c = \boldsymbol{s}_c\boldsymbol{s}_c^H$ is of the same form as (13) for $\boldsymbol{r}_c\boldsymbol{r}_c^H$ and represents the possibilities for differential transmission. Using (12) and expanding out (14) we get

$$\begin{aligned} \boldsymbol{R}_c &= \begin{bmatrix} \boldsymbol{r}\boldsymbol{r}^H & \boldsymbol{r}\boldsymbol{r}_\tau^H \\ \boldsymbol{r}_\tau\boldsymbol{r}^H & \boldsymbol{r}_\tau\boldsymbol{r}_\tau^H \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{H}\boldsymbol{s}\boldsymbol{s}^H\boldsymbol{H}^H & \boldsymbol{H}\boldsymbol{s}\boldsymbol{s}_\tau^H\boldsymbol{H}_\tau^H \\ \boldsymbol{H}_\tau\boldsymbol{s}_\tau\boldsymbol{s}^H\boldsymbol{H}^H & \boldsymbol{H}_\tau\boldsymbol{s}_\tau\boldsymbol{s}_\tau^H\boldsymbol{H}_\tau^H \end{bmatrix}\ . \end{aligned} \quad (15)$$

The matrix relation (15) represents a fundamental set of equations for understanding MIMO communication and interference suppression under differential signaling. Another version is obtained by vectorizing (14)

$$\boldsymbol{z}_c = \mathrm{vec}(\boldsymbol{R}_c) = [\boldsymbol{H}_c^* \otimes \boldsymbol{H}_c]\boldsymbol{x}_c\ ,\ \boldsymbol{x}_c = \mathrm{vec}(\boldsymbol{Q}_c) \quad (16)$$

where we have used the relation

$$\mathrm{vec}(\boldsymbol{ADB}) = [\boldsymbol{B}^T \otimes \boldsymbol{A}]\mathrm{vec}(\boldsymbol{D}) \quad (17)$$

where $\otimes$ denotes the Kronecker product [14]. An important special case of (17) for vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is

$$\mathrm{vec}(\boldsymbol{a}\boldsymbol{b}^H) = [\boldsymbol{b}^* \otimes \boldsymbol{a}]\mathrm{vec}(\boldsymbol{I}_1) = \boldsymbol{b}^* \otimes \boldsymbol{a} \quad (18)$$

which will be used later in the paper.

B. An Important Sub-System

We consider an important sub-system of (15) and (16) that forms the basis of the investigation in this paper:

$$\boldsymbol{r}\boldsymbol{r}_\tau^H = \boldsymbol{H}\boldsymbol{s}\boldsymbol{s}_\tau\boldsymbol{H}_\tau^H = \boldsymbol{H}\boldsymbol{s}\boldsymbol{s}_\tau\boldsymbol{H}^H\ , \quad (19)$$

where we have used the assumption that $\boldsymbol{H} = \boldsymbol{H}_\tau$. Vectorizing (19) we get

$$\begin{aligned} \boldsymbol{z} = \boldsymbol{H}_d\boldsymbol{x} \quad &; \quad \boldsymbol{H}_d = [\boldsymbol{H}_\tau^* \otimes \boldsymbol{H}] \\ \boldsymbol{z} = \mathrm{vec}(\boldsymbol{r}\boldsymbol{r}_\tau^H) \quad &, \quad \boldsymbol{x} = \mathrm{vec}(\boldsymbol{s}\boldsymbol{s}_\tau^H) \end{aligned} \quad (20)$$

and $\boldsymbol{H}_d$ is the D-MIMO channel matrix. We now specialize to the $n = 2$ case to get a concrete feel for the problem and which also forms the basis of the numerical results in this paper. We have

$$\boldsymbol{r}\boldsymbol{r}_\tau^H = \begin{bmatrix} r_1r_{1\tau}^* & r_1r_{2\tau}^* \\ r_2r_{1\tau}^* & r_2r_{2\tau}^* \end{bmatrix} \quad (21)$$

$$\boldsymbol{s}\boldsymbol{s}_\tau^H = \begin{bmatrix} s_1s_{1\tau}^* & s_1s_{2\tau}^* \\ s_2s_{1\tau}^* & s_2s_{2\tau}^* \end{bmatrix} \quad (22)$$

$$\boldsymbol{z} = \mathrm{vec}(\boldsymbol{r}\boldsymbol{r}_\tau^H) = \begin{bmatrix} r_1r_{1\tau}^* \\ r_2r_{1\tau}^* \\ r_1r_{2\tau}^* \\ r_2r_{2\tau}^* \end{bmatrix} \quad (23)$$

$$\boldsymbol{x} = \mathrm{vec}(\boldsymbol{s}\boldsymbol{s}_\tau^H) = \begin{bmatrix} s_1s_{1\tau}^* \\ s_2s_{1\tau}^* \\ s_1s_{2\tau}^* \\ s_2s_{2\tau}^* \end{bmatrix} \quad (24)$$

$$\begin{aligned} \boldsymbol{H}_d &= \boldsymbol{H}_\tau^* \otimes \boldsymbol{H} = \boldsymbol{H}^* \otimes \boldsymbol{H} && (25) \\ &= \begin{bmatrix} h_{11}^* & h_{12}^* \\ h_{21}^* & h_{22}^* \end{bmatrix} \otimes \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} && (26) \end{aligned}$$

$$= \begin{bmatrix} h_{11}^* \boldsymbol{H} & h_{12}^* \boldsymbol{H} \\ h_{21}^* \boldsymbol{H} & h_{22}^* \boldsymbol{H} \end{bmatrix} \quad (27)$$

$$= \begin{bmatrix} |h_{11}|^2 & h_{11}^* h_{12} & h_{12}^* h_{11} & |h_{12}|^2 \\ h_{11}^* h_{21} & h_{11}^* h_{22} & h_{12}^* h_{21} & h_{12}^* h_{22} \\ h_{21}^* h_{11} & h_{21}^* h_{12} & h_{22}^* h_{11} & h_{22}^* h_{12} \\ |h_{21}|^2 & h_{21}^* h_{22} & h_{22}^* h_{21} & |h_{22}|^2 \end{bmatrix} . \quad (28)$$

Remark 1 (System Rank) *$\boldsymbol{H}_d$ is full-rank if $\boldsymbol{H}$ is full-rank, which follows from the properties of the Kronecker product:* $\text{rank}(\boldsymbol{A} \otimes \boldsymbol{B}) = \text{rank}(\boldsymbol{A})\text{rank}(\boldsymbol{B})$ *[14].*

Remark 2 (Encoding/Decoding) *The first and last elements of $\boldsymbol{z}$ carry the information about the desired differential symbols, $\Delta\phi_1$ and $\Delta\phi_2$, contained in the first and last elements of $\boldsymbol{x}$. The remaining elements of $\boldsymbol{z}$ represent cross-terms that carry information about interference.*

Remark 3 (Interference) *If there is no inter-channel inference – $\boldsymbol{H}$ is diagonal – then there is no interference in the differential system (20) – $\boldsymbol{H}_d$ is diagonal. The off-diagonal entries of $\boldsymbol{H}_d$ represent the interference between the transmitted signals in $\boldsymbol{x}$ (see (24)) that corrupt the receiver measurements in $\boldsymbol{z}$ (see (23)).*

IV. Interference Suppression with Differential Reception

In this section, we develop an approach for linear interference suppression in MIMO systems that use differential encoding and decoding for the different spatial data streams. We explicitly describe our approach for the D-MIMO subsystem in (20); however, our results can be readily extended to the full D-MIMO system in (16). We start with the noisy underlying system equations (11)

$$\boldsymbol{r} = \sqrt{\rho}\boldsymbol{H}\boldsymbol{s} + \boldsymbol{v} \;;\; \boldsymbol{r}_\tau = \sqrt{\rho}\boldsymbol{H}_\tau \boldsymbol{s}_\tau + \boldsymbol{v}_\tau \quad (29)$$

$$\begin{aligned} \boldsymbol{r}\boldsymbol{r}_\tau^H &= \rho \boldsymbol{H}\boldsymbol{s}\boldsymbol{s}_\tau^H \boldsymbol{H}_\tau^H \\ &\quad + \sqrt{\rho}\boldsymbol{H}\boldsymbol{s}\boldsymbol{v}_\tau^H + \sqrt{\rho}\boldsymbol{v}\boldsymbol{s}_\tau^H \boldsymbol{H}_\tau^H + \boldsymbol{v}\boldsymbol{v}_\tau^H \end{aligned} \quad (30)$$

where $\boldsymbol{v} \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_n)$ and $\boldsymbol{v}_\tau \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_n)$ represent complex Gaussian noise vectors that are independent of each other and the signals $\boldsymbol{s}$ and $\boldsymbol{s}_\tau$, and ρ represents the SNR for each data stream. Vectorizing (30) yields the noisy version of the D-MIMO system equation (20)

$$\boldsymbol{z} = \rho \boldsymbol{H}_d \boldsymbol{x} + \boldsymbol{w} \quad (31)$$

$$\begin{aligned} \boldsymbol{w} &= \boldsymbol{w}_1 + \boldsymbol{w}_2 + \boldsymbol{w}_3 \\ &= \text{vec}(\sqrt{\rho}\boldsymbol{H}\boldsymbol{s}\boldsymbol{v}_\tau^H + \sqrt{\rho}\boldsymbol{v}\boldsymbol{s}_\tau^H \boldsymbol{H}_\tau^H + \boldsymbol{v}\boldsymbol{v}_\tau^H) \;. \end{aligned} \quad (32)$$

where $\boldsymbol{x} = \text{vec}(\boldsymbol{s}\boldsymbol{s}_\tau^H)$ is the vector of transmitted differential symbols, $\boldsymbol{z} = \text{vec}(\boldsymbol{r}\boldsymbol{r}_\tau^H)$ is a vector of received differential signals, and $\boldsymbol{w}$ is the effective noise vector that consists of three terms identified in (32). Signal and noise statistics are described in more detail in the next section.

Our goal is to design a $n^2 \times n^2$ (4×4 for the concrete case) matrix $\boldsymbol{F}_o$ that operates on the vector $\boldsymbol{z}$ to yield estimates of $\boldsymbol{x}$ in which the interference has been suppressed:

$$\boldsymbol{x}_{est} = \boldsymbol{F}_o \boldsymbol{z} \;. \quad (33)$$

We design $\boldsymbol{F}_o$ using the minimum mean squared error (MMSE) criterion, assuming knowledge of the D-MIMO channel matrix $\boldsymbol{H}_d$:

$$\begin{aligned} \boldsymbol{F}_o &= \arg\min_{\boldsymbol{F}} E[\|\boldsymbol{x}_{est} - \boldsymbol{x}\|^2] \\ &= \boldsymbol{H}_d^H \left(\rho^2 \boldsymbol{H}_d \boldsymbol{H}_d^H + \boldsymbol{\Sigma}_w\right)^{-1} \end{aligned} \quad (34)$$

where $\boldsymbol{\Sigma}_w = E[\boldsymbol{w}\boldsymbol{w}^H]$ is the covariance matrix of $\boldsymbol{w}$, and $\boldsymbol{H}_d \boldsymbol{H}_d^H = (\boldsymbol{H}_\tau^* \boldsymbol{H}_\tau^T \otimes \boldsymbol{H}\boldsymbol{H}^H)$. The differentially encoded transmitted symbols in $\boldsymbol{x}$ can then be estimated at the receiver by simply applying differential detectors, corresponding to the differential transmission scheme used, to the appropriate elements of $\boldsymbol{x}_{est}$; see Remark 2 for the $n = 2$ case.

A. Signal and Noise Statistics

We now characterize the second-order statistics of $\boldsymbol{x}$ and $\boldsymbol{w}$ in (31). We consider zero-mean signal constellations for the differential symbols, with different differential symbols independent across time and data streams. This results in the following second-order statistics for $\boldsymbol{s}$:

$$E[\boldsymbol{s}] = E[\boldsymbol{s}_\tau] = \boldsymbol{0} \;,\; E[\boldsymbol{s}\boldsymbol{s}_\tau^H] = \boldsymbol{0} \quad (35)$$

$$E[\boldsymbol{s}\boldsymbol{s}^H] = E[\boldsymbol{s}_\tau \boldsymbol{s}_\tau^H] = \boldsymbol{I}_n \quad (36)$$

which in turn results in the following second-order statistics for $\boldsymbol{x} = \text{vec}(\boldsymbol{s}\boldsymbol{s}_\tau^H)$

$$\begin{aligned} E[\boldsymbol{x}] &= E[\text{vec}(\boldsymbol{s}\boldsymbol{s}_\tau^H)] = \text{vec}(E[\boldsymbol{s}\boldsymbol{s}_\tau^H]) = \boldsymbol{0} \\ E[\boldsymbol{x}\boldsymbol{x}^H] &= E[\text{vec}(\boldsymbol{s}\boldsymbol{s}_\tau^H)\text{vec}(\boldsymbol{s}\boldsymbol{s}_\tau^H)^H] \\ &= E[(\boldsymbol{s}_\tau^* \otimes \boldsymbol{s})(\boldsymbol{s}_\tau^T \otimes \boldsymbol{s}^H)] = E[\boldsymbol{s}^* \boldsymbol{s}_\tau^T \otimes \boldsymbol{s}\boldsymbol{s}^H] \\ &= E[\boldsymbol{s}^* \boldsymbol{s}_\tau^T] \otimes E[\boldsymbol{s}\boldsymbol{s}^H] = \boldsymbol{I}_n \otimes \boldsymbol{I}_n = \boldsymbol{I}_{n^2} \;. \end{aligned} \quad (37)$$

Proposition 1 *Assuming that the signal and noise are independent, and using the assumptions on the statistics of $\boldsymbol{v}$ and $\boldsymbol{v}_\tau$, it can be shown that*

$$E[\boldsymbol{w}] = 0 \quad (38)$$

$$\begin{aligned} \boldsymbol{\Sigma}_w &= E[\boldsymbol{w}\boldsymbol{w}^H] \\ &= \rho\sigma^2(\boldsymbol{I}_n \otimes \boldsymbol{H}\boldsymbol{H}^H) + \rho\sigma^2(\boldsymbol{H}_\tau^* \boldsymbol{H}_\tau^T \otimes \boldsymbol{I}_n) \\ &\quad + \sigma^4 \boldsymbol{I}_{n^2} \end{aligned} \quad (39)$$

where the three terms in $\boldsymbol{\Sigma}_w$ in (39) represent the covariance matrices of the corresponding terms in (32).

The noise statistics follow from the following calculations on the joint statistics of $\boldsymbol{w}_1$, $\boldsymbol{w}_2$, and $\boldsymbol{w}_3$ in (32). Using (18), we first note that

$$\boldsymbol{w}_1 = \sqrt{\rho}\text{vec}(\boldsymbol{H}\boldsymbol{s}\boldsymbol{v}_\tau^H) = \sqrt{\rho}(\boldsymbol{v}_\tau^* \otimes \boldsymbol{H}\boldsymbol{s}) \quad (40)$$

$$\boldsymbol{w}_2 = \sqrt{\rho}\text{vec}(\boldsymbol{v}\boldsymbol{s}_\tau^H \boldsymbol{H}_\tau^H) = \sqrt{\rho}(\boldsymbol{H}_\tau^* \boldsymbol{s}_\tau^* \otimes \boldsymbol{v}) \quad (41)$$

$$\boldsymbol{w}_3 = \text{vec}(\boldsymbol{v}\boldsymbol{v}_\tau^H) = (\boldsymbol{v}_\tau^* \otimes \boldsymbol{v}) \;. \quad (42)$$

Now, the second-order statistics of $\{\boldsymbol{w}_i\}$ are

$$E[\boldsymbol{w}_1] = \sqrt{\rho}(E[\boldsymbol{v}_\tau^*] \otimes E[\boldsymbol{H}\boldsymbol{s}]) = \boldsymbol{0} \quad (43)$$

$$\begin{aligned} E[\boldsymbol{w}_1 \boldsymbol{w}_1^H] &= \rho E[(\boldsymbol{v}_\tau^* \otimes \boldsymbol{H}\boldsymbol{s})(\boldsymbol{v}_\tau^* \otimes \boldsymbol{H}\boldsymbol{s})^H] \\ &= \rho E[(\boldsymbol{v}_\tau^* \boldsymbol{v}_\tau^T \otimes \boldsymbol{H}\boldsymbol{s}\boldsymbol{s}^H \boldsymbol{H})] \\ &= \rho\sigma^2 E[\boldsymbol{v}_\tau^* \boldsymbol{v}_\tau^T] \otimes \boldsymbol{H} E[\boldsymbol{s}\boldsymbol{s}^H] \boldsymbol{H}^H \\ &= \rho\sigma^2 \boldsymbol{I}_n \otimes \boldsymbol{H}\boldsymbol{H}^H \end{aligned} \quad (44)$$

Similarly, we have

$$
\begin{aligned}
E[\boldsymbol{w}_2] &= E[\boldsymbol{w}_3] = \mathbf{0} && (45)\\
E[\boldsymbol{w}_2\boldsymbol{w}_2^H] &= \rho\sigma^2(\boldsymbol{H}_\tau^*\boldsymbol{H}_\tau^H \otimes \boldsymbol{I}_n) && (46)\\
E[\boldsymbol{w}_3\boldsymbol{w}_3^H] &= \sigma^2\boldsymbol{I}_n \otimes \sigma^2\boldsymbol{I}_n = \sigma^4\boldsymbol{I}_{n^2} && (47)
\end{aligned}
$$

Finally, it can be similarly shown that

$$
E[\boldsymbol{w}_1\boldsymbol{w}_2^H] = E[\boldsymbol{w}_1\boldsymbol{w}_3^H] = E[\boldsymbol{w}_2\boldsymbol{w}_3^H] = \mathbf{0} \tag{48}
$$

Combining the above calculations leads to the second-order statistics of $\boldsymbol{w}$ given in Prop. 1.

Proposition 2 *If $\boldsymbol{H}\boldsymbol{H}^H$ has the eigenvalue decomposition $\boldsymbol{H}\boldsymbol{H}^H = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H$ and $\boldsymbol{H}_\tau\boldsymbol{H}_\tau^H$ has the eigenvalue decomposition $\boldsymbol{H}_\tau\boldsymbol{H}_\tau^H = \boldsymbol{U}_\tau\boldsymbol{\Lambda}_\tau\boldsymbol{U}_\tau^H$, then the noise covariance matrix $\boldsymbol{\Sigma}_w$ admits the eigenvalue decomposition*

$$
\boldsymbol{\Sigma}_w = (\boldsymbol{U}_\tau^* \otimes \boldsymbol{U})\tilde{\boldsymbol{\Lambda}}(\boldsymbol{U}_\tau^* \otimes \boldsymbol{U})^H \tag{49}
$$

$$
\tilde{\boldsymbol{\Lambda}} = \rho\sigma^2(\boldsymbol{\Lambda} \oplus \boldsymbol{\Lambda}_\tau) + \sigma^4\boldsymbol{I}_{n^2} \tag{50}
$$

where $\boldsymbol{A} \oplus \boldsymbol{B} = (\boldsymbol{I} \otimes \boldsymbol{A}) + (\boldsymbol{B} \otimes \boldsymbol{I})$ is the Kronecker sum [15].

This follows from Theorem 13.16 in [15] and the fact that $\boldsymbol{U}$ and $\boldsymbol{U}_\tau$ are unitary, and thus $\boldsymbol{U}_\tau^* \otimes \boldsymbol{U}$ is also unitary. This result may be useful in analyzing the structure of $\boldsymbol{F}_o$ in (34).

B. Channel Estimation

In practice, $\boldsymbol{H}_d$ has to be estimated using training symbols and then an estimated version of $\boldsymbol{H}_d$ is plugged into (34) to determine $\boldsymbol{F}_o$. The training signals can be designed in a variety of ways. The simplest approach is to design the transmitted signals so that only one entry of $\boldsymbol{x}$ (see (24)) is non-zero in each differential training symbol; the corresponding column of $\boldsymbol{H}_d$ can then be estimated from the corresponding received differential measurements $\boldsymbol{z}$ (see (23) [16]. We present numerical results for $\boldsymbol{F}_o$ based on perfectly known $\boldsymbol{H}_d$ as well as estimated $\boldsymbol{H}_d$.

Note from (39) that we also need estimates of $\boldsymbol{H}\boldsymbol{H}^H$ and $\boldsymbol{H}_\tau^*\boldsymbol{H}_\tau^T$ to estimate $\boldsymbol{\Sigma}_w$ for $\boldsymbol{F}_o$ in (34). For the special case of interest $\boldsymbol{H}_\tau = \boldsymbol{H}$ we have

$$
\mathrm{vec}(\boldsymbol{H}\boldsymbol{H}^H) = [\boldsymbol{H}^* \otimes \boldsymbol{H}]\mathrm{vec}(\boldsymbol{I}) = \boldsymbol{H}_d\mathrm{vec}(\boldsymbol{I}) \tag{51}
$$

and thus the two matrices can be extracted from $\boldsymbol{H}_d$.

V. Quasi-Coherent Interference Suppression

In this section, we show that a quasi-coherent estimate of $\boldsymbol{H}$ can be obtained from $\boldsymbol{H}_d$ which can then be used for linear interference suppression on direct measurements $\boldsymbol{r}$ and $\boldsymbol{r}_\tau$ (rather than on $\boldsymbol{z} = \mathrm{vec}(\boldsymbol{r}\boldsymbol{r}_\tau^H)$) followed by differential detection from appropriate elements of $\boldsymbol{z}$.

A. Linear Interference Suppression at the Receiver

We have the following channel decomposition of $\boldsymbol{H}$

$$
\boldsymbol{H} = \boldsymbol{H}_o\boldsymbol{\Lambda}_\phi \tag{52}
$$

where $\boldsymbol{H}$ is the actual channel matrix

$$
\boldsymbol{H} = \begin{bmatrix} |h_{11}|e^{j\angle h_{11}} & |h_{12}|e^{j\angle h_{12}} \\ |h_{21}|e^{j\angle h_{21}} & |h_{22}|e^{j\angle h_{22}} \end{bmatrix} \tag{53}
$$

and $\boldsymbol{H}_o$ is what we can estimate from $\boldsymbol{H}_d$

$$
\boldsymbol{H}_o = \begin{bmatrix} |h_{11}| & |h_{12}|e^{j(\angle h_{12} - \angle h_{22})} \\ |h_{21}|e^{j(\angle h_{21} - \angle h_{11})} & |h_{22}| \end{bmatrix} \tag{54}
$$

and $\boldsymbol{\Lambda}_\phi$ is a diagonal matrix (that is unknown)

$$
\boldsymbol{\Lambda}_\phi = \mathrm{diag}(e^{j\angle h_{11}}, e^{j\angle h_{22}}) \ . \tag{55}
$$

To see how $\boldsymbol{H}_o$ can be estimated from $\boldsymbol{H}_d$, refer to (27). The first column of $h_{11}^*\boldsymbol{H}/|h_{11}|$ yields the first column of $\boldsymbol{H}_o$. Similarly, the second column of $h_{22}^*\boldsymbol{H}/|h_{22}|$ yields the second column of $\boldsymbol{H}_o$.

The MMSE filter matrix in this case is given by

$$
\begin{aligned}
\boldsymbol{F} &= \boldsymbol{H}^H(\rho\boldsymbol{H}\boldsymbol{H}^H + \sigma^2\boldsymbol{I}_n)^{-1} \\
&= \boldsymbol{\Lambda}_\phi^H\boldsymbol{H}_o^H(\rho\boldsymbol{H}_o\boldsymbol{H}_o^H + \sigma^2\boldsymbol{I}_n)^{-1} = \boldsymbol{\Lambda}_\phi^H\boldsymbol{F}_o \qquad (56)
\end{aligned}
$$

which operates on the baseband signal vector $\boldsymbol{r}$. We note that $\boldsymbol{F}_o$ in (56) is what can be computed at the receiver and used for interference suppression. Thus, processed signal vector from which the differentially encoded symbols are detected is given by

$$
\boldsymbol{y} = \boldsymbol{F}_o\boldsymbol{r} = \boldsymbol{F}_o\boldsymbol{H}\boldsymbol{s} + \boldsymbol{F}_o\boldsymbol{v} \ . \tag{57}
$$

We note the use of $\boldsymbol{F}_o$ (rather than $\boldsymbol{F}$) does not impact the ability to detect differential symbols since the i-th differentially encoded transmitted symbol in $s_i s_{i\tau}^*$ is detected from the product $y_i y_{i\tau}^*$. This corresponds to detecting the differentially encoded symbol vector via $\boldsymbol{y} \circ \boldsymbol{y}_\tau^*$ where $\circ$ denotes the Hadamard (element-wise) product.

B. Linear Interference Suppression at the Transmitter

Interference suppression using precoding at the transmitter is another attractive possibility. It turns out that $\boldsymbol{H}_o$ estimated at the receiver and fed back to the transmitter cannot be exploited due to the phase ambiguity. However, in reciprocal channels, if the transmitter first acts a receiver and estimates the channel matrix from differential measurements (based on training symbols from the receiver), it turns out that it results in the following decomposition of $\boldsymbol{H}$

$$
\boldsymbol{H} = \boldsymbol{\Lambda}_\phi\boldsymbol{H}_o \tag{58}
$$

In this case the transmitted signal is precoded as $\boldsymbol{s} \rightarrow \boldsymbol{G}\boldsymbol{s}$ where [4], [17]

$$
\begin{aligned}
\boldsymbol{G} &= \alpha\boldsymbol{F} \ , \ \alpha = \sqrt{\rho/\mathrm{tr}(\boldsymbol{F}\boldsymbol{\Lambda}_s\boldsymbol{F}^H)} \\
\boldsymbol{F} &= (\boldsymbol{H}^H\boldsymbol{H} + \zeta\boldsymbol{I})^{-1}\boldsymbol{H}^H \ , \ \zeta = \sigma^2/\rho \ , \qquad (59)
\end{aligned}
$$

and $\boldsymbol{s}$ is the transmitted symbol vector, ρ represents transmit power (SNR if $\sigma^2 = 1$) per data stream, and $\boldsymbol{\Lambda}_s = E[\boldsymbol{s}\boldsymbol{s}^H]$ is the diagonal covariance of transmitted symbols, which in our case is $\boldsymbol{\Lambda}_s = \boldsymbol{I}$. The composite system matrix with precoding is given by

$$
\boldsymbol{r} = \boldsymbol{H}\boldsymbol{G}\boldsymbol{s} + \boldsymbol{v} \tag{60}
$$

and the composite matrix $\boldsymbol{H}\boldsymbol{G}$ controls the interference. Note that in terms of $\boldsymbol{H}_o$, $\boldsymbol{F}$ is given by

$$
\boldsymbol{F} = (\boldsymbol{H}_o^H\boldsymbol{H}_o + \zeta\boldsymbol{I})^{-1}\boldsymbol{H}_o^H\boldsymbol{\Lambda}_\phi^* = \boldsymbol{F}_o\boldsymbol{\Lambda}_\phi^* \tag{61}
$$

where $\boldsymbol{F}_o$ is what we can actually compute based on the estimated $\boldsymbol{H}_o$ in (58). From (61) we note that the unknown phases in $\boldsymbol{\Lambda}_\phi^*$ are inconsequential from the viewpoint of differential signaling, and the receiver can directly detect the symbols differentially from $\boldsymbol{z} = \mathrm{vec}(\boldsymbol{r}\boldsymbol{r}_\tau^H)$ since interference suppression is done at the transmitter.

VI. Numerical Results

In this section, we present numerical results to illustrate the performance of the proposed D-MIMO transceiver architectures for an $n \times n$ MIMO system with $n = 2$ antennas.

Fig. 1 shows a diagram of the D-MIMO MMSE receiver (34), discussed in Sec. IV, that operates on the 4×1 differential measurements $\boldsymbol{z} = \mathrm{vec}(\boldsymbol{r}\boldsymbol{r}_\tau^H) = \boldsymbol{r}_\tau^* \otimes \boldsymbol{r}$ to detect the differential symbols. Results based on uncoded QPSK differential transmission for this receiver are presented in Fig. 2. The figures plot the probability of error P_e versus SNR for two D-MIMO MMSE receivers: one based on perfect channel state information (CSI) - perfect knowledge of $\boldsymbol{H}_d$, and one based on estimated $\boldsymbol{H}_d$ where the estimation is done via training symbols at the same SNR as that for data communication. The performance of a third D-MIMO receiver without interference suppression ($\boldsymbol{F}_o = \boldsymbol{I}_{n^2}$) is also shown for comparison. Finally, the performance of two ideal systems is shown for baseline comparison in which there is no interference: $\boldsymbol{H}$ is diagonal - see Remark 3. One is a coherent system corresponding to two non-interfering QPSK data streams, and the other is a corresponding differential system. The coherent system has the best performance and differential system has a 3dB loss compared to coherent system. The D-MIMO system with perfect CSI is next in line, followed by the D-MIMO with estimated channel. The worst performance is that of D-MIMO without interference suppression. Fig. 2(a)-(c) show the performance of the five systems for 3 different levels of interference. In Fig. 2(a), the interference is strongest: $|h_{12}|^2$ and $|h_{21}|^2$ are 3dB below $|h_{11}|^2 = |h_{22}|^2$, whereas in (b) the interference is 6dB below signal, and in (c) 10dB below signal. The P_e is computed numerically from 1000,000 symbols, and the phases of the entries of $\boldsymbol{H}$ change randomly every 1000 symbols. As evident, the D-MIMO receivers can deliver very competitive performance, whereas ignoring interference (D-MIMO w/o interference suppression) can result in unacceptably high P_e.

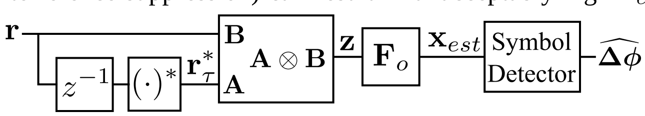
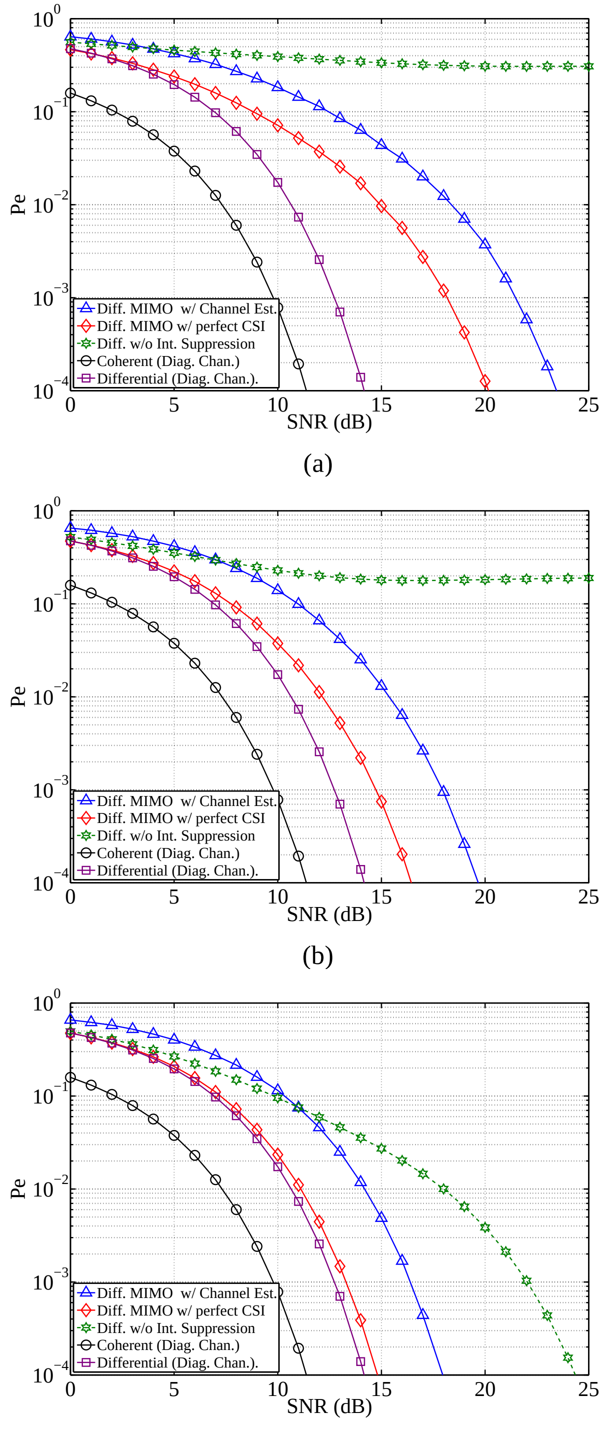


Fig. 1: D-MIMO MMSE receiver diagram

Fig. 3 shows a diagram of the of the quasi-coherent linear MMSE receiver discussed in Sec. V-A that performs linear MMSE interference suppression at the receiver before making the differential measurements to detect the differential symbols, where $\boldsymbol{A} \circ \boldsymbol{B}$ denotes the Hadamard product. Results parallel to Fig. 2 for this receiver are shown in Fig. 4. The two baseline ideal receivers, coherent and differential without interference, are the same and so is the D-MIMO receiver without interference suppression. The only difference is in the D-MIMO receivers with interference suppression: in this case, they are quasi-coherent linear MMSE receivers with perfect CSI ($\boldsymbol{H}$) and with estimated $\boldsymbol{H}$. The general trend is the same as in Fig. 2 and the performance of the two D-MIMO MMSE receivers is very comparable. The main difference seems to be



(a) (b) (c)

Fig. 2: P_e versus SNR for various receivers for different levels of interference power using interference suppression on differential measurements: (a) 3dB below signal, (b) 6dB below signal, (c) 10dB below signal.

that the D-MIMO receivers based on estimated $\boldsymbol{H}_d$ perform slightly worse than those based on estimated $\boldsymbol{H}_o$.

VII. Conclusion

We have presented two promising D-MIMO transceiver architectures that enable interference suppression in conjunction with spatially multiplexed differential signaling. While

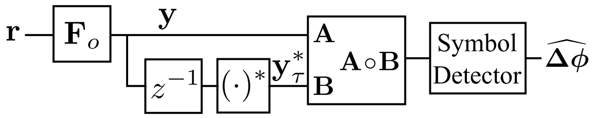


Fig. 3: Quasi-coherent linear MMSE receiver diagram

we have not explicitly discussed it, the proposed transceivers can deal with small frequency offsets as well. The results presented in this paper are based on a sub-system (19) of the general D-MIMO model in (15) and (16) that offers a rich structure and array of possibilities for further research. Extensions to multiuser transceivers and wideband scenarios that explicitly account for multipath propagation is another fruitful direction. The sampled approach to wideband MIMO channel modeling in [18] could be particularly relevant in this context. Development of the D-MIMO concept in beamspace for high-dimensional MIMO systems [2]–[4], such as those encountered at mmW frequencies, is also a promising direction. Finally, we note that at high frequencies such as mmW, the differential measurements at the receiver can also be realized in an analog fashion using interferometers thereby obviating the need for a local oscillator at the receiver [19].

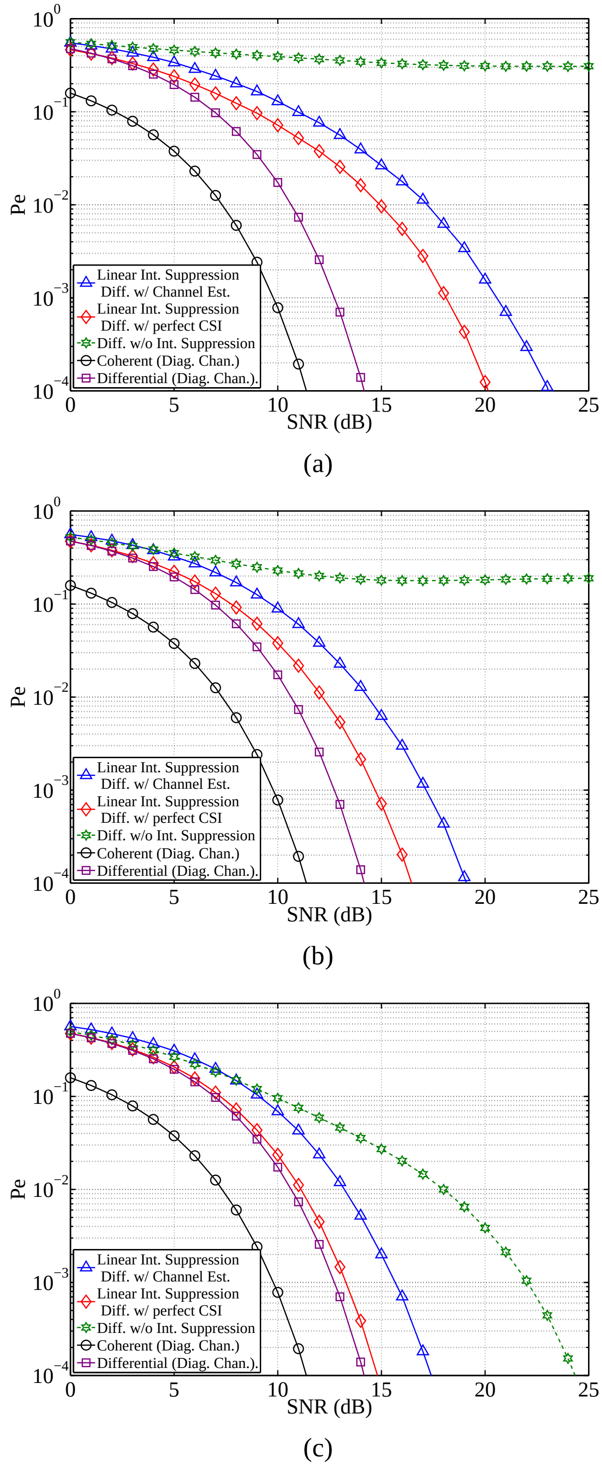


Fig. 4: P_e versus SNR for various receivers for different levels of interference power and using quasi-coherent interference suppression: (a) 3dB below signal, (b) 6dB below signal, (c) 10dB below signal.

References

[1] Z. Pi and F. Khan, "An introduction to millmeter-wave mobile broadband systems," *IEEE Commununications Magazine*, pp. 101–107, June 2011.

[2] A. Sayeed and N. Behdad, "Continuous aperture phased MIMO: Basic theory and applicatons," *Proc. 2010 Annual Allerton Conference on Communications, Control and Computers*, pp. 1196–1203, Sep. 2010.

[3] J. Brady, N. Behdad, and A. Sayeed, "Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis and measurements," *IEEE Transactions on Antenna and Propagation*, pp. 3814–3827, July 2013.

[4] A. Sayeed and J. Brady, "Beamspace MIMO for high-dimensional multiuser communication at millimeter-wave frequencies," *2013 IEEE Global Communications Conference*, pp. 3785–3789, Dec. 2013.

[5] H. Mehrpouyan, M. R. Khanzadi, M. Matthaiou, A. Sayeed, R. Schober, and Y. Hua, "Improving bandwidth efficiency in E-band communication systems," *IEEE Commun. Mag.*, vol. 1, pp. 121–128, Mar 2014.

[6] J. G. Proakis, *Digitial Communications*, McGraw Hill, New York, 4th edition, 2002.

[7] Vahid Tarokh and Hamid Jafarkhani, "A Differential Detection Scheme for Transmit Diversity," *IEEE Journal Selected Areas in Comm.*, vol. 18, no. 7, pp. 1169–1174, July 2000.

[8] L.-Y. Song and A.G. Burr, "Differential Quasi-Orthogonal Space-Time Block Codes," *IEEE Trans Wireless Comm.*, vol. 6, no. 1, pp. 64–68, Jan 2007.

[9] S.J. Alabed, J.M. Paredes, and A.B. Gershman, "A Low Complexity Decoder for Quasi-Orthogonal Space Time Block Codes," *IEEE Trans Wireless Comm.*, vol. 10, no. 3, pp. 988–994, March 2011.

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

[11] S.K. Cheung and R. Schober, "Differential spatial multiplexing," *IEEE Trans Wireless Comm.*, vol. 5, no. 8, pp. 2127–2135, Aug 2006.

[12] L.-Y. Song, A.G. Burr, and R.C. de Lamare, "Differential Bell-Labs Layered Space Time Architectures," in *Communications, 2007. ICC '07. IEEE International Conference on*, June 2007, pp. 5093–5097.

[13] M. Rezk and B. Friedlander, "Full-rate full-diversity differential mimo," in *Signals, Systems and Computers, 2008 42nd Asilomar Conference on*, Oct 2008, pp. 1739–1743.

[14] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circ. and Syst.*, vol. 25, no. 9, pp. 772–781, Sep. 1978.

[15] Alan J. Laub, *Matrix Analysis For Scientists And Engineers*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2004.

[16] J. Kotecha and A. M. Sayeed, "Optimal signal design for estimation of correlated MIMO channels," *IEEE Trans. Signal Processing*, pp. 546–557, Feb. 2004.

[17] M. Joham, W. Utschick, and J. Nossek, "Linear transmit processing in MIMO communication systems," *IEEE Trans. Signal Processing*, pp. 2700–2712, Aug. 2005.

[18] A. Sayeed and T. Sivanadyan, *Wireless Communication and Sensing in Multipath Environments using Multi-antenna Transceivers*, Handbook on Array Processing and Sensor Networks (K. J. R. Liu and S. Haykin, Eds.), IEEE-Wiley, 2010.

[19] R. Ridgway, D. Nippa, and S. Yen, "Data transmission using differential phase-shift keying on a 92 Ghz carrier," *IEEE Trans.s on Microwave Theory and Tech.*, pp. 3117–3126, Nov. 2010.