Zero-Forcing DFE Transceiver Design Over Slowly Time-Varying MIMO Channels Using ST-GTD

Chih-Hao Liu, *Student Member, IEEE*, and Palghat P. Vaidyanathan, *Fellow, IEEE*

***Abstract*—This paper considers the optimization of transceivers with decision feedback equalizers (DFE) for slowly time-varying memoryless multi-input multi-output (MIMO) channels. The data vectors are grouped into space-time blocks (ST-blocks) for the spatial and temporal precoding to take advantage of the diversity offered by time-varying channels. The space-time generalized triangular decomposition (ST-GTD) is proposed for application in time-varying channels. Under the assumption that the instantaneous channel state information at the transmitter (CSIT) and receiver (CSIR), and the channel prediction are available, we also propose the space-time geometric mean decomposition (ST-GMD) system based on ST-GTD. Under perfect channel prediction, the system minimizes both the arithmetic MSE at the feedback detector, and the average un-coded bit error rate (BER) in moderate high signal to noise ratio (SNR) region. For practical applications, a novel ST-GTD based system which does not require channel prediction but shares the same asymptotic BER performance with the ST-GMD system is also proposed. At the moderate high SNR region, our analysis and numerical results show that all the proposed systems have better BER performance than the conventional GMD-based systems over time-varying channels; the average BERs of the proposed systems are non-increasing functions of the ST-block size.**

***Index Terms*—Generalized triangular decomposition, GMD, space-time GTD, time-varying channels, transceivers.**

I. Introduction

In recent years, multi-input multi-output (MIMO) transceiver design has received a great deal of attention [1]–[3]. Most of the research on MIMO transceiver design focuses on time-invariant channels [4]–[8]. In practice, the wireless channels are time-varying due to users' mobility. In this paper, we consider transceiver design based on the block fading model in which the MIMO channel is constant over the coherence (block) interval of N_c symbol vectors. The channel varies across different coherence intervals independently or according to Jakes' model [9], [10]. Zero-forcing constraint is assumed throughout the paper. The case without zero-forcing is more involved, and is currently under study.

To exploit the array gain for the full channel capacity, both channel state information (CSI) at the transmitter (CSIT) and the receiver (CSIR) are required [3], [10]. When the channel varies at a much slower rate compared to the data rate of the systems, CSIT can be obtained from the receiver via feedback mechanism. However, the overhead becomes too large if the channel is varying at a faster rate. In time division duplex (TDD) systems, the uplink and downlink are multiplexed on the same channel, so channel reciprocity holds. Hence, the transmitter can estimate its own CSI at current time slot using the received signal from the reverse link, and use the estimated CSI to transmit data at next time slot provided that the channel does not change significantly [11]–[15]. Wiener filter prediction can further be exploited to improve the accuracy of CSIT [9], [11]. Both feedback and TDD schemes can offer instantaneous CSI at transmitter and receiver.

For time invariant channels, the geometric mean decomposition (GMD) based systems with "zero-forcing" and "minimum mean-square error (MSE)" decision feedback structures [4]–[7], are known to minimize the arithmetic mean (over the spatial domain) of the expected MSE at the input of the decision device and the average bit error rate (BER) in high signal-to-noise ratio (SNR) [7]. Moreover, [5] shows that the GMD-based system with zero-forcing constraint achieves optimal channel throughput asymptotically in high SNR. Unlike the singular value decomposition based systems which require bit allocation to achieve the optimal average BER [16], the GMD-based systems do not require bit allocation since all the effective subchannels have the same SINR [5].

In the case of time-varying channels, different data blocks pass through MIMO channels with different channel coefficients. If instantaneous CSIT and CSIR are available, the GMD-based system can be applied directly to time-varying channels. However, its average BER is not minimized since different coherence blocks have different arithmetic MSEs at the feedback loop detector. In [17], we proposed the GMD transceiver with a superimposed channel-independent temporal precoder (GMD-TP) which also only requires instantaneous CSIT and CSIR. We took the space-GMD and introduced the channel-independent temporal precoder to construct the GMD-TP. The temporal precoder equalizes the MSEs and hence SNRs across different coherence blocks (intervals) so that the average BER per space-time block (ST-block) is minimized.

In this paper, based on the generalized triangular decomposition (GTD) [18], we develop space-time GTD (ST-GTD) for the decomposition of time-varying MIMO channels which does GMD on the spatial domain and GTD on temporal domain. Using the special case of ST-GTD, namely ST-GMD which does GMD on both spatial and temporal domains, we develop the ST-GMD transceivers with zero-forcing constraint.

Manuscript received March 23, 2010; accepted August 02, 2010. Date of publication August 16, 2010; date of current version October 13, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Xiang-Gen Xia. This work is supported in parts by the ONR Grant N00014-08-1-0709 and California Institute of Technology.

The authors are with California Institute of Technology(Caltech), Pasadena, CA 91125 USA (e-mail: chliu@caltech.edu; ppvnath@systems.caltech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2010.2066973

The design of ST-GTD transceivers requires instantaneous CSIR, CSIT and channel prediction. The required prediction length depends on the size of a ST-block. Similar issues of channel prediction have been studied in several papers, e.g., [9] and [19]. The Wiener filter theory is usually adopted for the prediction of future channel coefficients based on the previous channel estimations. The accuracy of prediction depends highly on the channel model. Under perfect channel prediction assumption, ST-GMD transceiver is shown to jointly minimize the arithmetic MSE in each ST-block (which consists of several coherence blocks), and minimize the average per ST-block BER in high SNR. Next, in consideration of the feasibility of channel prediction, a causal ST-GTD based transceiver (CST-GTD) with stationary temporal processing is also proposed here. It does not require channel prediction because its temporal precoder is stationary. It is shown that the CST-GTD has smaller arithmetic MSE and average BER than the conventional GMD-based system in the high SNR region. The simulation also shows that the BER performance of CST-GTD approximates that of the ST-GMD transceiver asymptotically. In any case, ST-GMD transceiver serves as a performance benchmark for the general class of ST-GTD transceiver, including CST-GTD.

The novelty of the ST-GTD transceiver is the incorporation of the temporal precoder and the newly proposed "nested-loop" receiver structure. For each ST-block, these two components not only redistribute the MSEs among blocks, but also reduce the arithmetic MSE. This is in contrast to the linear block precoder in [20] and [21], and the temporal precoder in [17], which keep the same arithmetic MSE while equalizing the MSEs. At the moderate high SNR region, our analysis and numerical results show that all the proposed systems have better BER performance than the conventional GMD-based systems over time-varying channels; the average BERs of the proposed systems are non-increasing functions of the ST-block size. Moreover, our analysis shows that if the block size is a power of two, i.e., 2^n, then the average BERs of the proposed systems are non-increasing functions of n at the high SNR region and non-decreasing functions of n at the low SNR region. Our numerical studies in Section VI also demonstrate the use of channel prediction for ST-GMD transceivers in Jakes' channel model. In the cases that channel prediction is accurate enough, the performance of ST-GMD transceivers with imperfect channel prediction is still very close to that with perfect channel prediction.

The sections are structured as follows. In Section II, we introduce the time-varying channel model and review the GTD theorem [18]. The GMD-based DFE transceiver with the zero-forcing constraint [5], [7] is reviewed. In Section III-A, we develop space-time GTD based on the spatial GMD. Section III-B is devoted to the derivation the optimal ST-GTD transceiver which minimizes the arithmetic MSE. A practical suboptimal ST-GTD transceiver which does not require channel prediction is proposed in Section IV. In Section V, we analyze the performance of the proposed transceivers. Numerical examples of BER performances are given in Section VI. Concluding remarks are given in Section VII.

Notations: Upper case bold letters are reserved for matrices and lower case bold letters for vectors. $(\cdot)^T$ and $(\cdot)^\dagger$ denote the transpose and the conjugate transpose, respectively. x_i or $[\mathbf{x}]_i$ denotes the ith element of a vector $\mathbf{x}$; $A_{i,j}$ or $[\mathbf{A}]_{i,j}$ denotes the (i,j)th element of a matrix $\mathbf{A}$. $\mathbf{I}_M$ denotes the $M \times M$ identity matrix. $E(\cdot)$ stands for expectation. $\mathrm{diag}(\mathbf{x})$ is a diagonal matrix with the entries of $\mathbf{x}$ on the diagonal. $tr(\cdot)$ stands for trace.

II. Preliminaries and Reviews

A. System Model

In this paper, we consider the narrowband block fading MIMO channel model [10]. The channel remains constant over the coherence period of N_c transmitted signal vectors and varies independently [10] or according to Jakes' model [9] across different coherence intervals. For simplicity of analysis, we just pick one transmitted signal vector from each coherence block since the transmitted signal vectors in the same block go through the same channel. The channel model is given by

$$\mathbf{y}(k) = \mathbf{H}(k)\mathbf{x}(k) + \mathbf{w}(k) \tag{1}$$

where k is the coherence block index, $\mathbf{H}(k)$ is a $J \times N$ rank M channel matrix, and $\mathbf{x}(k)$ is an $N \times 1$ transmitted signal vector. The elements of $\mathbf{H}(k)$ are i.i.d. Gaussian random variables. In Jakes' model, the (i,j)th coefficients of the channel matrices from different k are related by

$$E\left[H_{i,j}(k)H^*_{i,j}(k')\right] = J_0\left(2\pi f_d |k-k'| N_c T_s\right) \tag{2}$$

where $J_0(\cdot)$ is the zeroth order Bessel function of first kind, f_d the Doppler spread and T_s the symbol period. The noise $\mathbf{w}(k)$ is a $J \times 1$ Gaussian random process vector with $E(\mathbf{w}(k)) = \mathbf{0}$ and $E(\mathbf{w}(k)\mathbf{w}^\dagger(k')) = \sigma_w^2 \delta(k-k')\mathbf{I}_J$. And $\mathbf{y}(k)$ is the $J \times 1$ received signal vector. At each coherence interval, $\mathbf{H}(k)$ is assumed to be known to the transmitter and receiver.

B. GTD Decomposition

In the following, we give a brief review of GTD theorem [18] and its application for the design of the GMD-based zero-forcing DFE transceiver [5].

Definition 1: Additive Majorization [8], [23]: For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ where $\mathbf{x} = [x_1, x_2, \cdots, x_n]$ and $\mathbf{y} = [y_1, y_2, \cdots, y_n]$, we say $\mathbf{x}$ is additively majorized by $\mathbf{y}$, and denote it as $\mathbf{x} \prec_+ \mathbf{y}$, if

$$\sum_{i=1}^{k} x_{[i]} \le \sum_{i=1}^{k} y_{[i]}, \quad \text{whenever} \quad 1 \le k \le n$$

and equality holds when $k = n$. Here, "[i]" denotes the ith largest component of the vector.

Definition 2: Multiplicative Majorization [23], [24]:: For $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}_+^n$ where $\mathbf{x} = [x_1, x_2, \cdots, x_n]$ and $\mathbf{y} = [y_1, y_2, \cdots, y_n]$, we say $\mathbf{x}$ is multiplicatively majorized by $\mathbf{y}$, and denote it as $\mathbf{x} \prec_\times \mathbf{y}$, if

$$\prod_{i=1}^{k} x_{[i]} \le \prod_{i=1}^{k} y_{[i]}, \quad \text{whenever} \quad 1 \le k \le n$$

and equality holds when $k = n$.

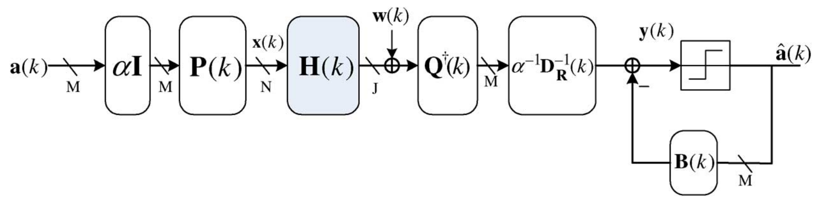


Fig. 1. The GMD-based system.

Theorem 1: The generalized triangular decomposition [18]: Let $\mathbf{H} \in \mathbb{C}^{J\times N}$ have rank M with non-zero singular values $\mathbf{d}_H = [\sigma_{H,0}, \sigma_{H,1}, \cdots, \sigma_{H,M-1}]^T$. Then there exists an upper triangular matrix $\mathbf{R} \in \mathbb{C}^{M\times M}$, and semi-unitary matrices $\mathbf{P} \in \mathbb{C}^{N\times M}$ and $\mathbf{Q} \in \mathbb{C}^{J\times M}$ in which all columns are orthonormal, such that $\mathbf{H} = \mathbf{QRP}^\dagger$ if and only if $\mathbf{r} \prec_\times \mathbf{d}_H$ where $r_i = |R_{ii}|$.

Suppose $R_{ii} = |R_{ii}|e^{j\theta_i}$. Without loss of generality, we can make the diagonal entries of $\mathbf{R}$ real and positive by extracting $e^{j\theta_i}$ from the ith row of $\mathbf{R}$ and multiplying the ith column of $\mathbf{Q}$ by $e^{j\theta_i}$. If one chooses $R_{ii} = \left(\prod_{i=0}^{M-1}\sigma_{H,i}\right)^{1/M}$, then GTD is reduced to GMD.

C. GMD-Based Transceivers With Zero-Forcing Constraint

Fig. 1 shows the GMD transceiver, which has been shown to be optimal in average BER at high SNR for linear time invariant channels [7]. Since both the transmitter and receiver have perfect CSI at current block time k, the $N\times M$ precoding matrix $\mathbf{P}(k)$, and the $J\times M$ feedforward matrix $\mathbf{Q}(k)$ can be obtained from the GMD of $\mathbf{H}(k)$ which is

$$\mathbf{H}(k) = \mathbf{Q}(k)\mathbf{R}(k)\mathbf{P}^\dagger(k) \tag{3}$$

where $\mathbf{P}(k) \in \mathbb{C}^{N\times M}$ and $\mathbf{Q}(k) \in \mathbb{C}^{J\times M}$ with orthonormal columns. $\mathbf{R}(k) \in \mathbb{C}^{M\times M}$ is an upper triangular matrix with $\mathbf{r}(k)$ on the diagonal. The $M\times 1$ vector $\mathbf{r}(k)$ has equal elements

$$r_i(k) = \sigma_k = \left(\prod_{i=0}^{M-1}\sigma_{H,i}(k)\right)^{1/M} \tag{4}$$

where $\sigma_{H,i}(k)$ is the ith singular value of $\mathbf{H}(k)$. The $M\times M$ feedback matrix $\mathbf{B}(k)$ is given by

$$\mathbf{B}(k) = \mathbf{D}_\mathbf{R}^{-1}(k)\mathbf{R}(k) - \mathbf{I}_M \tag{5}$$

where $\mathbf{D}_\mathbf{R}(k) = \mathrm{diag}(\mathbf{r}(k))$. $\mathbf{a}(k)$ is an $M\times 1$ symbol vector from the kth block with each element $[\mathbf{a}(k)]_i$ chosen from the alphabet χ of finite size. We assume $E(\mathbf{a}(k)\mathbf{a}^\dagger(k')) = \sigma_a^2\delta(k-k')\mathbf{I}_M$. The gain α is chosen to satisfy the total transmitting power constraint

$$P_0 = \mathrm{tr}\left(E\left(\mathbf{x}(k)\mathbf{x}^\dagger(k)\right)\right) \tag{6}$$

and hence satisfies $\alpha = \sqrt{P_0/M\sigma_a^2}$.

If there is no error propagation in the DFE loop, the received signal vector in front of the detector is given by

$$\mathbf{y}(k) = \mathbf{a}(k) + \mathbf{e}(k) \tag{7}$$

where $\mathbf{e}(k) = \alpha^{-1}\mathbf{D}_\mathbf{R}^{-1}(k)\mathbf{Q}^\dagger(k)\mathbf{w}(k)$. The error covariance of $\mathbf{e}(k)$ is

$$\mathbf{R}_{\mathbf{ee}} = \frac{\sigma_w^2}{\alpha^2}\mathrm{diag}\left(|\mathbf{r}(k)|\right)^{-2}. \tag{8}$$

The total MSE of the kth block at the detector is

$$\xi_{\mathrm{gmd}}(k) = \mathrm{tr}(\mathbf{R}_{\mathbf{ee}}) = \frac{\sigma_w^2}{\alpha^2}M\left(\prod_{i=0}^{M-1}\frac{1}{\sigma_{H,i}^2(k)}\right)^{1/M}. \tag{9}$$

III. Space-Time GTD Transceivers

A. Space-Time GTD

To facilitate space-time processing for the later sections, K blocks of symbol vectors are grouped into one space-time block as

$$\mathbf{a}_m = \left[\mathbf{a}^T(mK) \quad \ldots \quad \mathbf{a}^T(mK+K-1)\right]^T \tag{10}$$

where m is the ST-block index. The symbols, m and mK, will be omitted for convenience. The equivalent MIMO channel matrix for the mth ST-block is a $KJ\times KN$ block diagonal matrix given by

$$\mathbf{H} = \mathrm{diag}\left(\mathbf{H}(0), \mathbf{H}(1), \ldots, \mathbf{H}(K-1)\right). \tag{11}$$

Let $0 \le k \le K-1$. If GMD is applied to each $\mathbf{H}(k)$ separately (in spatial domain), we have $\mathbf{H}(k) = \mathbf{Q}(k)\mathbf{R}(k)\mathbf{P}^\dagger(k)$ as (3). $\mathbf{H}$ can be decomposed as

$$\mathbf{H} = \mathbf{QRP}^\dagger \tag{12}$$

where $\mathbf{Q}, \mathbf{P}$ and $\mathbf{R}$ are block diagonal matrices with $\mathbf{Q}(k), \mathbf{P}(k)$ and $\mathbf{R}(k)$ on the diagonals, respectively. Let

$$\mathbf{d} = [\sigma_0, \ldots, \sigma_{K-1}]^T \tag{13}$$

where σ_i is defined in (4). $\mathbf{R}$ can be expressed as

$$\mathbf{R} = \mathbf{D}_\mathbf{R}(\mathbf{I}_{KM} + \mathbf{B}) \tag{14}$$

where $\mathbf{D}_\mathbf{R} = \boldsymbol{\Sigma}\otimes\mathbf{I}_M$, $\boldsymbol{\Sigma} = \mathrm{diag}(\mathbf{d})$, and $\mathbf{B}$ is a block diagonal matrix with $\mathbf{B}(k)$ on the diagonal. $\mathbf{B}(k)$ are strictly upper triangular $M\times M$ matrices given by (5).

Since $\boldsymbol{\Sigma}$ is a diagonal matrix consisting of positive entries σ_k, these are also the singular values. Therefore, by Theorem 1, we can decompose $\boldsymbol{\Sigma}$ by GTD as

$$\boldsymbol{\Sigma} = \mathbf{Q}_1\mathbf{R}_1\mathbf{P}_1^\dagger, \tag{15}$$

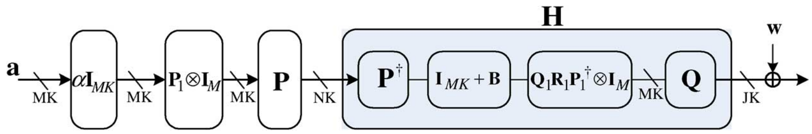


Fig. 2. The transmitter of the ST-GTD transceiver and the channel.

where $\mathbf{P}_1$ and $\mathbf{Q}_1$ are $K \times K$ unitary matrices, and $\mathbf{R}_1$ is a $K \times K$ upper triangular matrix. The necessary and sufficient condition for the GTD in (15) to be possible is

$$\mathbf{r}_1 \prec_{\times} \mathbf{d} \tag{16}$$

in which $\mathbf{r}_1$ is a $K \times 1$ vector consisting of diagonal elements of $\mathbf{R}_1$ and $\mathbf{d}$ is given by (13). We refer to the GTD of $\mathbf{\Sigma}$ as the temporal domain GTD because the σ_k depend on $\mathbf{H}(k)$ and the decomposition needs all σ_k all at one time. By (12), (14) and (15), the rank MK block diagonal matrix $\mathbf{H}$ of the from (11) can be decomposed as

$$\mathbf{H} = \mathbf{Q}\left(\left(\mathbf{Q}_1\mathbf{R}_1\mathbf{P}_1^{\dagger} \otimes \mathbf{I}_M\right)(\mathbf{I}_{MK} + \mathbf{B})\right)\mathbf{P}^{\dagger}, \tag{17}$$

if and only if $\mathbf{r}_1 \prec_{\times} \mathbf{d}$. The decomposition taking this form is referred to as the space-time GTD (ST-GTD). We denote

$$\mathbf{D}_{\mathbf{R}_1} = \mathrm{diag}(\mathbf{r}_1) \tag{18}$$

$$\mathbf{B}_1 = \mathbf{D}_{\mathbf{R}_1}^{-1}\mathbf{R}_1 - \mathbf{I}_K. \tag{19}$$

When the entries of $\mathbf{r}_1$ equal $\left(\prod_{k=0}^{K-1}\sigma_k\right)^{1/K}$, the time domain GTD in (15) reduces to the time domain GMD. We name this kind of ST-GTD, in particular, as ST-GMD.

The ST-GTD has some advantages over directly applying GTD on big matrix $\mathbf{H}$. Both algorithms first compute the SVD of $\mathbf{H}$, and do the decompositions on the diagonal matrix consisting of all singular values. The block diagonal structure of $\mathbf{H}$ helps to reduce the complexity in the SVD stage, from $O(K^3MNJ)$ to $O(KMNJ)$. Assuming that SVD of $\mathbf{H}$ is given, ST-GTD requires lower computational complexity, $O(KM(N+J)) + O(K^2)$, than the complexity of directly applying GTD on $\mathbf{H}$, which is $O(2K^2M^2)$ [18]. Moreover, ST-GTD decouples precoding into spatial and temporal domains. So, ST-GTD can be chosen in such a way that channel prediction is not necessary as we show later in Section IV.

B. Space-Time GTD Transceivers

In this subsection, we propose the ST-GTD ZF-DFE transceiver based on the ST-GTD introduced in the preceding subsection. The proposed precoder of ST-GTD transceiver is shown in Fig. 2. The proposed receiver is in Fig. 3 and its operation will be explained later. Here, it is assumed that the transmitter could predict the channels $\mathbf{H}(k)$ for $0 \leq k \leq K-1$ before sending a ST-block $\mathbf{a}$. We also assume that the receiver can track the channels perfectly and the decoding follows after the reception of a whole ST-block. There are well-studied methods [9], [19] which we can exploit here for the channel prediction. In any case, this system performance serves as a benchmark for performance comparisons and the theoretical foundation for the development of the transceiver which does not require channel prediction in next section.

Channel prediction is applicable when $\mathbf{H}(k)$ for different k are correlated [9], [19]. Before precoding a ST-block, one can apply a Wiener filter to predict $\mathbf{H}(1), \ldots, \mathbf{H}(K-1)$ based upon previous P channel matrices as

$$\hat{H}_{i,j}(k) = \mathbf{w}_{i,j,k}^{\dagger}\mathbf{h}_{i,j} \tag{20}$$

where $\hat{H}_{i,j}(k)$ is the (i,j)th element of the predicted channel matrix $\hat{\mathbf{H}}(k)$ of $\mathbf{H}(k)$ for $k = 1, \ldots, K-1$, $\mathbf{w}_{i,j,k} = [w_{i,j,k}(0), w_{i,j,k}(1), \ldots, w_{i,j,k}(P-1)]^T$ and $\mathbf{h}_{i,j} = [H_{i,j}(0), H_{i,j}(-1), \ldots, H_{i,j}(-P+1)]$. Suppose Jakes' model in (2) is used, then $\mathbf{w}_{i,j,k}$ is given by

$$\mathbf{w}_{i,j,k} = \mathbf{R}_{\mathbf{h}}^{-1}\mathbf{r}_{\mathbf{h}}(k) \tag{21}$$

where

$$[\mathbf{R}_{\mathbf{h}}]_{m,n} = J_0\left(2\pi f_d N_c T_s |n-m|\right)$$
$$[\mathbf{r}_{\mathbf{h}}(k)]_m = J_0\left(2\pi f_d N_c T_s |m+k|\right),$$

for $0 \leq m, n \leq P-1$. Interested readers can refer to [9] and [19] for more details.

If channel prediction is perfect, then the transmitter and receiver have perfect CSI of $\mathbf{H}$. ST-GTD can be applied to decompose $\mathbf{H}$ as (17) to get $\mathbf{P}$, $\mathbf{P}_1$, $\mathbf{Q}$, $\mathbf{D}_{\mathbf{R}}$, $\mathbf{B}$ and $\mathbf{B}_1$ for ST-GTD transceiver. Note that $\mathbf{D}_{\mathbf{R}}$ should be chosen as $\mathrm{diag}(\mathbf{d}) \otimes \mathbf{I}_M$ and $\mathbf{D}_{\mathbf{R}1} = \mathrm{diag}(\mathbf{r}_1)$ where $\mathbf{r}_1 \prec_{\times} \mathbf{d}$. If ST-GMD is applied instead of ST-GTD, we name the transceiver, in particular, the ST-GMD transceiver.

However, in practice, perfect channel prediction is not possible. So, the design of ST-GTD transceivers based on the predicted channel may not match the actual channel due to channel prediction error. To alleviate the mismatch, we modify the design procedure. Note that channel prediction is not required by the receiver to get $\mathbf{H}$ since the receiver can store the signal until the whole ST-block is received before it starts to decode. The only part that depends on channel prediction is the temporal precoder $\mathbf{P}_1$ at the transmitter. The computation of precoding matrix $\mathbf{P}_1$ requires the knowledge of the singular values of $\mathbf{H}(0), \ldots, \mathbf{H}(K-1)$. At time $k=0$, the precoder already needs $\mathbf{P}_1$ to precode one block for transmission. The implementation of the spatial precoder $\mathbf{P}$ without channel prediction is not a problem, since $\mathbf{P}$ is block diagonal matrix consisting of $\mathbf{P}(k)$ and the computation of $\mathbf{P}(k)$ requires only the current CSI $\mathbf{H}(k)$. Letting $\mathbf{a}' = (\mathbf{P}_1 \otimes \mathbf{I}_M)\mathbf{a}$, the precoded block at time k is given by

$$[\mathbf{t}]_{kM:kM+M-1} = \mathbf{P}(k)[\mathbf{a}']_{kM:kM+M-1}. \tag{22}$$

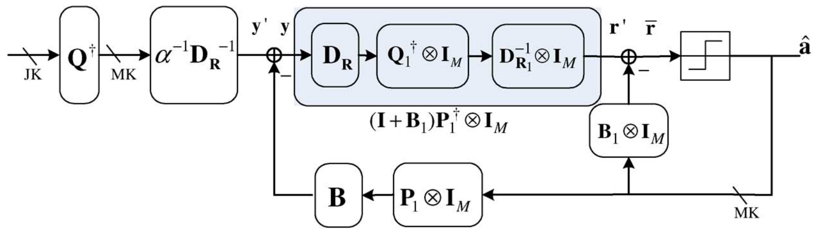


Fig. 3. The receiver of the ST-GTD transceiver.

To design $\mathbf{P}_1$, we firstly apply the Wiener prediction filter in (20) for channel prediction and construct the predicted channel matrix

$$\hat{\mathbf{H}} = \mathrm{diag}\left(\mathbf{H}(0), \hat{\mathbf{H}}(1), \ldots, \hat{\mathbf{H}}(K-1)\right) \tag{23}$$

where $\hat{\mathbf{H}}(k)$ are given by (20). Then, $\mathbf{P}_1$ is obtained from the ST-GTD or ST-GMD of $\hat{\mathbf{H}}$. In the case of perfect channel prediction, $\mathbf{Q}_1$, $\mathbf{P}_1$ and $\mathbf{R}_1$ were obtained from the temporal domain GTD of $\boldsymbol{\Sigma}$ which is obtained from $\mathbf{H}$, i.e., $\boldsymbol{\Sigma} = \mathbf{Q}_1\mathbf{R}_1\mathbf{P}_1^\dagger$. Here, $\mathbf{Q}_1$ and $\mathbf{R}_1$ are obtained from the QR decomposition of $\boldsymbol{\Sigma}\mathbf{P}_1$ as

$$\boldsymbol{\Sigma}\mathbf{P}_1 = \mathbf{Q}_1\mathbf{R}_1 \tag{24}$$

where $\mathbf{P}_1$ is from the ST-GTD of $\hat{\mathbf{H}}$, $\mathbf{Q}_1$ is a $K \times K$ unitary matrix and $\mathbf{R}_1$ is a $K \times K$ upper triangular matrix. Without loss of generality, we can make the vector $\mathbf{r}_1$ which consists of the diagonal entries of $\mathbf{R}_1$ positive and real. One can also verify that if the channel prediction is perfect both design procedures lead to the same ST-GTD or ST-GMD transceiver.

For each ST-block time, $\mathbf{a}$ is precoded by the linear precoders, transmitted through the channel $\mathbf{H}$ and preprocessed at the receiver by $\mathbf{Q}^\dagger$, and $\alpha^{-1}\mathbf{D}_{\mathbf{R}}^{-1}$. The estimation of $\mathbf{a}$ is obtained by the successive cancelation algorithm described next.

C. Successive Cancelation Detection Algorithm for ST-GTD Transceivers

Before summarizing the detection algorithm at the receiver in Fig. 3, we need some notations. For $0 \le i \le M-1$, define the $M \times M$ diagonal matrix $\mathbf{E}_i$ as

$$\mathbf{E}_i = \mathrm{diag}\left([\underbrace{0,\ldots,0}_{M-i}, \underbrace{1,\ldots,1}_{i}]\right) \tag{25}$$

and the $M \times 1$ vector $\mathbf{s}_i$ in which all the entries are zero except that the ith entry is 1. Based on $\mathbf{E}_i$ and $\mathbf{s}_i$, we define two operators, $\boldsymbol{\Theta}_i = \mathbf{I}_K \otimes \mathbf{E}_i$ and $\mathbf{S}_i = \mathbf{I}_K \otimes \mathbf{s}_i^T$, on the ST-block vector $\mathbf{a}$ which has the form (10). $\boldsymbol{\Theta}_i$ retains the last i symbols of each $\mathbf{a}(k)$ and makes the other symbols of $\mathbf{a}(k)$ zero. And $\mathbf{S}_i$ is such that

$$\mathbf{S}_i\mathbf{a} = [a_i(0), \ldots, a_i(K-1)]^T.$$

The detection algorithm for the receiver in Fig. 3 is as follows:

1) Initialize: $i = 0$.
2) *Outer loop feedback (space domain)*: Calculate

$$\mathbf{S}_{M-i-1}\mathbf{y} = \mathbf{S}_{M-i-1}\left(\mathbf{y}' - \mathbf{B}(\mathbf{P}_1 \otimes \mathbf{I}_M)\boldsymbol{\Theta}_i\hat{\mathbf{a}}\right).$$

3) *Inner loop feedback and detection (time domain)*: Compute

$$\mathbf{z}^{(i)} = \mathbf{S}_{M-i-1}\mathbf{r}' = (\mathbf{I}_K + \mathbf{B}_1)\mathbf{P}_1^\dagger\mathbf{S}_{M-i-1}\mathbf{y};\ \mathbf{S}_{M-i-1}\hat{\mathbf{a}}$$

can be decoded sequentially with the following procedures:
   a) $\hat{a}_{(K-1)M+M-1-i} = Qt(z_{K-1}^{(i)})$ where the function $q = Qt(t)$ sets q to the element in χ such that it is closest to t in Euclidean norm.
   b) For $c = 2, \ldots, K$, $\hat{a}_{(K-c)M+M-1-i} = Qt(z_{K-c}^{(i)} - \sum_{m=K-c+1}^{K-1} [\mathbf{B}_1]_{K-c,m}\hat{a}_{mM+M-i-1})$.
4) If $i = M-1$, then stop, else set $i = i+1$ and go to 2.

Step 1)-2) are clear by direct substitution. To justify Step 3), we assume that there is no error propagation, i.e., $\boldsymbol{\Theta}_i\hat{\mathbf{a}} = \boldsymbol{\Theta}_i\mathbf{a}$. By substitution, we have

$$\mathbf{y}' = (\mathbf{I}_{MK} + \mathbf{B})(\mathbf{P}_1 \otimes \mathbf{I}_M)\mathbf{a} + \mathbf{w}' \tag{26}$$

where

$$\mathbf{w}' = (1/\alpha)\mathbf{D}_{\mathbf{R}}^{-1}\mathbf{Q}^\dagger\mathbf{w}$$
$$\mathbf{w} = \left[\mathbf{w}(0)^T, \ldots, \mathbf{w}(K-1)^T\right]^T.$$

Then

$$\begin{aligned}\mathbf{S}_{M-i-1}\mathbf{y} &= \mathbf{S}_{M-i-1}\left(\mathbf{y}' - \mathbf{B}(\mathbf{P}_1 \otimes \mathbf{I}_M)\boldsymbol{\Theta}_i\hat{\mathbf{a}}\right)\\ &= \mathbf{P}_1\mathbf{S}_{M-i-1}\mathbf{a} + \mathbf{S}_{M-i-1}\mathbf{w}'.\end{aligned}$$

Substituting $\mathbf{S}_{M-i-1}\mathbf{y}$ into $\mathbf{S}_{M-i-1}\mathbf{r}'$, we have

$$\mathbf{S}_{M-i-1}\mathbf{r}' = (\mathbf{I}_K + \mathbf{B}_1)\mathbf{S}_{M-i-1}\mathbf{a} + \mathbf{w}_2 \tag{27}$$

where $\mathbf{w}_2 = \mathbf{S}_{M-i-1}((\mathbf{I}_K + \mathbf{B}_1)\mathbf{P}_1^\dagger \otimes \mathbf{I}_M)\mathbf{w}'$. Observe that the equivalent channel between $\mathbf{S}_{M-i-1}\mathbf{r}'$ and $\mathbf{S}_{M-i-1}\mathbf{a}$ is an upper triangular matrix. Hence, $\mathbf{S}_{M-i-1}\mathbf{a}$ can be detected sequentially by the VBLAST-like algorithm in Steps 3a) and 3b) above.

D. Mean-Square Error at the Detector

To analyze the performance of ST-GTD transceiver in Fig. 3, we assume perfect channel prediction. The performance of the

ST-GTD transceiver mainly depends on the noise component in $\bar{\mathbf{r}}$. To characterize the performance of the detector, we calculate the error covariance matrix of the noise component. We assume that there is no error propagation so that $\hat{\mathbf{a}} = \mathbf{a}$ which is a legitimate assumption at high SNR. Under this assumption, the signal vector $\bar{\mathbf{r}}$ can be expressed as

$$\bar{\mathbf{r}} = \mathbf{a} + \mathbf{e} \tag{28}$$

where

$$\mathbf{e} = \frac{1}{\alpha}\left(\mathbf{D}_{\mathbf{R}_1}^{-1}\mathbf{Q}_1^{\dagger} \otimes \mathbf{I}_M\right)\mathbf{q}.$$

The entries of the $MK \times 1$ vector $\mathbf{q} = \mathbf{Q}^{\dagger}\mathbf{w}$ are i.i.d. complex Gaussian with zero mean and variance σ_w^2. The error covariance matrix $\mathbf{R}_{\mathbf{ee}}$ is given by

$$\mathbf{R}_{\mathbf{ee}} = E(\mathbf{e}\mathbf{e}^{\dagger}) = \frac{\sigma_w^2}{\alpha^2}\left(\mathbf{D}_{\mathbf{R}_1}^{-1}\mathbf{D}_{\mathbf{R}_1}^{-\dagger} \otimes \mathbf{I}_M\right) \tag{29}$$

Denote the vector $\mathbf{r}_1$ which consists of the diagonal elements of $\mathbf{R}_1$ as $\mathbf{r}_1 = [\eta_0, \eta_1, \ldots, \eta_{K-1}]^T$. The total MSE of the ST-GTD transceiver over a ST-block is

$$\xi_{\mathrm{st-gtd}} = \frac{\sigma_w^2 M}{\alpha^2}\sum_{k=0}^{K-1}\frac{1}{|\eta_k|^2} \geq \frac{\sigma_w^2 MK}{\alpha^2}\left(\prod_{k=0}^{K-1}\frac{1}{\sigma_k^2}\right)^{1/K} \tag{30}$$

where $\eta_k = [\mathbf{r}_1]_k$. The last inequality comes from AM-GM inequality and $\mathbf{r_1} \prec_{\times} \mathbf{d}$. The equality holds when $|\eta_0| = |\eta_1| = \ldots = |\eta_{K-1}|$ and $\mathbf{r_1} \prec_{\times} \mathbf{d}$. In particular, if we choose

$$\eta_k = \left(\prod_{i=0}^{K-1}\sigma_i\right)^{1/K} \tag{31}$$

then $\mathbf{r}_1 \prec_{\times} \mathbf{d}$ is also satisfied, making the ST-GTD possible. This is the case when the ST-GMD is applied. We call this class of ST-GTD transceiver the ST-GMD transceiver. The total mean-square error of the ST-GMD transceiver is given by

$$\begin{aligned}\xi_{\mathrm{st-gmd}} &= \frac{\sigma_w^2 MK}{\alpha^2}\left(\prod_{k=0}^{K-1}\frac{1}{\sigma_k^2}\right)^{1/K} \\ &= \frac{\sigma_w^2 MK}{\alpha^2}\left(\prod_{k=0}^{K-1}\prod_{i=0}^{M-1}\frac{1}{\sigma_{H,i}^2(k)}\right)^{1/MK}.\end{aligned} \tag{32}$$

The class of ST-GMD transceivers is the optimal subclass of ST-GTD transceivers in terms of total mean-square error. Notice that the ST-GMD allows the ST-GTD transceiver to reach the optimal MSE in (32), which is the smallest achievable MSE by directly applying the GMD to the big matrix $\mathbf{H}$. Furthermore, the error covariance matrix of ST-GMD transceiver has equal diagonal elements. Hence, for every ST-block, the ST-GMD transceiver minimizes both the arithmetic and geometric MSE, and the average un-coded BER at the high SNR region according to [7].

IV. Space-Time GTD Transceivers With Fixed Temporal Precoder

In the previous section, the design of ST-GTD transceivers relies on the channel prediction. However, the channel prediction might not always be that accurate when the MIMO channels $\mathbf{H}(k)$ from block to block become more independent. The performance of the transceiver degrades when the predicted CSI at the transmitter is unreliable. In this section, we develop the ST-GTD transceiver which does not use channel prediction. We say that the transmitter is "causal."

As mentioned in Section III-B, the computation of precoding matrix $\mathbf{P}_1$ requires the knowledge of the singular values of $\mathbf{H}(0), \ldots, \mathbf{H}(K-1)$. Without channel prediction, the precoder only has the CSI at the current and previous times, and it is impossible to compute $\mathbf{P}_1$. To make the precoder causal, one can let $\mathbf{P}_1$ be a constant unitary matrix $\mathbf{W}$. In [21] and [20], the DFT or Hadamard matrix is chosen as the channel independent precoder for the OFDM system to equalize the MSEs over subchannels and hence minimize average BER. This motivates us to choose $\mathbf{W}$ to be a DFT or Hadamard matrix. $\mathbf{Q}_1$ and $\mathbf{R}_1$ are obtained from the QR decomposition of $\mathbf{\Sigma W}$ as

$$\mathbf{\Sigma W} = \mathbf{Q}_1\mathbf{R}_1 \tag{33}$$

where $\mathbf{Q}_1$ is also a $K \times K$ unitary matrix and $\mathbf{R}_1$ is a $K \times K$ upper triangular matrix. We call this kind of transceiver the causal ST-GTD transceiver (CST-GTD). It is in fact a subclass of ST-GTD with perfect channel prediction. The error covariance matrix of the noise signal in front of the detector is given by (29). And the total mean-square error $\xi_{\mathrm{cst-gtd}}$ is the same as (30).

A. Comparison of Mean-Square Error

Now, we compare the performance of the conventional GMD-based system [5], the ST-GMD transceiver with perfect channel prediction, and the CST-GTD transceiver. The total MSE of the GMD-based system in one ST-block is

$$\xi_{\mathrm{gmd}} = \frac{\sigma_w^2}{\alpha^2}M\sum_{k=0}^{K-1}\frac{1}{\sigma_k^2}. \tag{34}$$

The comparison of the three transceivers is given in Theorem 2.

Theorem 2: The total mean-square errors over one ST-block for the three transceivers are such that

$$\xi_{\mathrm{st-gmd}} \leq \xi_{\mathrm{st-gtd}} \leq \xi_{\mathrm{gmd}}. \tag{35}$$

Proof: The first inequality follows from (30). To prove the second inequality, we firstly define a function

$$f(\mathbf{x}) = \frac{\sigma_w^2 M}{\alpha^2}\sum_{k=0}^{K-1}e^{-2x_k} \tag{36}$$

where $f(\cdot) : \mathbb{R}^K \mapsto \mathbb{R}$ and $\mathbf{x} = [x_0, x_1, \ldots, x_{K-1}]^T$. Since $e^{-2x}, x \in \mathbb{R}$ is a convex function, $f(\mathbf{x})$ is a Schur-convex function by [23, Prop. 3.C.1]. Let

$$\begin{aligned}\mathbf{\Theta} &= [\log|\eta_0|, \log|\eta_1|, \ldots, \log|\eta_{K-1}|]^T \\ \mathbf{\Gamma} &= [\log|\sigma_0|, \log|\sigma_1|, \ldots, \log|\sigma_{K-1}|]^T.\end{aligned} \tag{37}$$

Since $\mathbf{r}_1 \prec_\times \mathbf{d}$, then $\mathbf{\Theta} \prec_+ \mathbf{\Gamma}$. $\xi_{\text{st-gtd}}$ in (30) and ξ_{gmd} in (34) can be expressed in term of $\mathbf{\Gamma}$ and $\mathbf{\Theta}$, respectively, as

$$\xi_{\text{st-gtd}} = f(\mathbf{\Theta}) \tag{38}$$
$$\xi_{\text{gmd}} = f(\mathbf{\Gamma}). \tag{39}$$

Since $\mathbf{\Theta} \prec_+ \mathbf{\Gamma}$, by the definition of Schur-convex function in [23], we have $\xi_{\text{st-gtd}} \leq \xi_{\text{gmd}}$. ■

Therefore, we have proven that the class of ST-GTD transceivers has performance superior to the conventional GMD-based system over time-varying channel in terms of total MSE within one ST-block or arithmetic MSE. In particular, for CST-GTD, we have

$$\xi_{\text{st-gmd}} \leq \xi_{\text{cst-gtd}} \leq \xi_{\text{gmd}}.$$

Also, note that ST-GMD transceiver with imperfect channel prediction can be treated as the ST-GTD transceiver with perfect channel prediction due to the channel mismatch caused by prediction error. Hence, the total mean-square error of ST-GMD transceiver with imperfect channel prediction $\xi_{\text{st-gmdic}}$ is such that

$$\xi_{\text{st-gmd}} \leq \xi_{\text{st-gmdic}} \leq \xi_{\text{gmd}}.$$

The temporal precoder $\mathbf{P}_1$ or $\mathbf{W}$, and the "nested-feedback-loop" receiver in ST-GTD or CST-GTD transceiver not only redistribute the MSEs of the blocks in each ST-block but also reduce the arithmetic MSE per ST-block. This is in contrast to the linear block precoder in [20] and [21], which keeps the same arithmetic MSE while equalizing the MSEs. Also, note that the conventional GMD-based system is actually a subclass of CST-GTD with the constant temporal precoding matrix $\mathbf{W}_K = \mathbf{I}_K$.

B. Comparison of Complexity

In this section, we compare the complexity of the conventional GMD-based system and the ST-GTD transceiver. We let these systems process one ST-block and compare the number of multiplications and additions. For the transmitter part, the complexity of the GMD-based system is $O(MNK)$ which comes from the spatial precoder $\mathbf{P}(k)$ in Fig. 1. Since the transmitters of ST-GTD incorporate an additional temporal precoder $\mathbf{P}_1$ as in Fig. 2, it has complexity $O(MNK + K^2M)$.

Next, we compare the receivers. For the GMD-based system in Fig. 1, the feedforward matrix $\mathbf{Q}^\dagger(k)$ has complexity $O(JM)$ and the feedback matrix $\mathbf{B}(k)$ has $O(M^2)$. The total complexity of the GMD-based receiver is $O(JMK+M^2K)$. The ST-GTD transceiver contains two additional temporal precoders in the feedback loop and an additional temporal feedback loop with feedback matrix $\mathbf{B}_1 \otimes \mathbf{I}_M$ which has complexity $O(MK^2)$. Hence, its total complexity is $O(JMK + MK^2 + M^2K)$.

V. Performance Analysis

In this section, we will compare the average BER and the ergodic channel capacity of the conventional GMD-based system, ST-GTD and ST-GMD transceivers. For both ST-GTD and ST-GMD transceivers, perfect channel prediction is assumed. We assume N' uses of the time-varying channels:

$$\mathbf{H}(n), \quad \text{for} \quad 0 \leq n \leq N' - 1. \tag{40}$$

Every K successive uses constitute one ST-block. So the mth ST-block uses the channels:

$$\mathbf{H}(mK + k), \quad \text{for} \quad 0 \leq k \leq K - 1 \tag{41}$$

where $0 \leq m \leq \lceil N'/K \rceil - 1$. The number of blocks, N', is assumed to be a large number and a multiple of the ST-block size K. Even number of bits, b, are allocated for every symbol $a_i(k)$ of each ST-block. For square QAM [22], the BER for each symbol in the kth block of the mth ST-block, assuming that there is just one bit error per symbol error, is approximately

$$P_e \approx cQ\left(\frac{A}{\sqrt{\nu(mK+k)}}\right) \tag{42}$$

where $Q(\cdot)$ is the Q-function defined in [22], $A = \sqrt{(3E_{av})/(2^b - 1)}$, E_{av} is the average symbol power, $\nu(mK + k)$ is the per symbol MSE of the kth block in the mth ST-block and $c = (4/b)(1 - 2^{-b/2})$. Notice that the symbol error rate (SER) equals bP_e. The average BER over the entire transmission is hence given by

$$\mathcal{P} = \frac{1}{N'} \sum_{m=0}^{(N'/K)-1} \sum_{k=0}^{K-1} cQ\left(\frac{A}{\sqrt{\nu(mK+k)}}\right). \tag{43}$$

The function $Q(A/\sqrt{y})$ for $y \in \mathbb{R}$ plays a crucial role in BER analysis. An important property of it is restated as the following lemma.

Lemma 1: The function $f(y) = Q(A/\sqrt{y})$ is monotone increasing. It is convex when $y \leq A^2/3$ and concave when $y > A^2/3$.

Proof: See [21]. ■

We define the SNR of the nth block as $\Gamma(n)$. The SNR expressions for the ST-GTD and ST-GMD transceivers are given respectively by

$$\begin{aligned}\Gamma_{\text{st-gtd}}(n) &= \frac{\alpha^2 E_{av} |\eta_n^2|}{\sigma_w^2} \\ \Gamma_{\text{st-gmd}}(n) &= \frac{\alpha^2 E_{av}}{\sigma_w^2 \left(\prod_{i=0}^{K-1} \frac{1}{\sigma^2(mK+i)}\right)^{1/K}}\end{aligned} \tag{44}$$

which follow from (29). We also define two SNR regions:

$$\mathcal{R}_{high} = \left\{\Gamma : \Gamma \geq \frac{2^b - 1}{3}\right\}, \quad \mathcal{R}_{low} = \left\{\Gamma : \Gamma < \frac{2^b - 1}{3}\right\}. \tag{45}$$

Before starting the analysis, we prove another useful lemma:

Lemma 2: If $\mathbf{A} \in \mathbb{R}_{+}^{n\times n}$, $\mathbf{B} \in \mathbb{R}_{+}^{m\times m}$ are doubly stochastic matrices, the $\mathbf{C} = \mathbf{A} \otimes \mathbf{B}$ is an $mn \times mn$ doubly stochastic matrix.

Proof: See the Appendix. ■

A. BER Performance Comparison of the Transceivers

Now, we will compare the BER of the entire class of ST-GTD transceivers including ST-GMD, ST-GMD with imperfect channel prediction and CST-GTD transceivers with the conventional GMD-based system. The following lemma is helpful for further analysis.

Lemma 3: The function $\Delta(y) = Q(c_2 \exp(y))$ is monotone decreasing where $c_2 = \alpha A/\sigma_w > 0$. It is convex when $c_2^2 e^{2y} \geq 1$ and concave when $c_2^2 e^{2y} < 1$.

Proof: The proof is similar to Lemma 1. ■

In the following theorems, "high SNR" means the SNRs $\Gamma(n)$ of the transceivers are such that $\Gamma(n) \in \mathcal{R}_{high}$ and "low SNR" means $\Gamma(n) \in \mathcal{R}_{low}$.

Theorem 3: Let $\mathcal{P}_{\rm st-gmd}$, $\mathcal{P}_{\rm st-gtd}$ and $\mathcal{P}_{\rm gmd}$ be the average BER of ST-GMD, ST-GTD and the conventional GMD-based transceivers, respectively. Then,

$$\begin{aligned} &\mathcal{P}_{\rm st-gmd} \leq \mathcal{P}_{\rm st-gtd} \leq \mathcal{P}_{\rm gmd}, \quad \text{at high SNR} \\ &\mathcal{P}_{\rm st-gmd} \geq \mathcal{P}_{\rm st-gtd} \geq \mathcal{P}_{\rm gmd}, \quad \text{at low SNR.} \end{aligned} \tag{46}$$

Proof: We first prove the second inequality for both high and low SNR. Let

$$g(\mathbf{z}) = \frac{1}{K}\sum_{k=0}^{K-1} Q(c_2 e^{z_k}) \tag{47}$$

where $g(\cdot) : \mathbb{R}^K \longmapsto \mathbb{R}$, $\mathbf{z} = [z_0, \ldots, z_{K-1}]^T$ and $c_2 = \alpha A/\sigma_w$. By Lemma 3 and [23], $g(\mathbf{z})$ is Schur-convex when $c_2^2 e^{2z_k} \geq 1$ and Schur-concave when $c_2^2 e^{2z_k} \leq 1$ for all $0 \leq k \leq K-1$. For $0 \leq m \leq (N'/K) - 1$, define $K \times 1$ vectors as

$$\mathbf{x}(m) = [\log(\sigma_{mK}), \ldots, \log(\sigma_{mK+K-1})]^T \tag{48}$$
$$\mathbf{y}(m) = [\log(|\eta_{mK}|), \ldots, \log(|\eta_{mK+K-1}|)]^T. \tag{49}$$

So, the BERs of the GMD and the ST-GTD transceivers are given, respectively, by

$$\mathcal{P}_{\rm gmd} = \frac{cK}{N'}\sum_{m=0}^{(N'/K)-1} g(\mathbf{x}(m)) \tag{50}$$
$$\mathcal{P}_{\rm st-gtd} = \frac{cK}{N'}\sum_{m=0}^{(N'/K)-1} g(\mathbf{y}(m)). \tag{51}$$

At the high SNR region, where $\Gamma_{\rm gmd}(n), \Gamma_{\rm st-gtd}(n) \in \mathcal{R}_{high}$, we have $c_2^2 e^{2x_k(m)} \geq 1$ and $c_2^2 e^{2y_k(m)} \geq 1$ for all k, m. In this domain, the function $g(\cdot)$ is Schur-convex. It is known that $\mathbf{r}_1 \prec_{\times} \mathbf{d}$, so $\mathbf{y}(m) \prec_{+} \mathbf{x}(m)$; Hence, we have $g(\mathbf{y}(m)) \leq g(\mathbf{x}(m))$ for $\forall m$. Therefore, $\mathcal{P}_{\rm st-gtd} \leq \mathcal{P}_{\rm gmd}$. At the low SNR region, where $\Gamma_{\rm gmd}(n)$, $\Gamma_{\rm st-gmd}(n) \in \mathcal{R}_{low}$, we can prove $\mathcal{P}_{\rm st-gtd} \geq \mathcal{P}_{\rm gmd}$ similarly. The first inequality can also been proven by following similar steps. ■

Let $\mathcal{P}_{\rm st-gmdic}$ and $\mathcal{P}_{\rm cst-gtd}$ denote the average BER of the ST-GMD transceiver with imperfect channel prediction, and CST-GTD, respectively. At the high SNR region, from Theorem 3, we can conclude, in particular, that

$$\mathcal{P}_{\rm st-gmd} \leq \mathcal{P}_{\rm cst-gtd} \leq \mathcal{P}_{\rm gmd} \tag{52}$$
$$\mathcal{P}_{\rm st-gmd} \leq \mathcal{P}_{\rm st-gmdic} \leq \mathcal{P}_{\rm gmd}. \tag{53}$$

B. Block Size and the BER Performance

In this subsection, the relationship between the size of ST-block and the BER performance is explored.

Theorem 4: Let $\mathcal{P}_{\rm st-gmd}^{(K)}$ denote the BER of the ST-GMD transceiver with ST-block size K and $\mathcal{P}_{\rm st-gmd}^{(qK)}$ denote that with ST-block size qK, for $q, K \in \mathbb{N}$. The number of blocks transmitted, N', is assumed to be the multiple of qK. Then,

$$\mathcal{P}_{\rm st-gmd}^{(qK)} \begin{cases} \leq \mathcal{P}_{\rm st-gmd}^{(K)}, & \text{at high SNRm} \\ \geq \mathcal{P}_{\rm st-gmd}^{(K)}, & \text{at low SNR.} \end{cases} \tag{54}$$

Proof: Let

$$\psi(\mathbf{w}) = \frac{c}{N'}\sum_{n=0}^{N'-1} Q\left(c_2 \exp(w_n)\right) \tag{55}$$

where $\mathbf{w} = [w_0, w_1, \ldots, w_{N'-1}]^T$ and $c_2 = \alpha A/\sigma_w$. $\psi(\mathbf{w})$ is Schur-convex if all $c_2^2 e^{2w_n} \geq 1$, Schur-concave if all $c_2^2 e^{2w_n} < 1$. Let $\mathbf{x} = [x_0, x_1, \ldots, x_{N'-1}]^T$ where

$$x_n = \log \sigma_n. \tag{56}$$

For the ST-block size K,

$$\mathcal{P}_{\rm st-gmd}^{(K)} = \psi(\mathbf{y}), \tag{57}$$

where $\mathbf{y} = [y_0, y_1, \ldots, y_{N'-1}]^T$ and

$$y_{mK+k} = \frac{1}{K}\sum_{i=0}^{K-1} \log(\sigma_{mK+i}), \tag{58}$$

for $0 \leq k \leq K-1$ and $0 \leq m \leq (N'/K) - 1$. From (56) and (58),

$$\mathbf{y} = (\mathbf{I}_{N'/K} \otimes \mathbf{L}_K)\mathbf{x} \tag{59}$$

where $\mathbf{L}_K$ is a $K \times K$ matrix with equal elements, $1/K$.

For the ST-block size qK,

$$\mathcal{P}_{\rm st-gmd}^{(qK)} = \psi(\mathbf{z}) \tag{60}$$

where $\mathbf{z} = [z_0, z_1, \ldots, z_{N'-1}]^T$ and

$$z_{m'qK+k'} = \frac{1}{qK}\sum_{i=0}^{qK-1} \log(\sigma_{m'qK+i}), \tag{61}$$

for $0 \leq k' \leq qK-1$ and $0 \leq m' \leq N'/(qK) - 1$. From (59), (61),

$$\begin{aligned} \mathbf{z} &= \left(\mathbf{I}_{N'/(qK)} \otimes \mathbf{L}_q \otimes \mathbf{L}_K\right)\mathbf{x} \\ &= \left[\left(\mathbf{I}_{N'/(qK)} \otimes \mathbf{L}_q\right) \otimes \mathbf{I}_K\right]\left[\mathbf{I}_{N'/K} \otimes \mathbf{L}_K\right]\mathbf{x} \\ &= \left(\mathbf{I}_{N'/(qK)} \otimes \mathbf{L}_q \otimes \mathbf{I}_K\right)\mathbf{y}. \end{aligned} \tag{62}$$

By Lemma 2, $\mathbf{I}_{N'/(qK)} \otimes \mathbf{L}_q \otimes \mathbf{I}_K$ is a doubly stochastic matrix; So, $\mathbf{z} \prec_+ \mathbf{y}$ [23].

At the high SNR region, where $\Gamma^{(K)}_{\text{st-gmd}}(n)$, $\Gamma^{(qK)}_{\text{st-gmd}}(n) \in \widehat{\mathcal{R}}_{high}$, we have $c_2^2 e^{2y_n} \geq 1$ and $c_2^2 e^{2z_n} \geq 1$ for all n. In this domain, the function $\psi(\cdot)$ is Schur-convex. Since $\mathbf{z} \prec_+ \mathbf{y}$, one can conclude that $\psi(\mathbf{z}) \leq \psi(\mathbf{y})$. At the low SNR region, where $\Gamma^{(K)}_{\text{st-gmd}}(n)$, $\Gamma^{(qK)}_{\text{st-gmd}}(n) \in \widehat{\mathcal{R}}_{low}$, we can prove $\psi(\mathbf{y}) \leq \psi(\mathbf{z})$ similarly. ∎

At the high SNR region, from Theorem 4, we can conclude that $\mathcal{P}^{(qK)}_{\text{st-gmd}}$ is a non-increasing function of q. As the ST-block size gets larger, the BER performance of ST-GMD improves monotonically. Larger ST-block size is more favorable because it gains more diversity from the time-varying channels. However, it implies longer decoding delay at the receiver. At the low SNR region, the relationship is the other way around, so it is better to have small ST-block size.

C. Performance Comparison in Capacity

In the conventional GMD-based system, the average BER per ST-block is dominated by the block with the largest MSE. To achieve the optimal per ST-block average BER and hence minimize the average BER, bit allocation is required. The proposed ST-GMD transceiver does not require bit allocation among blocks since all SNRs of different blocks in each ST-block are the same. In this subsection, from the perspective of capacity, we will show the asymptotic optimality of ST-GMD transceiver.

With uniform power loading, the ergodic channel capacity [25] for the equivalent channel $\mathbf{H}$ in (11) of a ST-block is given by

$$\begin{aligned} C_{\text{upl}} &= E_{\mathbf{H}}\left\{\log\left(\det(\mathbf{I}+\rho\mathbf{H}^H\mathbf{H})\right)\right\} \\ &= E_{\mathbf{H}}\left\{\sum_{k=0}^{K-1}\sum_{i=0}^{M-1}\log\left(1+\rho\sigma_{H,i}^2(k)\right)\right\} \end{aligned} \quad (63)$$

where $\rho = (\sigma_a \alpha)^2/\sigma_w^2$ and $\sigma_{H,i}(k)$ is given in (4). In the conventional GMD-based system and the ST-GMD transceiver, the channel $\mathbf{H}$ is converted into equivalent parallel subchannels. Hence, the ergodic channel capacities of the equivalent subchannels obtained by using the conventional GMD-based system and the ST-GMD transceiver are respectively given by

$$\begin{aligned} C_{\text{gmd}} &= ME_{\mathbf{H}}\left\{\sum_{k=0}^{K-1}\log\left(1+\rho\sigma_k^2\right)\right\} \\ C_{\text{st-gmd}} &= MKE_{\mathbf{H}}\left\{\log\left(1+\rho\left(\prod_{k=0}^{K-1}\sigma_k^2\right)^{1/K}\right)\right\}. \end{aligned} \quad (64)$$

where σ_k is given by (4). For high SNR,

$$\lim_{\rho\to\infty} C_{\text{upl}} - C_{\text{gmd}} = \sum_{k=0}^{K-1} E_{\mathbf{H}}\left\{\log\left(\frac{\prod_i \rho\sigma_{H,i}^2(k)}{\prod_i \rho\sigma_{H,i}^2(k)}\right)\right\} = 0. \quad (65)$$

So the GMD-based system does not have capacity loss. For ST-GMD transceivers, $\lim_{\rho\to\infty} C_{\text{gmd}} - C_{\text{st-gmd}} = ME_{\mathbf{H}}\{\log(\prod_k \rho\sigma_k^2/\prod_k \rho\sigma_k^2)\} = 0$ Together with (65), we have

$$\lim_{\rho\to\infty} C_{\text{upl}} - C_{\text{st-gmd}} = 0. \quad (66)$$

Therefore, for high SNR, the ST-GMD transceiver is asymptotically optimal in capacity and per ST-block average BER simultaneously. Note that the design of ST-GMD transceiver is possible only when the channel prediction is good.

VI. Numerical Results

In this section, we present the numerical results on the average BERs of the GMD-based system, ST-GMD and CST-GTD transceivers. We also demonstrate how the ST-block size affects the BER performance. The channel model described in Section II-A is adopted. The noise is AWGN. The MIMO channel matrices $\mathbf{H}(k)$ are 3×3 complex Gaussian random matrices. The elements of $\mathbf{H}(k)$ are i.i.d. complex Gaussian random variables with zero mean and unit variance. Uniform bit allocation is adopted with $b = 4$ for each subchannel. The modulation scheme is 16-QAM. Both transmitter and receiver have perfect CSI at current time k. We assume that perfect channel prediction is available only for the ST-GMD transceiver. The temporal precoding matrix of the CST-GTD is a $K \times K$ DFT matrices. $N' = 2^{20}$ data blocks are sent through the channels for BER performance evaluation.

Example 1: The ST-block size is $K = 16$. $\mathbf{H}(k)$ are independent for different k. Fig. 4 shows the BER performance of the conventional GMD-based system, CST-GTD and ST-GMD transceivers for different SNRs. For the high SNR region, Fig. 4 satisfies $\mathcal{P}_{\text{st-gmd}} < \mathcal{P}_{\text{cst-gtd}} < \mathcal{P}_{\text{gmd}}$, which verifies Theorem 3. At BER 10^{-4}, the SNR gains of the ST-GMD and the CST-GTD over the GMD-based system are 5 dB and 4.7 dB, respectively. The performance of CST-GTD is close to ST-GMD. At BER 10^{-5}, the SNR gain of the ST-GMD transceiver over the CST-GTD transceiver is about 0.6 dB. At the low SNR region, $\mathcal{P}_{\text{st-gmd}}$ and $\mathcal{P}_{\text{cst-gtd}}$ are greater than $\mathcal{P}_{\text{gmd}}$. This is because of the error propagation. For the space-time processing at these receivers, the errors might propagate through the entire ST-block, i.e., K blocks.

Example 2: In this example, various choices of ST-block size are compared. $\mathbf{H}(k)$ are independent for different k. Fig. 5 shows $\mathcal{P}_{\text{st-gmd}}$ and $\mathcal{P}_{\text{cst-gtd}}$. At the low SNR region, $\mathcal{P}_{\text{st-gmd}}$ and $\mathcal{P}_{\text{cst-gtd}}$ increase with respect to K as in shown Fig. 7. But at the high SNR region, $\mathcal{P}_{\text{st-gmd}}$ and $\mathcal{P}_{\text{cst-gtd}}$ decrease with respect to K which is best illustrated by Fig. 6. These results verify Theorem 4. Notice that the CST-GTD transceiver almost has the same performance as the ST-GMD transceiver when $K = 32$. The SNR gap is only 0.08 dB at BER 10^{-5}.

Example 3: Here, the BER performances of the three transceivers are evaluated using Jakes' channel model. $\mathbf{H}(k)$ for different k are correlated and the cross-correlation is given by (2). Fig. 8 shows the average BER performances of the conventional GMD-based system, CST-GTD and ST-GMD transceivers for different values of the product $f_d N_c T_s$, which appears in (2). As $f_d N_c T_s$ gets larger, the channels are changing at faster rates, and $\mathbf{H}(k)$ for different k become more uncorrelated. The ST-block

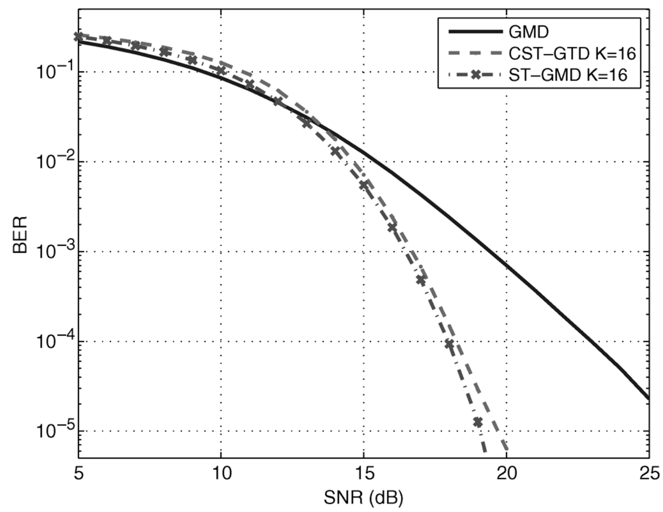


Fig. 4. BER performance of GMD, ST-GMD and CST-GTD.

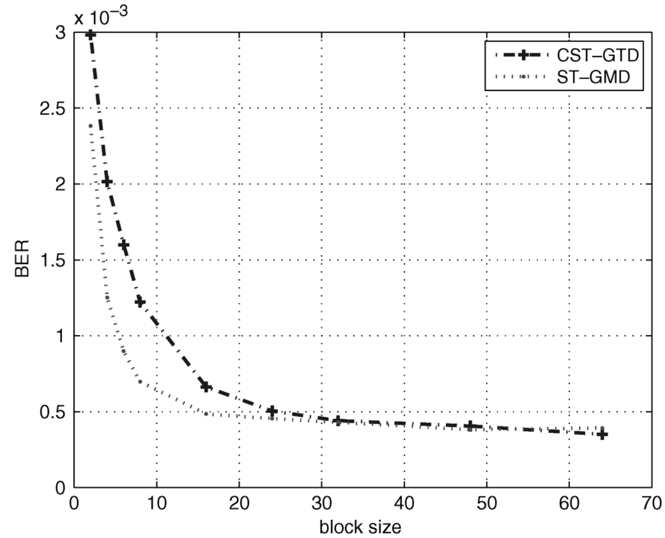


Fig. 6. BER performance versus block size K at SNR $=$ 17 dB.

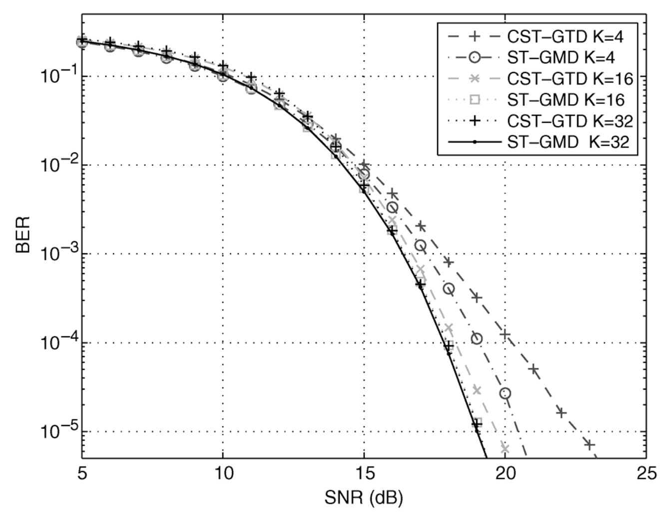


Fig. 5. BER performance versus block size for ST-GMD and CST-GTD.

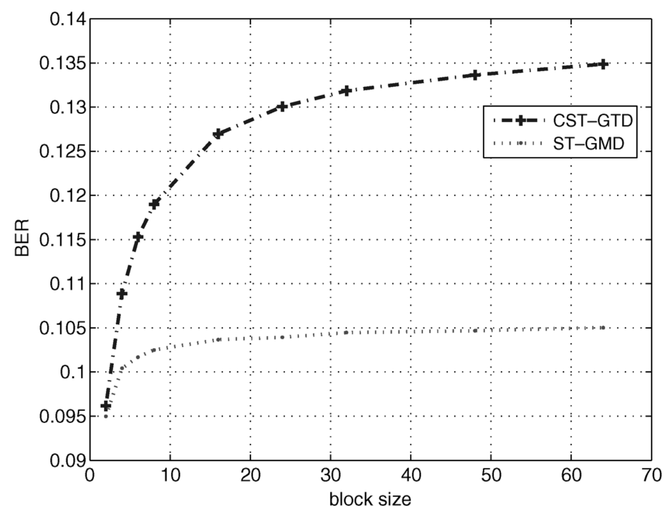


Fig. 7. BER performance versus block size K at SNR $=$ 10 dB.

size is $K = 16$ and the BERs are evaluated at SNR $=$ 17 dB. For small $f_d N_c T_s$, the channels are almost like time invariant channels, the BER improvements of CST-GTD and ST-GMD transceivers over the conventional GMD-based system are small since there is not much temporal diversity for the temporal precoders to exploit. As $f_d N_c T_s$ increases, the average BERs of CST-GTD and ST-GMD transceivers drop quickly due to the rich temporal diversity offered by the time-varying channels.

Example 4: This example demonstrates the BER performance of ST-GMD transceiver with imperfect channel prediction. $\mathbf{H}(k)$ follows the Jakes' model in Section II-A. The ST-block size $K = 4$, and $f_d N_c T_s = 0.1$. The ST-GMD transceiver is designed according to the procedure for the case of imperfect channel prediction in Section III-A. Fig. 9 illustrates the BER performance of the ST-GMD transceiver based on channel prediction. Its BER curve follows the curve of the ST-GMD transceiver with perfect channel prediction closely for most of the SNR values, and deviates at very high SNR region. The BER degradation results from the channel prediction error. At BER 10^{-5}, the SNR loss from imperfect channel prediction is 0.35 dB.

VII. Conclusion

We have proposed two MIMO transceivers with zero-forcing decision feedback structure for the MIMO slowly time-varying channels. They harvest the rich temporal diversity due to the time-varying nature of the channels to minimize the average BER. The issue of available CSIT for slowly time-varying channel can be addressed using feedback mechanisms or TDD schemes. Under the assumption of perfect channel prediction, the ST-GMD transceiver is shown to be the best in terms of arithmetic MSE and average BER in high SNR. The ST-GMD transceiver serves as a benchmark for performance. The CST-GTD transceivers only requires the instantaneous CSIT and CSIR as the GMD-based systems does. It has the same asymptotic BER performance as the ST-GMD transceiver and has smaller arithmetic MSE than the conventional GMD-based systems. The dependency of BER on the ST-block size has also

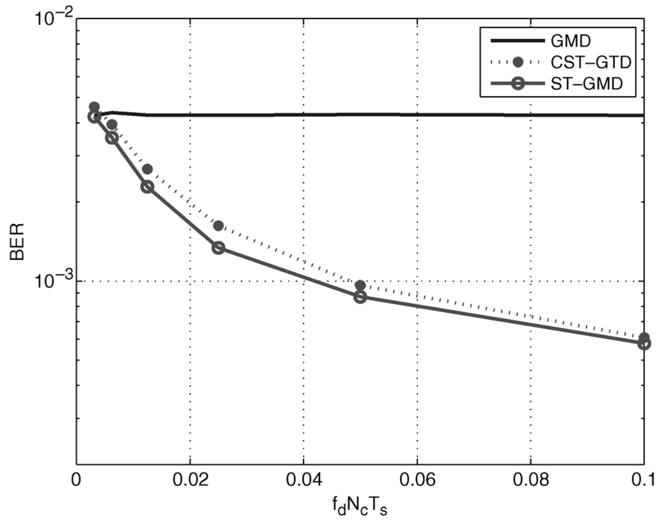


Fig. 8. BER performance versus $f_dN_cT_s$ at SNR $= 17$ dB.

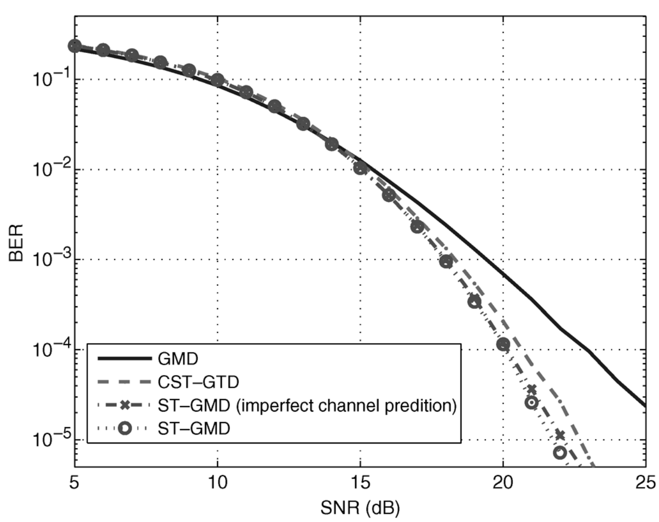


Fig. 9. BER performance of ST-GMD with channel prediction.

been analyzed. Simulations show that only moderate ST-block size is required for good average BER performance.

Appendix
Proof of Lemma 2

Let $[\mathbf{A}]_{ij} = a_{ij}$, $[\mathbf{B}]_{ij} = b_{ij}$ and $[\mathbf{C}]_{ij} = c_{ij}$. By the definition of Kronecker product,

$$c_{i_1 m+j_1, i_2 m+j_2} = a_{i_1,i_2} b_{j_1,j_2} \geq 0$$

where $0 \leq i_1, i_2 \leq n-1$ and $0 \leq j_1, j_2 \leq m-1$. Since $\mathbf{A}$ and $\mathbf{B}$ are doubly stochastic,

$$\sum_{i=0}^{n-1} a_{ij} = \sum_{j=0}^{n-1} a_{ij} = \sum_{i=0}^{m-1} b_{ij} = \sum_{j=0}^{m-1} b_{ij} = 1. \qquad (67)$$

Consider

$$\sum_{i=0}^{mn-1} c_{i,j} = \sum_{i_1=0}^{n-1}\sum_{j_1=0}^{m-1} c_{i_1 m+j_1, i_2 m+j_2} = \sum_{i_1}\sum_{j_1} a_{i_1,i_2} b_{j_1,j_2}$$

where $j = i_2 m + j_2$. By (67), we have

$$\sum_{i=0}^{mn-1} c_{i,j} = \sum_{i_1} a_{i_1,i_2} \sum_{j_1} b_{j_1,j_2} = 1. \qquad (68)$$

Similarly, we can prove $\sum_{j=0}^{mn} c_{i,j} = 1$. Therefore, $\mathbf{C}$ is also a doubly stochastic matrix.

References

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," AT&T Bell Labs., Murray Hill, NJ, Tech. Rep., 1995.

[2] G. J. Foschinini, "Layered space-time architecture for wireless communications in a fading environment when using multi-element antennas," *Bell Labs. Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[3] J. B. Anderson, "Array gain and capacity for known random channels with multiple elements arrays at both ends," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp. 2172–2178, Nov. 2000.

[4] J. Zhang, A. Kavcic, and K. M. Wong, "Equal-diagonal QR decomposition and its application to precoder design for successive-cancellation detection," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 154–172, Jan. 2005.

[5] Y. Jiang, J. Li, and W. W. Hager, "Joint transceiver design for MIMO communications using geometric mean decomposition," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 3791–3803, Oct. 2005.

[6] Y. Jiang, J. Li, and W. W. Hager, "Uniform channel decomposition for MIMO communications," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4283–4294, Nov. 2005.

[7] F. Xu, T. N. Davidson, J. K. Zhang, and K. M. Wong, "Design of block transceivers with decision feedback detection," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 965–978, Mar. 2006.

[8] D. P. Palomar and Y. Jiang, "MIMO transceiver design via majorization theory," *Found. Trends Commun. Inf. Theory*, vol. 3, no. 4, pp. 331–551, Nov. 2006.

[9] S. Zhou and G. B. Giannakis, "How accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over Rayleigh MIMO channels?," *IEEE Trans. Wireless Commun.*, vol. 3, no. 3, pp. 1285–1294, Jul. 2004.

[10] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[11] F. Dietrich, R. Hunger, M. Joham, and W. Utschick, "Linear precoding over time-varying channels in TDD systems," in *Proc. ICASSP*, Hong Kong, Apr. 2003, vol. 5, pp. 117–120.

[12] S. K. Jayaweera and H. V. Poor, "Capacity of multiple-antenna systems with both receiver and transmitter channel state information," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2697–2709, Oct. 2003.

[13] G. Lebrun, J. Gao, and M. Faulkner, "MIMO transmission over a time-varying channel using SVD," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 757–764, Mar. 2005.

[14] D. Samardzija and N. Mandayam, "Impact of pilot design on achievable data rates in multiple antenna multiuser TDD systems," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1370–1379, Sep. 2007.

[15] P. Komulainen, A. Tolli, M. Latva-aho, and M. Juntti, "Downlink assisted uplink zero forcing for TDD multiuser MIMO systems," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, no. 8, Mar. 2009 [Online]. Available: http://downloads.hindawi.com/journals/wcn/2009/894726.pdf

[16] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 357–366, Mar. 1998.

[17] C. H. Liu and P. P. Vaidyanathan, "ZF-DFE transceiver for time-varying MIMO channels with channel-independent temporal precoder," in *Proc. ISCAS*, Paris, France, May 2010, pp. 3525–3528.

[18] Y. Jiang, W. W. Hager, and J. Li, "Generalized triangular decomposition," *Math. Comput.*, vol. 77, no. 262, pp. 1037–2056, Apr.. 2008.

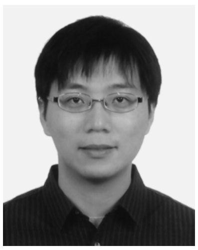
[19] A. Duel-Hallen, S. Hu, and H. Hallen, "Long-range prediction of fading signals," *IEEE Signal Precess. Mag.*, vol. 17, pp. 62–75, May 2000.

[20] Y. Ding, T. Davidson, Z. Luo, and K. Wong, "Minimum BER block precoders for zero-forcing equalization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2410–2423, Sep. 2003.

[21] Y. P. Lin and S. M. Phoong, "BER minimized OFDM systems with channel independent precoders," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2369–2380, Sep. 2003.

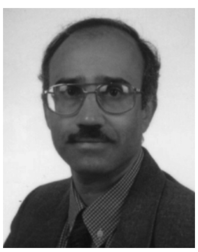
[22] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 2001.

[23] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York: Academic, 1979.
[24] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1991.
[25] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.



Chih-Hao Liu (S'10) was born in Chiayi, Taiwan, on December 5, 1980. He received the B.S. degree in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2003 and M.S. degrees in communication engineering from National Taiwan University, Taipei, Taiwan, in 2005.

He is currently working towards the Ph.D. degree in the field of digital signal processing at Caltech. His research interests include digital communications and signal processing.



Palghat P. Vaidyanathan (S'80–M'83–SM'88–F'91) was born in Calcutta, India, on October 16, 1954. He received the B.Sc. (Hons.) degree in physics and the B.Tech. and M.Tech. degrees in radiophysics and electronics, all from the University of Calcutta, India, in 1974, 1977, and 1979, respectively, and the Ph.D. degree in electrical and computer engineering from the University of California at Santa Barbara in 1982.

He was a Postdoctoral Fellow at the University of California, Santa Barbara, from September 1982 to March 1983. In March 1983, he joined the Electrical Engineering Department of the California Institute of Technology as an Assistant Professor, and since 1993, he has been Professor of electrical engineering there. His main research interests are in digital signal processing, multirate systems, wavelet transforms, signal processing for digital communications, genomic signal processing, radar signal processing, and sparse array signal processing. He has authored more than 400 papers in journals and conferences and is the author of the three books *Multirate Systems and Filter Banks* (Prentice-Hall, 1993), *Linear Prediction Theory* (Morgan and Claypool, 2008), and (with Phoong and Lin) *Signal Processing and Optimization for Transceiver Systems* (Cambridge Univ. Press, 2010). He has written several chapters for various signal processing handbooks.

Dr. Vaidyanathan served as Vice-Chairman of the Technical Program committee for the 1983 IEEE International symposium on Circuits and Systems, and as the Technical Program Chairman for the 1992 IEEE International symposium on Circuits and Systems. He was an Associate editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS from 1985 to 1987, and for the journal IEEE SIGNAL PROCESSING LETTERS, and a consulting editor for the journal *Applied and Computational Harmonic Analysis*. He was a Guest Editor in 1998 for special issues of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, on the topics of filter banks, wavelets, and subband coders. He was a recipient of the Award for Excellence in Teaching at the California Institute of Technology for the academic years 1983–1984, 1992–1993, and 1993–1994. He also received the NSF's Presidential Young Investigator Award in 1986. In 1989, he received the IEEE ASSP Senior Award for his paper on multirate perfect-reconstruction filter banks. In 1990, he was recipient of the S. K. Mitra Memorial Award from the Institute of Electronics and Telecommunications Engineers, India, for his joint paper in the IETE journal. In 2009, he was chosen to receive the IETE students' journal award for his tutorial paper in the *IETE Journal of Education*. He was also the coauthor of a paper on linear-phase perfect reconstruction filter banks in the IEEE TRANSACTIONS ON SIGNAL PROCESSING, for which the first author (T. Nguyen) received the Young Outstanding Author Award in 1993. He received the 1995 F. E. Terman Award of the American Society for Engineering Education, sponsored by the Hewlett Packard Company, for his contributions to engineering education. He has given several plenary talks, including at the IEEE ISCAS 2004, Sampta 2001, EUSIPCO 1998, SPCOM 1995, and Asilomar 1988 conferences on signal processing. He has been chosen a Distinguished Lecturer for the IEEE Signal Processing Society for the academic year 1996–1997. In 1999, he was chosen to receive the IEEE CAS Society's Golden Jubilee Medal. He is a recipient of the IEEE Signal Processing Society's Technical Achievement Award for 2002.