Uniform Decomposition of Mutual Information using MMSE Decision Feedback Detection

Jian-Kang Zhang, Timothy N. Davidson and Kon Max Wong
Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada.
Email: {jkzhang, davidson, wong}@mail.ece.mcmaster.ca

***Abstract*— We consider efficient techniques for the design of a transceiver for a matrix channel with minimum mean square error (MMSE) decision-feedback (DF) detection when perfect channel information is available at both the transmitter and receiver. By combining the canonical property of the MMSE-DF detector and our recently developed equal-diagonal QRS decomposition of a matrix we obtain a uniform decomposition of mutual information in which each of the synthesized scalar subchannels has the same mutual information (under the assumption of error-free feedback). To assist our analysis of this uniform decomposition, we provide a new QR interpretation of the MMSE-DF receiver. This enables us to show that the natural detection order is optimal (in an SINR sense), and that for the proposed transmitter, the MMSE-DF detector is asymptotically equivalent to the maximum likelihood detector when the SNR is high. We also derive a low-complexity quadratic recursive algorithm for the characterization of all eligible S-factors in the QRS decomposition. When coupled with our QR interpretation of the MMSE-DF detector, this enables us to efficiently design the optimal transmitter and to efficiently implement the MMSE-DF receiver.**

I. Introduction

In many block-by-block communication systems, the discrete-time baseband model for the received signal is

$$\mathbf{y} = \sqrt{\mathrm{snr}}\,\mathbf{H}\mathbf{T}\mathbf{x} + \boldsymbol{\xi}, \tag{1}$$

where $\mathbf{H}$ is an $M \times N$ complex matrix that models the channel, $\mathbf{T}$ is an $N \times K$ linear precoding matrix ($N \geq K$), $\mathbf{x}$ is the block of K transmitted symbols, which is assumed to be zero-mean, white and of unit variance, and $\boldsymbol{\xi}$ is an $M \times 1$ white Gaussian noise vector with unit variance. It is well known that if the channel matrix $\mathbf{H}$ is known at both the transmitter and the receiver, then the Gaussian mutual information between the input and the output of the model in (1) is $I_G(\mathbf{x};\mathbf{y}) = \log\det(\mathbf{I} + \mathrm{snr}\mathbf{T}^H\mathbf{H}^H\mathbf{H}\mathbf{T})$. For a given transmitted power constraint $\mathrm{tr}(\mathbf{T}^H\mathbf{T}) \leq p$, $I_G(\mathbf{x};\mathbf{y})$ is maximized when the transmitter $\mathbf{T}$ is the water-filling solution [1]; i.e., $\mathbf{T} = \mathbf{V}\boldsymbol{\Xi}\mathbf{W}$, where $\mathbf{V}$ is the eigen-vector matrix of $\mathbf{H}^H\mathbf{H}$, $\boldsymbol{\Xi}$ denotes the optimal power loading matrix and $\mathbf{W}$ is an arbitrary unitary matrix. Since the rotation of the white Gaussian vector $\mathbf{x}$ does not change the channel capacity, we have a unitary matrix degree of freedom, $\mathbf{W}$, within the water-filling solution family which can be designed so as to improve other aspects of system performance. We will propose an efficient design technique for $\mathbf{W}$ for applications in which an MMSE-DF receiver is employed.

A previous joint design technique for the transmitter and the DF receiver shows that the receiver that minimizes the (arithmetic) MSE between the input symbols and their estimates also achieves the minimum geometric MSE [2]. To design the transmitter, Yang and Roy [2] chose to optimize the geometric MSE rather than the (arithmetic) MSE. It turned out that their transmitter is same as the capacity-achieving input spectrum for the coordinated multiuser channel obtained by Brandenburg and Wyner [1]. However, this optimal transmitter is not guaranteed to simultaneously minimize the (arithmetic) MSE. In this paper we describe a new quadratic recursive algorithm for efficiently computing (and implementing) a matrix $\mathbf{W}$ that simultaneously minimizes the arithmetic and geometric MSEs and also maximizes the Gaussian mutual information. This algorithm advances our recent development of an equal-diagonal QRS decomposition of a matrix [3], which was originally developed for jointly designing a transceiver with zero-forcing (ZF) DF detector [3].

Notation: Matrices are denoted by uppercase boldface characters (e.g., $\mathbf{A}$), while column vectors are denoted by lowercase boldface characters (e.g., $\mathbf{b}$). The (i,j)-th entry of $\mathbf{A}$ is denoted by $A_{i,j}$ or $[\mathbf{A}]_{i,j}$. The i-th entry of $\mathbf{b}$ is denoted by b_i. The columns of an $M \times N$ matrix $\mathbf{A}$ are denoted by $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_N$. The term $\mathbf{A}_k$ denotes a matrix consisting of the first k columns of $\mathbf{A}$, i.e., $\mathbf{A}_k = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_k]$, and by convention $\mathbf{A}_0 = 1$. The matrix remaining after deletion of columns $k_1, k_2, \ldots, k_i$ from $\mathbf{A}$ is denoted by $\overline{\mathbf{A}}_{k_1,k_2\cdots,k_i}$. For a matrix $\mathbf{A}$, $\mathbf{A}^{\perp}$ denotes the orthonormal complement, $\mathbf{A}^T$ the transpose, and $\mathbf{A}^H$ the conjugate transpose.

II. QR interpretation of MMSE-DF detection

Consider a communication system of the form in (1) in which $\mathbf{T}$ is absorbed into $\mathbf{H}$; i.e., $\mathbf{y} = \sqrt{\mathrm{snr}}\,\mathbf{H}\mathbf{x} + \boldsymbol{\xi}$. At the receiver, our task is to detect (estimate) the vector $\mathbf{x}$ given the noisy observation $\mathbf{y} = [y_1, \cdots, y_M]^T$. For notational convenience, we define $\mathcal{I}_{\mathbf{H}} = \mathbf{I} + \mathrm{snr}\mathbf{H}^H\mathbf{H}$, and we will call $\mathcal{I}_{\mathbf{H}}^{1/2}$ the *mutual information matrix*. The DF receiver (with perfect feedback) makes successive decisions on the vector $\mathbf{z} = \mathbf{F}\mathbf{y} - \mathbf{B}\mathbf{x}$, [4]–[6], [8]–[10], where $\mathbf{F}$ and $\mathbf{B}$ are the feedforward and feedback matrices, respectively. For the MMSE-DF receiver, $\mathbf{F} = \sqrt{\mathrm{snr}}\,(\mathbf{B}+\mathbf{I})\mathcal{I}_{\mathbf{H}}^{-1}\mathbf{H}^H$ and $\mathbf{B}+\mathbf{I}$ is the upper triangular matrix with unit diagonal entries in the Cholesky decomposition $\mathcal{I}_{\mathbf{H}} = (\mathbf{B}+\mathbf{I})^H\mathbf{D}(\mathbf{B}+\mathbf{I})$, where $\mathbf{D}$ is a positive-definite diagonal matrix. Using this decomposition, we observe that the feedforward matrix $\mathbf{F}$ consists of three

parts. Specifically, $\mathbf{F} = \sqrt{\mathrm{snr}}\,\mathbf{D}^{-1}(\mathbf{B}+\mathbf{I})^{-H}\mathbf{H}^H$, where $\sqrt{\mathrm{snr}}\,\mathbf{H}^H$ is the channel matched filter, $(\mathbf{B}+\mathbf{I})^{-H}$ decorrelates the error covariance matrix, and $\mathbf{D}^{-1}$ is a diagonal scaling. MMSE decision feedback detection can be also interpreted using the QR decomposition of a certain virtual channel matrix [11], [12]. However, that virtual channel matrix has twice the number of rows as $\mathbf{H}$. In order to place our recently developed equal-diagonal QRS decomposition of a matrix in a similar context, we now provide a new QR interpretation of the MMSE-DF receiver that uses the QR decomposition of $\mathcal{I}_{\mathbf{H}}^{1/2}$; i.e., $\mathcal{I}_{\mathbf{H}}^{1/2} = \mathbf{QR}$, where $\mathbf{Q}$ is an $N \times N$ orthonormal matrix and $\mathbf{R}$ is an $N \times N$ upper triangular matrix with $R_{k,k} > 0$ for all $k = 1, 2, \cdots, N$. If we observe that $\mathbf{H}^H\mathbf{H} = \mathrm{snr}^{-1}(\mathcal{I}_{\mathbf{H}} - \mathbf{I}) = \mathrm{snr}^{-1}(\mathbf{R}^H\mathbf{R} - \mathbf{I})$, then we can write $\mathbf{R}^{-H}\mathbf{H}^H\mathbf{y} = \mathrm{snr}^{-1/2}(\mathbf{R} - \mathbf{R}^{-H})\mathbf{x} + \mathbf{R}^{-H}\mathbf{H}^H\boldsymbol{\xi}$. Therefore, after having been processed by MMSE-DF detector, the original model (1) with $\mathbf{T}$ absorbed into $\mathbf{H}$ is transformed into the following model,

$$\breve{\mathbf{y}} = \breve{\mathbf{R}}\mathbf{x} + \breve{\boldsymbol{\xi}}, \tag{2}$$

where $\breve{\mathbf{y}} = \mathbf{R}^{-H}\mathbf{H}^H\mathbf{y}$, $\breve{\mathbf{R}} = \mathrm{snr}^{-1/2}(\mathbf{R} - \mathrm{diag}(\mathbf{R}^{-H}))$, $\breve{\boldsymbol{\xi}} = \mathrm{snr}^{-1/2}\left(\mathrm{diag}(\mathbf{R}^{-H}) - \mathbf{R}^{-H}\right)\mathbf{x} + \mathbf{R}^{-H}\mathbf{H}^H\boldsymbol{\xi}$. When combined with the fast QRS decomposition described below, (2) exposes an efficient algorithm for implementing the MMSE-DF receiver.

III. Fast closed-form equal diagonal QRS decomposition of matrix

In this section, we first state our equal diagonal QRS decomposition theorem [3] and then give a simplified version of the quadratic recursive algorithm to characterize all the S-factors in this decomposition. In particular, we derive specific closed-form recursions for the S, Q and R-factors for a nonsingular matrix and its inverse.

A. Equal diagonal QRS decomposition

Let the singular value decomposition (SVD) of an $M \times N$ matrix $\mathbf{H}$ be represented by

$$\mathbf{H} = \mathbf{U}\begin{pmatrix} \boldsymbol{\Lambda}^{1/2} & \mathbf{0}_{r\times(N-r)} \\ \mathbf{0}_{(M-r)\times r} & \mathbf{0}_{(M-r)\times(N-r)} \end{pmatrix}\mathbf{V}^H, \tag{3}$$

where r is the rank of $\mathbf{A}$, $\mathbf{U}$ and $\mathbf{V}$ are unitary and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_r)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$. In [3] we proved the following equal-diagonal QRS decomposition lemma.

Lemma 1: For an arbitrary $M \times N$ matrix $\mathbf{H}$ with rank r, there exists a unitary matrix $\mathbf{S}$ such that $\mathbf{HS}$ has an equal-diagonal R-factor; i.e, $\mathbf{HS} = \mathbf{QR}$, where $\mathbf{Q}$ is an $M \times r$ column-wise orthonormal matrix and $\mathbf{R} = [\ \mathbf{R}_{r\times r}\ \ \mathbf{0}_{r\times(N-r)}\ ]$ with $[\mathbf{R}_{r\times r}]_{k,k} = (\lambda_1\lambda_2\cdots\lambda_r)^{1/2r}$, where $\sqrt{\lambda_k}$ is the k-th singular value of $\mathbf{H}$. ■

The key to obtaining such decomposition is to find the unitary matrix $\mathbf{S}$. Once we have had $\mathbf{S}$, we can apply the conventional QR decomposition to the matrix $\mathbf{HS}$ to obtain the equal-diagonal QRS decomposition. The following recursive algorithm is a simplified version of the algorithm in [3] to find the S-factor.

Algorithm 1 (Construction of the S-factor):

1) *SVD*. Perform the SVD (3).
2) *Initialization*. Determine the first column $\breve{\mathbf{s}}_1$ of $\breve{\mathbf{S}}$ such that the following constraints are satisfied,

$$\breve{\mathbf{s}}_1^H\boldsymbol{\Lambda}\breve{\mathbf{s}}_1 = D \tag{4a}$$
$$\breve{\mathbf{s}}_1^H\breve{\mathbf{s}}_1 = 1 \tag{4b}$$

3) *Recursion* (reduce the dimension and decouple constraints). Set $\breve{\mathbf{s}}_{k+1} = \breve{\mathbf{S}}_k^{\perp}\mathbf{z}_{k+1}$, where $\mathbf{z}_{k+1}$ is any vector that satisfies

$$\mathbf{z}_{k+1}^H\mathbf{C}^{(k)}\mathbf{z}_{k+1} = D \tag{4c}$$
$$\mathbf{z}_{k+1}^H\mathbf{z}_{k+1} = 1, \tag{4d}$$

where $\mathbf{C}^{(k)} = \left((\breve{\mathbf{S}}_k^{\perp})^H\boldsymbol{\Lambda}^{-1}\breve{\mathbf{S}}_k^{\perp}\right)^{-1}$.

4) *Complete the S-factor*. $\mathbf{S} = [\mathbf{V}_r\breve{\mathbf{S}}, \overline{\mathbf{V}}_{1,\cdots,r}]$. ■

We would like to make the following remarks on this quadratic recursive algorithm.

(a) The structure of the matrix $\mathbf{C}^{(k)}$ in Step 3) provides a substantial reduction in the computational cost of the recursion over that of the original recursion in [3].
(b) Characterization of the S-factor: In [3] we proved that Algorithm 1 has a solution, and that when $r > 2$ it has an infinite number of solutions. Moreover, the solution obtained in the ith recursion does not affect whether the set of equations in the next recursion is solvable. Therefore, Algorithm 1 characterizes all S-factors such that the matrix $\mathbf{HS}$ has an equal-diagonal R-factor. By choosing a particular solution for each recursion one can significantly reduce the complexity of computing the S-factor, as we show in the next section.
(c) Characterization of the Q-factor: Notice that if $\mathbf{Q}, \mathbf{R}$ and $\mathbf{S}$ are the Q-R-S factors of an invertible matrix $\mathbf{H}$, then, $\mathbf{S}, \mathbf{R}^{-1}$ and $\mathbf{Q}$ are the Q-R-S factors of its inverse, $\mathbf{H}^{-1}$. This observation implies that there is a companion algorithm for $\mathbf{H}^{-1}$ which has the same structure as Algorithm 1. This inverse algorithm actually characterizes the Q-factor in the equal diagonal QRS decomposition of the original matrix $\mathbf{H}$, and this observation leads to an efficient algorithm for finding the S and Q factors.
(d) Uniform decomposition of information: In principal, the quadratic recursive algorithm is based on Schur's decomposition of the determinant of a positive definite matrix into the product of the determinant of an arbitrary principal submatrix and the determinant of its Schur complement. Therefore, Algorithm 1 implicitly describes a process for successively and evenly distributing the total information quantity $D = \det(\boldsymbol{\Lambda})$ over each one-dimensional subspace (or subchannel). We expand on this observation in Section IV.

B. Fast closed-form Q-R-S factors

In this subsection, we will give closed-form equal diagonal QRS decompositions for both a matrix and its inverse by carefully choosing such particular solutions in the quadratic

recursion equations (4a)–(4c) that all the matrices $\mathbf{C}^{(k)}$ are diagonalized simultaneously. To this end, we first define canonical eigen-diagonal matrix sequences from the original singular value matrix.

Definition 1: Let $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_r)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$ and $D = (\lambda_1\lambda_2\cdots\lambda_r)^{1/r}$. A *canonical eigen-diagonal matrix sequence* $\{\ell_k, \mathbf{\Lambda}^{(k)} = \mathrm{diag}(\lambda_1^{(k)}, \lambda_2^{(k)}, \cdots, \lambda_{r-k}^{(k)})\}_{k=1}^{r-1}$ generated by $\mathbf{\Lambda}$ is defined as follows:

1) $\mathbf{\Lambda}^{(1)} = \mathbf{\Lambda}$.
2) If $\frac{\lambda_1^{(k)}\lambda_{r-k+1}^{(k)}}{D} \geq \lambda_2^{(k)}$, we let $\lambda_1^{(k+1)} = \frac{\lambda_1^{(k)}\lambda_{r-k}^{(k)}}{D}, \lambda_2^{(k+1)} = \lambda_2^{(k)}, \cdots, \lambda_{r-k}^{(k+1)} = \lambda_{r-k}^{(k)}$ and $\ell_k = 1$.
3) If $\frac{\lambda_1^{(k)}\lambda_{r-k+1}^{(k)}}{D} \leq \lambda_{r-k}^{(k)}$, we let $\lambda_1^{(k+1)} = \lambda_2^{(k)}, \lambda_2^{(k+1)} = \lambda_3^{(k)}, \cdots, \lambda_{r-k-1}^{(k+1)} = \lambda_{r-k}^{(k)}, \lambda_{r-k}^{(k+1)} = \frac{\lambda_1^{(k)}\lambda_{r-k+1}^{(k)}}{D}$ and $\ell_k = r-k$.
4) If there exists the largest positive integer ℓ_k with $2 \leq \ell_k \leq r-k-1$ such that $\lambda_{\ell_k}^{(k)} \geq \frac{\lambda_1^{(k)}\lambda_{r-k+1}^{(k)}}{D} \geq \lambda_{\ell_k+1}^{(k)}$, then, we let $\lambda_1^{(k+1)} = \lambda_2^{(k)}, \lambda_2^{(k+1)} = \lambda_3^{(k)}, \cdots, \lambda_{\ell_k-1}^{(k+1)} = \lambda_{\ell_k}^{(k)}, \lambda_{\ell_k}^{(k+1)} = \frac{\lambda_1^{(k)}\lambda_{r-k}^{(k)}}{D}, \lambda_{\ell_k+1}^{(k+1)} = \lambda_{\ell_k+1}^{(k)}, \cdots, \lambda_{r-k}^{(k+1)} = \lambda_{r-k}^{(k)}$.

Correspondingly, a *basic rotation sequence* $\{\alpha_k, \beta_k\}_{k=1}^{r-1}$ and its *basic inverse rotation sequence* $\{\widetilde{\alpha}_k, \widetilde{\beta}_k\}_{k=1}^{r-1}$ are defined, respectively, as

$$\alpha_k = \sqrt{\frac{D-\lambda_{r-k+1}^{(k)}}{\lambda_1^{(k)}-\lambda_{r-k+1}^{(k)}}}, \quad \beta_k = \sqrt{\frac{\lambda_1^{(k)}-D}{\lambda_1^{(k)}-\lambda_{r-k+1}^{(k)}}} \tag{5a}$$

$$\widetilde{\alpha}_k = \sqrt{\frac{\widetilde{\lambda}_{r-k+1}^{(k)}-\widetilde{D}}{\widetilde{\lambda}_{r-k+1}^{(k)}-\widetilde{\lambda}_1^{(k)}}}, \quad \widetilde{\beta}_k = \sqrt{\frac{\widetilde{D}-\widetilde{\lambda}_1^{(k)}}{\widetilde{\lambda}_{r-k+1}^{(k)}-\widetilde{\lambda}_1^{(k)}}}, \tag{5b}$$

where, for simplicity and symmetry of the algorithm structure, $\widetilde{D}$ and $\widetilde{\lambda}_{r-k+1}^{(k)}$ denote D^{-1} and $(\lambda_{r-k+1}^{(k)})^{-1}$, respectively. ■

The canonical eigen-diagonal matrix sequence and the basic rotation sequence have the following properties, which will play a key role in deriving our fast closed-form equal-diagonal QRS decompositions.

Lemma 2: Sequences $\{\alpha_k, \beta_k\}_{k=1}^{r-1}$ and $\{\widetilde{\alpha}_k, \widetilde{\beta}_k\}_{k=1}^{r-1}$ satisfy conditions

$$\alpha_k^2 + \beta_k^2 = 1, \quad \lambda_1^{(k)}\alpha_k^2 + \lambda_{r-k+1}^{(k)}\beta_k^2 = D \tag{6a}$$

$$\widetilde{\alpha}_k^2 + \widetilde{\beta}_k^2 = 1, \quad \widetilde{\lambda}_1^{(k)}\widetilde{\alpha}_k^2 + \widetilde{\lambda}_{r-k+1}^{(k)}\widetilde{\beta}_k^2 = \widetilde{D} \tag{6b}$$

$$\widetilde{\alpha}_k = \sqrt{\frac{\lambda_1^{(k)}}{D}}\alpha_k, \ \widetilde{\beta}_k = \sqrt{\frac{\lambda_{r-k+1}^{(k)}}{D}}\beta_k. \tag{6c}$$

■

Now we are in a position to formally state our main result. A proof is provided in the appendix.

Theorem 1: Let $\mathbf{\Lambda}^{1/2} = \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_r})$ with $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_r > 0$ and let the basic rotation sequences $\{\alpha_k, \beta_k\}_{k=1}^{r-1}$ and $\{\widetilde{\alpha}_k, \widetilde{\beta}_k\}_{k=1}^{r-1}$ be defined by (5a) and (5b), respectively. Then, we have the following closed-form equal diagonal QRS decompositions,

$$\mathbf{\Lambda}^{1/2}\,\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \mathbf{S}_{\widetilde{\alpha},\widetilde{\beta}}\,\mathbf{R}_{\boldsymbol{\alpha},\mathbf{\Lambda},\boldsymbol{\beta}}, \tag{7}$$

$$\mathbf{\Lambda}^{-1/2}\mathbf{S}_{\widetilde{\boldsymbol{\alpha}},\widetilde{\boldsymbol{\beta}}} = \mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}\,\mathbf{R}_{\widetilde{\boldsymbol{\alpha}},\mathbf{\Lambda}^{-1},\widetilde{\boldsymbol{\beta}}}, \tag{8}$$

where $\mathbf{R}_{\boldsymbol{\alpha},\mathbf{\Lambda},\boldsymbol{\beta}}$ and $\mathbf{R}_{\widetilde{\boldsymbol{\alpha}},\mathbf{\Lambda}^{-1},\widetilde{\boldsymbol{\beta}}}$ are upper triangular matrices with $[\mathbf{R}_{\boldsymbol{\alpha},\mathbf{\Lambda},\boldsymbol{\beta}}]_{k,k} = (\lambda_1\lambda_2\cdots\lambda_r)^{\frac{1}{2r}} = [\mathbf{R}_{\widetilde{\boldsymbol{\alpha}},\mathbf{\Lambda}^{-1},\widetilde{\boldsymbol{\beta}}}]_{k,k}^{-1}$ and $\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ is an $r \times r$ unitary matrix that can be represented by the following two kinds of decompositions:

1) Column decomposition: The n-th column of $\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ is determined by

$$\mathbf{s}_1 = \mathbf{b}_{r-2,\alpha_1,\beta_1} \tag{9a}$$

$$\mathbf{s}_r = \prod_{k=1}^{r-2}\mathbf{\Gamma}_{\ell_k,\alpha_k,\beta_k,r-k}\mathbf{b}_{-1,\alpha_{r-1},\beta_{r-1}} \tag{9b}$$

$$\mathbf{s}_n = \prod_{k=1}^{n-1}\mathbf{\Gamma}_{\ell_k,\alpha_k,\beta_k,r-k}\mathbf{b}_{r-n-1,\alpha_n,\beta_n} \tag{9c}$$

for $n = 2, 3, \cdots, r-1$, where

$$\mathbf{\Gamma}_{m,\alpha,\beta,n} = \begin{pmatrix} \mathbf{0}_{1\times(m-1)} & \beta & \mathbf{0}_{1\times(n-m)} \\ \mathbf{I}_{m-1} & \mathbf{0}_{(m-1)\times 1} & \mathbf{0}_{(m-1)\times 1} \\ \mathbf{0}_{(n-m)\times(m-1)} & \mathbf{0}_{(n-m)\times 1} & \mathbf{I}_{n-m} \\ \mathbf{0}_{1\times(m-1)} & \alpha & \mathbf{0}_{1\times(n-m)} \end{pmatrix}, \tag{9d}$$

$\mathbf{b}_{-1,\alpha,\beta} = \left(\begin{smallmatrix}\beta\\ \alpha\end{smallmatrix}\right)$ and $\mathbf{b}_{k,\alpha,\beta} = \begin{pmatrix}\alpha \\ \mathbf{0}_{k\times 1} \\ -\beta\end{pmatrix}$.

2) Product decomposition: $\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ can be factored into $\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \prod_{i=1}^{r-1}\mathbf{S}_{\alpha_i,\ell_i,\beta_i}$, where

$$\mathbf{S}_{\alpha_i,\ell_i,\beta_i} = \begin{pmatrix} \mathbf{I}_{i-1} & \mathbf{0}_{(i-1)\times 1} & \mathbf{0}_{(i-1)\times(r-i)} \\ \mathbf{0}_{(r+1-i)\times(i-1)} & \mathbf{b}_{r-1-i,\alpha_i,\beta_i} & \mathbf{\Gamma}_{\ell_i,\alpha_i,\beta_i,r-i} \end{pmatrix}.$$

The same results holds for $\mathbf{S}_{\widetilde{\boldsymbol{\alpha}},\widetilde{\boldsymbol{\beta}}}$ by replacing the basic rotation sequences $\{\alpha_k, \beta_k\}_{k=1}^{r-1}$ in $\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ by its inverse basic rotation sequences $\{\widetilde{\alpha}_k, \widetilde{\beta}_k\}_{k=1}^{r-1}$. ■

We would like to make the following observations.

1) Different eigen-diagonal matrix sequences result in different rotation sequences and thus, different specific Q-R-S triples.
2) Combining Lemma 1, Theorem 1 and Algorithm 1 we have a specific Q-R-S triple in the equal-diagonal QRS decomposition of a general matrix $\mathbf{H}$; i.e., $\mathbf{Q} = \mathbf{U}_r\mathbf{S}_{\widetilde{\boldsymbol{\alpha}},\widetilde{\boldsymbol{\beta}}}$, $\mathbf{R}_{r\times r} = \mathbf{R}_{\boldsymbol{\alpha},\mathbf{\Lambda},\boldsymbol{\beta}}$ and $\mathbf{S} = [\mathbf{V}_r\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}, \overline{\mathbf{V}}_{1,2,\cdots,r}]$.

IV. Uniform Decomposition of Mutual Information

In this section we combine the canonical property of the MMSE-DF detector [6], [7], [13] with an equal-diagonal QRS decomposition of the mutual information matrix $\mathcal{I}_{\mathbf{HT}}^{1/2}$ to obtain a uniform factorization of the Gaussian mutual information for the MMSE-DF detector. Then, we give some perspective to such factorization from both information theoretic and signal estimation and detection error viewpoints, with an emphasis on building a comprehensive understanding of the MMSE-DF receiver and the resulting optimal transmitter.

It is known that under the assumption that the channel matrix $\mathbf{H}$ is available to both the receiver and the transmitter, the Gaussian mutual information for the precoded channel model (1) is given by $I_G(\mathbf{x};\mathbf{y}) = \log\det(\mathbf{I}+\mathrm{snr}\mathbf{T}^H\mathbf{H}^H\mathbf{H}\mathbf{T})$. The standard approach to the design of a transmitter that maximizes $I_G(\mathbf{x};\mathbf{y})$ subject to a power constraint uses the singular vector decomposition (SVD) to diagonalize the original matrix channel matrix $\mathbf{H}$ in (1). That results in an optimal $\mathbf{T}$ of the form $\mathbf{T} = \mathbf{V}[\ \breve{\mathbf{\Xi}}, \mathbf{0}_{K\times(N-K)}\ ]^T\mathbf{W} = \tilde{\mathbf{T}}\mathbf{W}$, where $\mathbf{V}$ is the singular value matrix of $\mathbf{H}$ given in (3), $\breve{\mathbf{\Xi}} = \mathrm{diag}(\mu_1, \mu_2, \cdots, \mu_K)$ and $\mathbf{W}$ is a $K \times K$ unitary matrix. For this choice of $\mathbf{T}$, $I_G(\mathbf{x};\mathbf{y}) = \sum_{k=1}^{K}\log(1+\mathrm{snr}\lambda_k\mu_k^2)$,

and optimal power loading is performed over the eigen-subchannels.

An alternative decomposition of $I_G(\mathbf{x};\mathbf{y})$ is the QR decomposition of the mutual information matrix. It was proved [13] that under an assumption of error-free feedback, the mutual information between the $(K-i)$th symbol (or user) x_{K-i} and $\mathbf{y}$ conditioned on $\mathbf{x}_K^{K-i+1} = [x_K, x_{K-1}, \cdots, x_{K-i+1}]$ for the model (1) can be expressed as $I_G(x_{K-i};\mathbf{y}|\mathbf{x}_K^{K-i+1}) = \log\big([\mathbf{R}_{\mathcal{I}_{\mathbf{HT}}^{1/2}}]^2_{K-i,K-i}\big)$. Therefore, $I_G(\mathbf{x};\mathbf{y}) = \sum_{i=0}^{K-1} I(x_{K-i};\mathbf{y}|\mathbf{x}_K^{K-i+1}) = \sum_{i=0}^{K-1}\log\big([\mathbf{R}_{\mathcal{I}_{\mathbf{HT}}^{1/2}}]^2_{K-i,K-i}\big) = \log\det(\mathbf{I}+\mathrm{snr}\mathbf{T}^H\mathbf{H}^H\mathbf{H}\mathbf{T})$. This shows that the MMSE-DF receiver is information lossless. For a given matrix $\mathbf{H}$, its singular values are fixed under any unitary transformation and hence, its eigen-subchannel decomposition of mutual information does not change. However, the R-factor diagonal entries of the mutual information matrix change with the unitary transformation, and as a result the capacity of each R-factor subchannel of the MMSE-DF detector will change too. (These subchannels are implicit in (2).) In other words, different unitary matrices $\mathbf{W}$ result in the different R-factors, and hence different R-factor subchannel capacities and different detection error performance for the MMSE-DF detector. Therefore, a natural question is that what is the "best" distribution of the total mutual information to each R-factor subchannel (2) for the MMSE-DF detector? As proved in the appendix, a candidate is the uniform distribution.

Theorem 2: Under an assumption of error-free feedback, the total mutual information for the system in (1) can be uniformly decomposed into the sum of that of each R-factor subchannel (2) for the MMSE-DF detector by rotating the input signal vector with the S-factor of its mutual information matrix. ■

In order to provide some properties of the uniform decomposition of mutual information, we provide the following three propositions. Let $\mathbf{e} = \widetilde{\mathbf{x}} - \mathbf{x} = (\sqrt{\mathrm{snr}}\,\mathbf{F}\mathbf{H}\mathbf{T} - \mathbf{I} - \mathbf{B})\mathbf{x} + \mathbf{F}\boldsymbol{\xi}$ denote the error vector between the transmitted symbols and the corresponding receiver estimates.

Proposition 1: The arithmetic MSE [2] of the MMSE-DF receiver is $\mathrm{MSE}_a(\mathbf{HT}) = \frac{1}{K}\mathrm{tr}\big(\mathrm{E}[\mathbf{e}\mathbf{e}^H]\big)$. Let C denote the capacity of the channel in (1). Then, $C \geq I_G(\mathbf{x};\mathbf{y}) \geq -\log(\mathrm{MSE}_a(\mathbf{HT}))$. Maximizing this lower bound is equivalent to minimizing $\mathrm{MSE}_a(\mathbf{HT})$. The maximum value of this lower bound is C, which is achieved by choosing $\mathbf{T}$ such that $I_G(\mathbf{x};\mathbf{y})$ is maximized, and then using the remaining unitary degree of freedom to achieve the uniform decomposition of mutual information. ■

Proof: It is clear that $C \geq I_G(\mathbf{x};\mathbf{y})$. Since the error covariance matrix [13] of the MMSE-DFE is $\mathrm{E}[\mathbf{e}\mathbf{e}^H] = \mathrm{diag}([\mathbf{R}_{\mathcal{I}_{\mathbf{HT}}}]^{-2}_{1,1}, [\mathbf{R}_{\mathcal{I}_{\mathbf{HT}}}]^{-2}_{2,2}, \cdots, [\mathbf{R}_{\mathcal{I}_{\mathbf{HT}}}]^{-2}_{K,K})$, we have $\mathrm{MSE}_a(\mathbf{HT}) = \sum_{1\leq k\leq K}[\mathbf{R}_{\mathcal{I}_{\mathbf{HT}}}]^{-2}_{k,k}$. On the other hand, $I_G(\mathbf{x};\mathbf{y}) = \log\det(\mathbf{I}+\mathrm{snr}\mathbf{T}^H\mathbf{H}^H\mathbf{H}\mathbf{T}) = \log\big(\prod_{k=1}^K[\mathbf{R}_{\mathcal{I}_{\mathbf{HT}}}]^2_{k,k}\big) \geq -K\log\big(\frac{1}{K}\sum_{k=1}^K[\mathbf{R}_{\mathcal{I}_{\mathbf{HT}}}]^{-2}_{k,k}\big) = -K\log(\mathrm{MSE}_a(\mathbf{HT}))$ with equality holding if and only if all $[\mathbf{R}_{\mathcal{I}_{\mathbf{HT}}}]^2_{k,k}$ are equal. (Here we have used the arithmetic-geometric mean inequality.) Therefore, maximizing the lower bound on the Gaussian mutual information is equivalent to minimizing the arithmetic MSE. This can be shown to be equivalent to first maximizing the Gaussian mutual information and then uniformly distributing total channel capacity into each R-factor subchannel by choosing the free unitary matrix in the water-filling solution family as the S-factor. □

The following two propositions further describe the optimality of the uniform distribution of mutual information into R-factor subchannels for the MMSE-DF detector.

Proposition 2: Suppose we wish to use an ordered MMSE-DF detector for a channel where the mutual information matrix has the equal-diagonal R-factor. Then, the optimal detection order (that ensures that the high SINR components are detected first) is the natural order, i.e., the i-th symbol to be detected is the symbol x_{N+1-i}.

Definition 2: Define the minimum distance of a finite constellation $\mathcal{X}$ as

$$d_{\min}(\mathcal{X}) = \min_{x\neq x', x, x'\in\mathcal{X}} |x-x'| = \sqrt{\min_{\mathbf{x},\mathbf{x}'\in\mathcal{X}^N, \mathbf{x}\neq\mathbf{x}'} ||\mathbf{x}-\mathbf{x}'||^2}.$$

■

Definition 3: Define the free distance of an $M\times N$ channel matrix $\mathbf{H}$ as

$$d_{\mathrm{free}}(\mathbf{H}) = \sqrt{\min_{\mathbf{x},\mathbf{x}'\in\mathcal{X}^N, \mathbf{x}\neq\mathbf{x}'} (\mathbf{x}-\mathbf{x}')^H\mathbf{H}^H\mathbf{H}(\mathbf{x}-\mathbf{x}')}.$$

■

The following proposition shows the asymptotic behavior of the free distance, and hence that of the MMSE-DF detector for a channel with the equal-diagonal R-factor of its mutual information matrix.

Proposition 3: Let the SVD of $\mathbf{H}$ be given by (3) and the signaling point in constellation $\mathcal{X}$ be PAM or PSK or QAM. When $\mathbf{T} = \mathbf{V}\boldsymbol{\Xi}\mathbf{S} = \widetilde{\mathbf{T}}\mathbf{S}$ with $\mathrm{tr}(\mathbf{T}^H\mathbf{T}) \leq p$, where $\mathbf{S}$ is obtained by applying Algorithm 1 to $\mathcal{I}^{1/2}_{\mathbf{H}\widetilde{\mathbf{T}}}$, we have

$$\lim_{\mathrm{snr}\to\infty} d_{\mathrm{free}}(\mathbf{HT}) = \sqrt{\frac{p}{r}}\Big(\prod_{k=1}^r \lambda_k\Big)^{1/2r} d_{\min}(\mathcal{X}). \tag{10}$$

Therefore, under the Gaussian approximation (e.g., [14]), the MMSE-DF detector is asymptotically equivalent to the ML detector in the sense that

$$\lim_{\mathrm{snr}\to\infty} \frac{\ln \mathrm{P}_{\mathrm{ML}}(\mathrm{snr})}{\ln \mathrm{P}_{\mathrm{MMSE-DFE}}(\mathrm{snr})} = 1, \tag{11}$$

where $P_{\mathrm{ML}}(\mathrm{snr})$ and $P_{\mathrm{MMSE-DFE}}(\mathrm{snr})$ denote the block error probability of the ML detector and MMSE-DF detector, respectively. ■

V. Conclusion

In this paper we have shown that if the precoder in (1) is chosen to be $\mathbf{T} = \mathbf{V}\boldsymbol{\Xi}\mathbf{S} = \widetilde{\mathbf{T}}\mathbf{S}$, where $\widetilde{\mathbf{T}}$ performs a water-filling power allocation over the eigen vectors of $\mathbf{H}^H\mathbf{H}$ and $\mathbf{S}$ is determined by applying Algorithm 1 to $\mathcal{I}^{1/2}_{\mathbf{H}\widetilde{\mathbf{T}}}$, then we obtain a uniform decomposition of the maximum Gaussian mutual information of the channel matrix $\mathbf{H}$ into implicit scalar subchannels, each of which has the same mutual information.

APPENDIX

A. Proof of Theorem 1

First, we show that $\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ defined by (9) is an orthonormal matrix. Let $\mathbf{s}_m$ and $\mathbf{s}_n$ denote the m-th and n-th columns ($n \le m$) of $\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$, respectively. Since the matrix $\boldsymbol{\Gamma}_{\ell_k,\alpha_k,\beta_k,r-k}$ is column-wise orthonormal, we have that

$$\mathbf{s}_n^T\mathbf{s}_m = \mathbf{b}_{r-n-1,\,\alpha_n,\,\beta_n}^T \prod_{k=n}^{m-1} \boldsymbol{\Gamma}_{\ell_k,\,\alpha_k,\,\beta_k,\,r-k}\mathbf{b}_{r-m-1,\,\alpha_m,\,\beta_m}. \quad (12)$$

In addition, we notice that $\mathbf{b}_{r-n-1,\,\alpha_n,\,\beta_n}^T\boldsymbol{\Gamma}_{\ell_n,\,\alpha_n,\,\beta_n,\,r-n} = \mathbf{0}_{1\times(r-n)}$. Combining this with (12) yields $\mathbf{s}_n^T\mathbf{s}_m = 0$ for $n < m$. For $n = m$, equation (12) becomes $\mathbf{s}_n^T\mathbf{s}_m = \mathbf{b}_{r-n-1,\,\alpha_n,\,\beta_n}^T\mathbf{b}_{r-n-1,\,\alpha_n,\,\beta_n} = 1$, where we have used (6a). Therefore, $\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ is an orthonormal matrix. We can prove that $\mathbf{S}_{\widetilde{\boldsymbol{\alpha}},\widetilde{\boldsymbol{\beta}}}$ is also an orthonormal matrix in a similar way.

We now show that $\mathbf{S}_{\widetilde{\boldsymbol{\alpha}},\widetilde{\boldsymbol{\beta}}}^T\boldsymbol{\Lambda}^{1/2}\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ is an upper triangular matrix with equal diagonal entries. Let $\widetilde{\mathbf{s}}_m$ denote the m-th column of $\mathbf{S}_{\widetilde{\boldsymbol{\alpha}},\widetilde{\boldsymbol{\beta}}}^T$. Note that $\boldsymbol{\Gamma}_{\ell_k,\widetilde{\alpha}_k,\widetilde{\beta}_k,r-k}^T\sqrt{\boldsymbol{\Lambda}^{(k)}}\boldsymbol{\Gamma}_{\ell_k,\alpha_k,\beta_k,r-k} = \sqrt{\boldsymbol{\Lambda}^{(k+1)}}$, where $\boldsymbol{\Lambda}^{(k)}$ is the canonical eigen-diagonal matrix sequence in Definition 1. Therefore, when $m > n$, we can use the fact that $\sqrt{\lambda_1^{(n)}}\widetilde{\beta}_n\alpha_n - \sqrt{\lambda_{r-n+1}^{(n)}}\widetilde{\alpha}_n\beta_n = 0$ to show that $\widetilde{\mathbf{s}}_m^T\boldsymbol{\Lambda}^{1/2}\mathbf{s}_n = 0$ When $m = n$, we have $\widetilde{\mathbf{s}}_n^T\boldsymbol{\Lambda}^{1/2}\mathbf{s}_n = \sqrt{\lambda_1^{(n)}}\widetilde{\alpha}_n\alpha_n + \sqrt{\lambda_{r-n+1}^{(n)}}\widetilde{\beta}_n\beta_n = \sqrt{D}$. Similarly, we can prove the product decomposition of $\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$. □

B. Proof of Proposition 3

We first note that $\mathbf{T}^H\mathbf{H}^H\mathbf{H}\mathbf{T} = \mathrm{snr}^{-1}(\mathcal{I}_{\mathbf{HT}} - \mathbf{I})$. Now consider two different signal vectors: $\mathbf{x} = [x_1, x_2, \cdots, x_K]^T$ and $\mathbf{x}' = [x_1', x_2', \cdots, x_K']^T$. If $x_k = x_k'$ for $k = 2, \cdots, K$, but $x_1 \neq x_1'$, then

$$\begin{aligned}(\mathbf{x}-\mathbf{x}')^H\mathbf{T}^H\mathbf{H}^H\mathbf{H}\mathbf{T}(\mathbf{x}-\mathbf{x}') &= \mathrm{snr}^{-1}\left((\mathbf{x}-\mathbf{x}')^H\mathcal{I}_{\mathbf{HT}}(\mathbf{x}-\mathbf{x}') - |x_1 - x_1'|^2\right)\\ &= \mathrm{snr}^{-1}([\mathbf{R}]_1^2 - 1)|x_1 - x_1'|^2. \quad (13)\end{aligned}$$

Hence, by taking the minima of both sides of (13), we have

$$d_{\mathrm{free}}^2(\mathbf{HT}) \le \mathrm{snr}^{-1}([\mathbf{R}]_1^2 - 1)\cdot d_{\min}^2(\mathcal{X}). \quad (14)$$

Since $\lim_{\mathrm{snr}\to\infty}\mathrm{snr}^{-1}[\mathbf{R}]_1^2 = \frac{p}{r}\left(\prod_{k=1}^r\lambda_k\right)^{1/r}$, (14) implies that

$$\lim_{\mathrm{snr}\to\infty} d_{\mathrm{free}}(\mathbf{HT}) \le \sqrt{\frac{p}{r}}\Big(\prod_{k=1}^r\lambda_k\Big)^{1/2r}\cdot d_{\min}(\mathcal{X}). \quad (15)$$

On the other hand, we note that

$$\begin{aligned}(\mathbf{x}-\mathbf{x}')^H\mathbf{T}^H\mathbf{H}^H\mathbf{H}\mathbf{T}(\mathbf{x}-\mathbf{x}') &= \mathrm{snr}^{-1}\big((\mathbf{x}-\mathbf{x}')^H\mathcal{I}_{\mathbf{HT}}(\mathbf{x}-\mathbf{x}') - \|\mathbf{x}\|^2\big)\\ &= \mathrm{snr}^{-1}\Big(\sum_{i=1}^K\Big|\sum_{j=i}^K R_{i,j}\cdot(x_j - x_j')\Big|^2 - \|\mathbf{x}\|^2\Big). \quad (16)\end{aligned}$$

Assume $\mathbf{x} \neq \mathbf{x}'$. Let k be an integer such that $x_i = x_i'$, for $i > k$, but $x_k \neq x_k'$. Then, from (16), using the upper triangularity of $\mathbf{R}$, we have

$$\begin{aligned}&(\mathbf{x}-\mathbf{x}')^H\mathbf{T}^H\mathbf{H}^H\mathbf{H}\mathbf{T}(\mathbf{x}-\mathbf{x}')\\ &\quad= \mathrm{snr}^{-1}\Big(\sum_{i=1}^k\Big|\sum_{j=i}^k R_{i,j}(x_j - x_j')\Big|^2 - \|\mathbf{x}\|^2\Big)\\ &\quad\ge \mathrm{snr}^{-1}\left(([\mathbf{R}]_1^2 - 1)|x_k - x_k'|^2 - \|\mathbf{x}\|^2\right)\\ &\quad\ge \mathrm{snr}^{-1}\left(([\mathbf{R}]_1^2 - 1)\cdot d_{\min}^2(\mathcal{X}) - \|\mathbf{x}\|^2\right). \quad (17)\end{aligned}$$

Taking the minima of both sides of (17) yields

$$d_{\mathrm{free}}^2(\mathbf{HT}) \ge \mathrm{snr}^{-1}\left(([\mathbf{R}]_1 - 1)\cdot d_{\min}(\mathcal{X}) - \|\mathbf{x}\|_{\max}^2\right). \quad (18)$$

Since the constellation $\mathcal{X}$ is finite, the quantity $\|\mathbf{x}\|_{\max}^2$ is bounded. Therefore, using (18) we have that

$$\lim_{\mathrm{snr}\to\infty} d_{\mathrm{free}}(\mathbf{HT}) \ge \sqrt{\frac{p}{r}}\left(\prod_{k=1}^r\lambda_k\right)^{1/2r}\cdot d_{\min}(\mathcal{X}). \quad (19)$$

Combining (15) with (19), we complete the proof of (10). Moreover, we know from [15] and [8] that

$$\lim_{\mathrm{snr}\to\infty}\frac{\ln \mathrm{P}_{\mathrm{MLD}}(\mathrm{snr})}{\ln Q\left(\frac{\sqrt{\mathrm{snr}}\,d_{\mathrm{free}}(\mathbf{HT})}{2}\right)} = \lim_{\mathrm{snr}\to\infty}\frac{\ln \mathrm{P}_{\mathrm{MMSE-DFE}}(\mathrm{snr})}{\ln Q\left(\frac{\sqrt{\mathrm{snr}}\,d_{\mathrm{free}}(\mathbf{HT})}{2}\right)} = 1.$$

REFERENCES

[1] L. H. Brandenburg and A. D. Wyner, "Capacity of the Gaussain channel with memory," *Bell System Tech. J.*, vol. 53, pp. 745–779, May–June 1974.

[2] J. Yang and S. Roy, "Joint transmitter–receiver optimization for multi-input multi-output systems with decision feedback," *IEEE Trans. Informat. Theory*, vol. 40, pp. 1334–1347, Sep. 1994.

[3] J.-K. Zhang, A. Kavčić, and K. M. Wong, "Equal-diagonal QR decomposition and its application to precoder design for successive cancellation detection," *IEEE Trans. Inform. Theory*, vol. 51, pp. 154–172, Jan. 2005.

[4] A. Duel-Hallen, "Decorrelating decision-feedback multiuser detector for synchronous code-division multiple-access channels," *IEEE Trans. Commun.*, vol. 41, pp. 285–290, Feb. 1993.

[5] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney, "MMSE decision-feedback equalizers and coding–Part I: Equalization results; Part II: Coding results," *IEEE Trans. Comm.*, vol. 43, pp. 2582–2604, Oct 1995.

[6] J. M. Cioffi and G. D. Forney, "Generalized decision-feedback equalization packet transmission with ISI and Gaussain noise", in *Communications, Computation, Control and Signal Processing*, A. Paulraj, V. Roychowdhury and C. Schaper, Eds, Kluwer, Boston, 1997.

[7] G. D. Forney, Jr., "Shannon meets Wiener II: On MMSE estimation in successive decoding schemes," *Proc. 2004 Allerton Conf.*, Oct. 2004.

[8] M. K. Varanasi, "Decision feedback multiuser detection: A systematic approach," *IEEE Trans. Inform. Theory*, vol. 45, pp. 110–121, Jan. 1999.

[9] G. Woodward, R. Ratasuk, M. L. Honig, and P. B. Rapajic, "Minimum mean-squared error multiuser decision feedback detection for DS-CDMA," *IEEE Trans. Commun.*, vol. 50, pp. 2104–2112, Dec. 2002.

[10] T. Guess, "Optimal sequences for CDMA with decision-feedback receivers," *IEEE Trans. Informat. Theory*, vol. 49, pp. 886–900, Apr. 2003.

[11] B. Hassibi, "An fast square root algorithm for BLAST," in *Proc. Asilomar Conf. Signals, Systems, Computers*, pp. 1255–1259, Nov. 2000.

[12] H. El-Gamal, G. Caire and M. O. Damen, "Lattice coding and decoding achieve the optimal diversity–multiplexing tradeoff of MIMO channels," *IEEE Trans. Informat. Theory*, vol. 50, pp. 968–985, June 2004.

[13] T. Guess and M. K. Varanasi, "An information-theoretic framework for deriving canonical decision-feedback receivers Gaussian channels," *IEEE Trans. Informat. Theory*, vol. 51, pp. 173–187, Jan. 2005.

[14] H. V. Poor and S. Verdu, "Probability of error in MMSE multiuser detection," *IEEE Trans. Informat. Theory*, vol. 43, pp. 858–871, May 1997.

[15] G. D. Forney, Jr., and G. Ungerboeck, "Modulation and coding for linear Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2384–2415, Oct. 1998.