Constrained Detection for Spatial-Multiplexing Multiple-Input–Multiple-Output Systems

Tao Cui, *Student Member, IEEE*, Chintha Tellambura, *Senior Member, IEEE*, and Yue Wu

***Abstract*—A family of detectors that exploit signal constraints is developed for maximum-likelihood detection for multiple-input–multiple-output (MIMO) systems. Real constrained detectors and decision-feedback detectors are proposed for real constellations by forcing the relaxed solution to be real. A generalized minimum mean square error (GMMSE) and constrained least squares MIMO detectors are also developed for unitary and nonunitary signal constellations. Using these constrained detectors, we propose a new ordering scheme to achieve a tradeoff between interference suppression and noise enhancement. Moreover, to mitigate the inherent error propagation, the decision-feedback MIMO detectors are integrated with signal constraints. The simulation results show that our combined detector achieves a significant performance gain over vertical Bell Laboratories layered space-time (V-BLAST) detection.**

***Index Terms*—Decision-feedback detector (DFD), linear detector, maximum likelihood, multiple-input–multiple-output (MIMO).**

I. Introduction

Multiple-input–multiple-output (MIMO) wireless communication systems with spatial multiplexing can potentially achieve remarkably high spectral efficiencies in rich scattering multipath environments. Consequently, efficient signal detection algorithms for spatial-multiplexing MIMO systems have attracted much interest. A prime example is the vertical Bell Laboratories layered space-time (V-BLAST) detector [1].

Although the optimal maximum-likelihood detector (MLD) achieves the minimum error probability for independent identically distributed (i.i.d.) random symbols, which is a requirement that holds in many cases, the complexity of the MLD exponentially grows with the number of transmit antennas and the number of bits that index each scalar constellation point, making the MLD computationally prohibitive in most cases. Therefore, various computationally efficient suboptimal detection algorithms based on linear or feedback receiver structures using the zero-forcing (ZF) or the minimum mean square error (MMSE) criterion have been developed. For instance, V-BLAST [1] involves symbol ordering and sequential detection. Prior to the detection of a symbol, interference from previously detected symbols is subtracted (the canceling step), and the received vector elements are linearly weighted to null the interference from the yet undetected symbols (the nulling step). The equivalence between V-BLAST and a generalized decision-feedback equalizer has been demonstrated [2]. A large number of extensions to the basic V-BLAST have been investigated in the literature. However, these suboptimal receivers perform much worse than the MLD. On the other hand, the sphere decoder (SD) [3]–[5] offers the optimal MLD performance at reduced complexity, particularly in the high SNR region. Nevertheless, its worst-case complexity is exponential in the number of transmit antennas, and its average complexity is high in a low SNR or for large systems [6]. The performance and complexity gaps between the MLD and the existing suboptimal receivers have motivated the development of alternative detectors.

The MIMO detection problem requires minimizing a quadratic cost function over the discrete set of all possible transmit vectors. In the relaxation approach, this discrete set is embedded in a larger bounded multidimensional continuous space, and the minimization is performed over this continuous space subject to certain constraints. The resulting minimum solution is mapped back into the original discrete space. Several such constrained detectors have been developed [7]–[11]. For example, a generalized MMSE (GMMSE) detector for code-division multiple-access (CDMA) systems has been proposed [7], where the constrained optimization problem resulting from the relaxation of the binary phase-shift keying (BPSK) vectors inside the unit hypersphere is solved via the convex duality theorem and the gradient descent. In [8], a tighter relaxation is used in orthogonal frequency division multiplexing (OFDM)/spatial division multiple-access (SDMA) systems employing unitary constellations by restricting the binary vectors in the hypersphere, resulting in the constrained least squares (CLS) detector. In [9], semidefinite relaxation (SDR) has been developed for BPSK-CDMA systems. The SDR has been also extended to general M phase-shift keying (M-PSK) and quadrature amplitude modulation (QAM) constellations in [10], [12], and [13].

In this paper, constrained linear detectors and decision-feedback detectors (DFDs) are developed for spatial-multiplexing MIMO systems. Real constrained detectors and DFDs are proposed for real constellations by suppressing the imaginary interference component. We also generalize

Manuscript received June 29, 2006; revised October 18, 2006, June 21, 2007, and July 30, 2007. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada, the Informatics Circle of Research Excellence, and the Alberta Ingenuity Fund. This paper was presented in part at the IEEE Global Telecommunications Conference, November 2005, St. Louis, MO, USA. The review of this paper was coordinated by Prof. L. Lampe.

T. Cui was with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada. He is now with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: taocui@caltech.edu).

C. Tellambura and Y. Wu are with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: chintha@ece.ualberta.ca; yuewu@ece.ualberta.ca).

Digital Object Identifier 10.1109/TVT.2007.907291

the CLS detector of [8] by dividing the signal vector to several subgroups and applying the unitary constraints to these subgroups. Similarly, the GMMSE detector [7] is extended to nonunitary constellations. A new ordering scheme is also proposed by using the constrained detectors, which maximizes the signal-to-interference-plus-noise ratio (SINR) at each step. Reference [14] shows that the first detected symbol limits the overall performance of V-BLAST. We, thus, combine the constrained detector and the decision feedback to improve the quality of the first few detected symbols. An earlier version of our work appears in [15].

This paper is organized as follows. Section II describes the spatial-multiplexing MIMO system model. Real constrained detectors, modulus constrained subgroup detectors, and coordinate ascent improvement are developed in Section III. Constrained decision-feedback receivers are developed in Section IV. Simulation results are given in Section V, and this paper concludes in Section VI.

1) Notation: Bold symbols denote matrices or vectors. $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^*$ denote the transpose, the conjugate transpose, and the conjugate, respectively. $(\cdot)^\dagger$ denotes the pseudoinverse. $\mathrm{Re}\{x\}$ and $\mathrm{Im}\{x\}$ denote the real and imaginary parts of x, respectively. $\|(\cdot)\|^2$ is the squared norm of $(\cdot)$. The sets of real numbers and complex numbers are $\mathbb{R}$ and $\mathbb{C}$, respectively, and the set of all complex $K \times 1$ vectors is denoted by $\mathbb{C}^K$. A circularly complex Gaussian variable with mean μ and variance σ^2 is denoted by $z \sim \mathcal{CN}(\mu, \sigma^2)$. The $N \times N$ identity matrix and the diagonal matrix formed by vector $\mathbf{a}$ are $\mathbf{I}_n$ and $\mathrm{diag}(\mathbf{a})$, respectively.

II. System Model

We consider a standard MIMO system with n transmit antennas and m receive antennas and with spatial multiplexing, where the individual antennas transmit independent signals rather than jointly encoded ones. That is, the input data stream is demultiplexed into n equal-rate substreams, and each is simultaneously sent through one of the n antennas over a rich scattering channel. A finite modulation constellation $\mathcal{Q}$ is used. We consider a flat fading MIMO channel, where each receive antenna collects signals from all the n transmit antennas. The discrete-time equivalent baseband received signals can, thus, be written as

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{x} = [x_1, \dots, x_n]^T$, $x_i \in \mathcal{Q}$, is the transmitted signal vector, $\mathbf{r} = [r_1, \dots, r_m]^T$, $r_i \in \mathbb{C}$, is the received signal vector, $\mathbf{H} = [h_{i,j}] \in \mathbb{C}^{m \times n}$ is the channel matrix, and $\mathbf{n} = [n_1, \dots, n_m]^T$, $n_i \in \mathbb{C}$, is an additive white Gaussian noise vector. The elements of $\mathbf{H}$ are i.i.d. complex Gaussian, $h_{i,j} \sim \mathcal{CN}(0,1)$. The components of $\mathbf{n}$ are i.i.d. with $n_i \sim \mathcal{CN}(0, \sigma_n^2)$. We assume that the channel is perfectly known to the receiver, and that $n \leq m$. If $n > m$, we can readily transform the rank deficient problem into a full rank problem, as shown in [16]. Note that (1) models any linear, synchronous, and flat fading channels. Therefore, all our detectors can be readily applied to CDMA systems.

Given the standard model (1), the MLD that minimizes the average error probability is given by

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathcal{Q}^n} \|\mathbf{r} - \mathbf{H}\mathbf{x}\|^2. \tag{2}$$

Due to the discrete nature of $\mathcal{Q}$, (2) is a nondeterministic-polynomial-time-hard problem, and an exhaustive search for $\hat{\mathbf{x}}$ has a complexity exponential in n.

III. Constrained Detectors

A. Classic Receivers

We briefly review several classic receivers. Since $\mathcal{Q} \subseteq \mathbb{C}$, a simple relaxation is to allow each $x_i \in \mathbb{C}$. This relaxation results in the well-known ZF (decorrelating) detector, and (2) becomes

$$\hat{\mathbf{x}}_{\mathrm{ZF}} = D\left[\arg\min_{\mathbf{x} \in \mathbb{C}^n} \|\mathbf{r} - \mathbf{H}\mathbf{x}\|^2\right] \tag{3}$$

where $D[x]$ denotes the threshold detection rule that yields the constellation symbol that is closest to x. For a vector $\mathbf{x}$, $D[\mathbf{x}]$ individually operates on each element. The minimization part in (3) has the least squares solution, and the ZF detector can, thus, be written as

$$\hat{\mathbf{x}}_{\mathrm{ZF}} = D\left[(\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{r}\right]. \tag{4}$$

If the same relaxation is combined with the minimization of the MSE between the transmitted signals and the detected signals $E\{\|\mathbf{x} - \hat{\mathbf{x}}\|^2\}$, then the MMSE prefilter output is $\hat{\mathbf{x}} = \mathbf{G}\mathbf{r}$ ($\mathbf{G}$ is a prefilter matrix). Using the orthogonality principle [17], one can determine the prefilter matrix, and the MMSE linear receiver is then given by

$$\hat{\mathbf{x}}_{\mathrm{MMSE}} = D\left[\left(\mathbf{H}^H\mathbf{H} + \sigma_n^2\mathbf{I}_n\right)^{-1}\mathbf{H}^H\mathbf{r}\right]. \tag{5}$$

However, the ZF and MMSE linear receivers do not guarantee the optimal solution (2) due to the looseness of the relaxation.

For additional details, see [7] and [18].

B. Real Constrained Detectors

A real constellation $\mathcal{Q}$ has all real elements, e.g., BPSK and pulse amplitude modulation. If the real signals are transmitted through a complex channel like (1), the received signals are complex, and the ZF and MMSE solutions from (4) and (5) are usually complex vectors. However, the receiver has *a priori* knowledge that the transmitted signals are real. Moreover, the imaginary part may cause additional interference. To impose a real constraint on (4) and (5), we relax $\mathcal{Q}$ to $\mathbb{R}$. Note that the complex system (1) can be transformed into a real system as follows:

$$\tilde{\mathbf{r}} = \begin{bmatrix} \mathrm{Re}\{\mathbf{r}\} \\ \mathrm{Im}\{\mathbf{r}\} \end{bmatrix} = \begin{bmatrix} \mathrm{Re}\{\mathbf{H}\} \\ \mathrm{Im}\{\mathbf{H}\} \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathrm{Re}\{\mathbf{n}\} \\ \mathrm{Im}\{\mathbf{n}\} \end{bmatrix} = \tilde{\mathbf{H}}\mathbf{x} + \tilde{\mathbf{n}}. \tag{6}$$

Note that the entries of $\tilde{\mathbf{n}}$ have zero means and variance $\sigma_n^2/2$. The ZF and MMSE linear detectors for the equivalent real system (6) can be obtained as

$$\hat{\mathbf{x}}_{\text{R-ZF}} = D\left[(\tilde{\mathbf{H}}^H\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^H\tilde{\mathbf{r}}\right] \tag{7}$$

and

$$\hat{\mathbf{x}}_{\text{R-MMSE}} = D\left[\left(\tilde{\mathbf{H}}^H\tilde{\mathbf{H}} + \sigma_n^2/2\mathbf{I}_n\right)^{-1}\tilde{\mathbf{H}}^H\tilde{\mathbf{r}}\right] \tag{8}$$

where R-ZF and R-MMSE denote the real constrained ZF and MMSE detectors, respectively. Since $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{r}}$ are real, $\hat{\mathbf{x}}_{\text{R-ZF}}$ and $\hat{\mathbf{x}}_{\text{R-MMSE}}$ are also real. Therefore, the real constraint is implicitly imposed.

Note that as the solution obtained by either (5) or (8) has a bias toward zero, the prefilter output $\hat{\mathbf{x}}$ should be scaled to maintain the average constellation power before applying the threshold decision. By exploiting the power constraint or the modulus constraint of each constellation, the performance of the MMSE receivers can be improved, and $\hat{\mathbf{x}}$ needs not to be scaled [17].

C. Modulus Constrained Subgroup Detectors

When $\mathcal{Q}$ is complex, we can exploit the modulus constraints of $\mathcal{Q}$. We first consider a unitary constellation with unity modulus $|x_i|^2 = 1$, i.e., M-PSK. This pointwise constraint directly leads to the candidate vectors being in the hypersphere $\mathbf{x}^H\mathbf{x} = n$, e.g., the CLS [8]. However, to achieve better performance, tighter constraints are required. We, thus, partition the vector $\mathbf{x}$ into $g > 1$ groups, and each group forms a subvector $\mathbf{x}_i$ with size $s_i, i = 1, \ldots, g$, where $\sum_{i=1}^{g} s_i = n$. We relax each $\mathbf{x}_i$ on an s_i-dimensional hypersphere $\mathbf{x}_i^H\mathbf{x}_i = s_i$. The constrained MLD is, thus, given by

$$\hat{\mathbf{x}}_{\text{CML}} = D\left[\mathop{\arg\min}_{\mathbf{x}_1^H\mathbf{x}_1 = s_1, \ldots, \mathbf{x}_g^H\mathbf{x}_g = s_g} \|\mathbf{r} - \mathbf{H}\mathbf{x}\|^2\right] \tag{9}$$

where CML denotes the constrained MLD.[1] The minimization problem in (9) can be written as

$$\begin{aligned}&\min_{\mathbf{x}} \|\mathbf{r} - \mathbf{H}\mathbf{x}\|^2\\ &\text{s.t. } \mathbf{x}_1^H\mathbf{x}_1 = s_1, \ldots, \mathbf{x}_g^H\mathbf{x}_g = s_g.\end{aligned} \tag{10}$$

The Lagrangian $\mathcal{L}(\mathbf{x}, \lambda_1, \ldots, \lambda_g)$ for this minimization problem is

$$\mathcal{L}(\mathbf{x}, \lambda_1, \ldots, \lambda_g) = \|\mathbf{r} - \mathbf{H}\mathbf{x}\|^2 + \sum_{i=1}^{g} \lambda_i\left(\mathbf{x}_i^H\mathbf{x}_i - s_i\right). \tag{11}$$

By taking partial derivatives with respect to $\mathbf{x}$, the solution for $\mathbf{x}$ can be derived as

$$\hat{\mathbf{x}}(\lambda_1, \ldots, \lambda_g) = (\mathbf{H}^H\mathbf{H} + \mathbf{\Lambda})^{-1}\mathbf{H}^H\mathbf{r} \tag{12}$$

where $\mathbf{\Lambda}$ is a diagonal matrix and is given by

$$\mathbf{\Lambda} = \text{diag}\{\underbrace{\lambda_1, \ldots, \lambda_1}_{s_1}, \ldots, \underbrace{\lambda_g, \ldots, \lambda_g}_{s_g}\}. \tag{13}$$

Note that (12) is a minimizer of (11) only when $\mathbf{H}^H\mathbf{H} + \mathbf{\Lambda}$ is semidefinite. When $g = 1$, there is only one λ_1, and (13) reduces to the CLS solution in [8]. When $\lambda_1 = \cdots = \lambda_g = \sigma_n^2$, the CML detector reduces to the MMSE detector (5). Compared with the CLS detector [8], our new relaxation is tighter, and the continuous space is smaller. Note that (12) reduces to the ZF linear detector if $\mathbf{\Lambda} = \mathbf{0}$.

To obtain the CML solution in (12), the optimal values for $\lambda_1, \ldots, \lambda_g$ have to be computed so that the unitary constraints are fulfilled. Substituting $\hat{\mathbf{x}}(\lambda_1, \ldots, \lambda_g)$ into (10), we need the zeros of the set of equations, i.e.,

$$\begin{aligned}F_1(\lambda_1, \ldots, \lambda_g) &= \|\hat{\mathbf{x}}_1(\lambda_1, \ldots, \lambda_g)\|^2 - s_1 = 0\\ &\vdots\\ F_g(\lambda_1, \ldots, \lambda_g) &= \|\hat{\mathbf{x}}_g(\lambda_1, \ldots, \lambda_g)\|^2 - s_g = 0.\end{aligned} \tag{14}$$

However, the solution of (14) does not necessarily make $\mathbf{H}^H\mathbf{H} + \mathbf{\Lambda}$ semidefinite. Therefore, solving (14) does not guarantee the optimal solution of (10).

The multidimensional Newton–Raphson root finding method [19] can be used to solve (14). This method needs the partial derivative of F_i with respect to λ_j, $\partial F_i/\partial\lambda_j$, $1 \le i$, $j \le g$. For simplicity, we show only the differentiation of F_1 with respect to λ_1 and F_g with respect to λ_1. $\partial F_i/\partial\lambda_i$ can be obtained by permuting the columns of $\mathbf{H}$ such that $\mathbf{x}_i$ corresponds to the first s_i entries of $\mathbf{x}$. $\partial F_i/\partial\lambda_j$, $j \ne i$, can be obtained by permuting the columns of $\mathbf{H}$ such that $\mathbf{x}_i$ corresponds to the last s_i entries of $\mathbf{x}$, and $\mathbf{x}_j$ corresponds to the first s_j entries of $\mathbf{x}$. We can obtain $\partial F_1/\partial\lambda_1$ as given in (15), shown at the bottom of the page, where $\mathbf{A} = \mathbf{H}_1^H\mathbf{H}_1 + \lambda_1\mathbf{I}$, $\mathbf{B} = \mathbf{H}_1^H\mathbf{H}_2$, $\mathbf{C} = \mathbf{H}_2^H\mathbf{H}_2 + \Lambda_2$, $\Lambda_2 = \text{diag}(\lambda_2, \ldots, \lambda_2, \ldots, \lambda_g, \ldots, \lambda_g)$, $\mathbf{Q} = \mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^H$, $\mathbf{H}_1$ corresponds to the first s_1 columns of $\mathbf{H}$, and $\mathbf{H}_2$ corresponds to the last $n - s_1$ columns of $\mathbf{H}$. We also have

$$\frac{\partial F_g(\lambda_1, \cdots, \lambda_g)}{\partial\lambda_1} = \mathbf{r}^H\mathbf{H}\begin{pmatrix}\mathbf{\Phi} & \mathbf{\Psi}\\ \mathbf{\Omega} & \mathbf{\Xi}\end{pmatrix}\mathbf{H}^H\mathbf{r} \tag{16}$$

$$\frac{\partial F_1(\lambda_1, \ldots, \lambda_g)}{\partial\lambda_1} = \mathbf{r}^H\mathbf{H}\begin{pmatrix}-2\mathbf{Q}^{-3} & \mathbf{Q}^{-2}\mathbf{B}\mathbf{C}^{-1}\mathbf{Q}^{-1} + \mathbf{Q}^{-1}\mathbf{B}\mathbf{C}^{-1}\mathbf{Q}^{-2}\\ \mathbf{Q}^{-2}\mathbf{C}^{-1}\mathbf{B}^H\mathbf{Q}^{-1} + \mathbf{Q}^{-1}\mathbf{C}^{-1}\mathbf{B}^H\mathbf{Q}^{-2} & -\mathbf{Q}^{-2}\mathbf{B}\mathbf{C}^{-2}\mathbf{B}^H\mathbf{Q}^{-1} - \mathbf{Q}^{-1}\mathbf{B}\mathbf{C}^{-2}\mathbf{B}^H\mathbf{Q}^{-2}\end{pmatrix}\mathbf{H}^H\mathbf{r} \tag{15}$$

[1] Note that the CML is not the true maximum likelihood even if (9) is exactly solved.

where

$$\mathbf{\Phi} = -\mathbf{E}^H\mathbf{D} - \mathbf{D}^H\mathbf{E} \tag{17a}$$

$$\mathbf{\Psi} = -\mathbf{E}^H\mathbf{C}^{-1} - \mathbf{E}^H\mathbf{C}^{-1}\mathbf{B}^H\mathbf{D}^H - \mathbf{D}^H\mathbf{C}^{-1}\mathbf{B}^H\mathbf{E}^H \tag{17b}$$

$$\mathbf{\Omega} = -\mathbf{C}^{-1}\mathbf{E}^H - \mathbf{DBC}^{-1}\mathbf{E} - \mathbf{EBC}^{-1}\mathbf{D} \tag{17c}$$

$$\mathbf{\Xi} = -\mathbf{C}^{-2}\mathbf{B}^H\mathbf{E}^H - \mathbf{EBC}^{-2} - \mathbf{EBC}^{-2}\mathbf{B}^H\mathbf{D}^H \tag{17d}$$

$$-\mathbf{DBC}^{-2}\mathbf{B}^H\mathbf{E}^H \tag{17e}$$

$\mathbf{A} = \mathbf{H}_1^H\mathbf{H}_1 + \Lambda_1\mathbf{I}$, $\mathbf{B} = \mathbf{H}_1^H\mathbf{H}_2$, $\mathbf{C} = \mathbf{H}_2^H\mathbf{H}_2 + \lambda_g\mathbf{I}$, $\Lambda_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_1, \ldots, \lambda_{g-1}, \ldots, \lambda_{g-1})$, and $\mathbf{Q} = \mathbf{A} - \mathbf{BC}^{-1}\mathbf{B}^H$. $\mathbf{H}_1$ corresponds to the first $n - s_g$ columns of $\mathbf{H}$, and $\mathbf{H}_2$ corresponds to the last s_g columns of $\mathbf{H}$. $\mathbf{\Delta} = \mathrm{diag}(\underbrace{1, \ldots, 1}_{s_1}, \ldots, 0, \ldots, 0)$, $\mathbf{E} = \mathbf{C}^{-1}\mathbf{B}^H\mathbf{Q}^{-1}\mathbf{\Delta}\mathbf{Q}^{-1}$, and $\mathbf{D} = \mathbf{C}^{-1}\mathbf{B}^H\mathbf{Q}^{-1}$.

Since system (14) has multiple roots, an initial estimate is needed to guarantee the convergence to the desired root. In the CLS case (a 1-D case where only λ_1 exists), from [20], it can be shown that the global minimum is achieved by using the maximal real root λ_1^*, which can be found by using the 1-D Newton method. However, in the multidimensional case, no such theorem exists that specifies the root that minimizes (10). There are two possible initial estimates for λ_i's. First, the initial values may be chosen as $\lambda_1 = \cdots = \lambda_g = \sigma_n^2$ [this choice is inspired by the fact that the performance of the MMSE detector (5) in the low SNR is dominated by the noise variance]. Second, we solve the CLS first and use the solution λ_1^* for $g = 1$ as the initial estimate. If the Newton method does not converge after a specified number of iterations, we simply set $\lambda_1 = \cdots = \lambda_g = \sigma_n^2$ or $\lambda_1 = \cdots = \lambda_g = \lambda_1^*$. Our simulation results show that the probability that the Newton method for (14) does not converge increases with the increase in the number of groups g. Since the Newton method returns the root close to the initial estimate, the global minimum of (10) may not be achieved. Therefore, our approach is suboptimal for solving (10). However, based on our simulation results, we find that if the CLS solution is used as the initial estimate, our detector always performs better than the original CLS detector.

For a nonunitary constellation such as QAM, we assume $\rho_{\max}$ and $\rho_{\min}$ as the largest and smallest moduli of the constellation, respectively. As before, we partition the vector $\mathbf{x}$ into g groups and use the constraint $\rho_{\max}$. We, thus, relax each $\mathbf{x}_i$ in an s_i-dimensional hypersphere $\mathbf{x}_i^H\mathbf{x}_i \leq \rho_{\max}^2 s_i$. The CML detector is modified as

$$\hat{\mathbf{x}}_{\mathrm{CML}} = D\left[\underset{\mathbf{x}_1^H\mathbf{x}_1 \leq \rho_{\max}^2 s_1, \ldots, \mathbf{x}_g^H\mathbf{x}_g \leq \rho_{\max}^2 s_g}{\arg\min} \|\mathbf{r} - \mathbf{Hx}\|^2\right]. \tag{18}$$

The Lagrangian function for the minimization problem in (18) can be expressed as

$$\mathcal{L}(\mathbf{x}, \lambda_1, \ldots, \lambda_g) = \|\mathbf{r} - \mathbf{Hx}\|^2 + \sum_{i=1}^{g} \lambda_i \left(\mathbf{x}_i^H\mathbf{x}_i - \rho_{\max}^2 s_i\right) \tag{19}$$

where λ_i is the Lagrangian multiplier associated with the ith inequality constraint, and $\lambda_i \geq 0$. The Lagrange dual function is the minimum value of the Lagrangian (19) over $\mathbf{x}$, and

$$g(\lambda_1, \ldots, \lambda_g) = \inf_{\mathbf{x} \in \mathbb{C}^n} \mathcal{L}(\mathbf{x}, \lambda_1, \ldots, \lambda_g). \tag{20}$$

The minimization of (19) for $\mathbf{x}$ has the same solution as (13). Substituting it back to (19), we obtain

$$g(\lambda_1, \ldots, \lambda_g) = -\mathbf{r}^H\mathbf{H}(\mathbf{H}^H\mathbf{H} + \mathbf{\Lambda})^{-1}\mathbf{H}^H\mathbf{r} - \rho_{\max}^2 \sum_{i=1}^{g} \lambda_i s_i, \quad \lambda_i \geq 0. \tag{21}$$

It can be readily verified that the objective function and the constraints are convex. There exists a strictly feasible point. Therefore, the constraints meet Slater's condition, and strong duality holds for (18) [21]. The maximum value of $g(\lambda_1, \ldots, \lambda_g)$ is equal to the minimum of (18). We solve $\lambda_1, \ldots, \lambda_g$ by maximizing (21) first and substituting them back into (12) to obtain the solution to (18). In (21), the set $S = \{[\lambda_1, \ldots, \lambda_i] | \lambda_i \geq 0, i = 1, \ldots, g\}$ is convex. A g-dimensional subgradient algorithm [22] can, thus, be used to solve (21). For simplicity, we show only the differentiation of $g(\lambda_1, \ldots, \lambda_g)$ with respect to λ_1. $\partial g / \partial \lambda_i$, $i > 1$ can be obtained by permuting the columns of $\mathbf{H}$ such that $\mathbf{x}_i$ corresponds to the first s_i entries of $\mathbf{x}$. We can obtain

$$\frac{g(\partial\lambda_1, \ldots, \lambda_g)}{\partial\lambda_1} = \mathbf{r}^H\mathbf{H}\begin{pmatrix} \mathbf{Q}^{-2} & -\mathbf{Q}^{-2}\mathbf{BC}^{-1} \\ -\mathbf{C}^{-1}\mathbf{B}^H\mathbf{Q}^{-2} & \mathbf{C}^{-1}\mathbf{B}^H\mathbf{Q}^{-2}\mathbf{BC}^{-1} \end{pmatrix} \times \mathbf{H}^H\mathbf{r} - \rho_{\max}^2 s_1 \tag{22}$$

where $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{Q}$ are defined in (15). The gradient descent algorithm starts at $\lambda_1 = \lambda_2 = \cdots = \lambda_g = \sigma_n^2$. With a diminishing stepsize, the gradient descent algorithm converges to the optimal solution [22]. If $g = 1$, the CML detector (18) reduces to the GMMSE in [7]. Therefore, the CML detector generalizes the GMMSE.

For tighter constraints and better performance, $\rho_{\min}$ can also be considered by posing another g constraints $\rho_{\min}^2 s_i \leq \mathbf{x}_i^H\mathbf{x}_i$, for $i = 1, \ldots, g$, on (18). However, the resulting nonconvex optimization problem, in general, is hard to solve.

Remarks:

1) The constrained detectors that use constellation modulus information can be combined with the real constraint in Section III-B. For real constellations, the CML detectors (9) and (18) can be directly applied to the equivalent system (6) by taking into account the real and modulus constraints. We denote the combined receiver as R-CML.
2) The proposed approach can also be extended to the MMSE linear detectors. Let the prefilter matrix in the MMSE be $\mathbf{G}$ and the prefilter output be $\hat{\mathbf{x}} = \mathbf{Gr}$. Denote $\mathcal{P}$ as the average constellation power. The constrained MMSE can be obtained by solving

$$\min_{\mathbf{G}} E\left\{\|\mathbf{x} - \hat{\mathbf{x}}\|^2\right\} \text{ s.t. } E\left\{\hat{\mathbf{x}}_1^H\hat{\mathbf{x}}_1\right\} = s_1\mathcal{P}, \ldots \quad E\left\{\hat{\mathbf{x}}_g^H\hat{\mathbf{x}}_g\right\} = s_g\mathcal{P}. \tag{23}$$

We omit the details of solving (23) here.

D. *Coordinate Ascent Improvement*

Although the proposed detectors perform worse than the MLD (see the simulation results in Section V), the number of symbol errors in them in a high SNR is rather small (i.e., $\tilde{x}_k = x_k$ for most k, and for erroneous decisions, $\tilde{x}_k \neq x_k$). This fact suggests that an iterative approach may improve the performance of our constrained detectors. This detector is a block version of a coordinate ascent algorithm [19]. As we did before with the CML detectors, we partition $\mathbf{x}$ into g groups, each with s_i symbols. In each iteration, for $i = 1, \ldots, g$, we minimize one group by fixing the other $g-1$ groups. The algorithm can be summarized in the following two steps.

- Initialization: Set iteration number $k = 0$, and obtain the initial data detection by using a suboptimal detector. This initial solution is denoted as $\hat{\mathbf{x}}^{(0)}$.
- Iteration: $k = k+1$, and set iteration number $i = 1$. For $i = 1, \ldots, g$, compute

$$\mathbf{r}^{(i)} = \mathbf{r} - \mathbf{H}_{\bar{i}}\hat{\mathbf{x}}^{(k-1)} \tag{24}$$

where $\mathbf{H}_{\bar{i}}$ is formed by zeroing the columns from $f_i = \sum_{j=1}^{i-1} s_j + 1$ to $e_i = \sum_{j=1}^{i} s_j$. The data vector $\mathbf{x}_i = [x_{f_i}, \ldots, x_{e_i}]^T$ is detected by using

$$\hat{\mathbf{x}}_i = \arg\min_{\mathbf{x}_i \in \mathcal{Q}^{s_i}} \left\| \mathbf{r}^{(i)} - \mathbf{H}_i\mathbf{x}_i \right\|^2 \tag{25}$$

where $\mathbf{H}_i$ is formed by the columns from f_i to e_i. $\hat{\mathbf{x}}^{(k)} = [\hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_g]^T$, and the iteration continues until $\hat{\mathbf{x}}^{(k)} = \hat{\mathbf{x}}^{(k-1)}$.

Remarks:

1) The variation of the number of groups g from 1 to n results in different performance levels. In particular, if the number of groups is one $(g = 1)$, (25) reduces to the MLD problem (2). When $g = n$, the iterative improvement algorithm is similar to the parallel interference cancellation (PIC) that is used in CDMA systems. However, the iterative algorithm for MIMO with $g = n$ performs worse than the PIC does in CDMA for the following reason. In the CDMA case, the nondiagonal terms in $\mathbf{H}$ due to the nonorthogonality of the spreading codes are typically small; however, this condition does not hold for the MIMO channel $\mathbf{H}$. Our new iterative improvement, however, generalizes the PIC detector.
2) When the number of symbols in the tth group (s_i) is small, an exhaustive search or the SD can solve (25). In any case, the worst-case complexity of our new iterative algorithm is $O(K\sum_{i=1}^{g}|Q|^{s_i})$, where K is the total number of iterations. The complexity is between those of the SD and the PIC. The choice of g and s_i depends on many factors, such as the type of the suboptimal detector, the desired bit error rate (BER), and the complexity. For a practical system design, the number of groups and their sizes should be empirically chosen to achieve a good performance-complexity tradeoff at a given SNR.
3) For soft decoding of linear block codes, Chase [23] has proposed a class of suboptimal decoders. These have been adapted for MIMO detection [24], [25].
4) In general, the reliability of the first few detected symbols is less than that of the later detected symbols [14], [26], and the overall error rate is highly affected by the reliability of the first stage. Therefore, it makes sense to choose $s_1 \leq s_2 \leq \cdots \leq s_g$.

IV. CONSTRAINED DFDs

A. *V-BLAST Detection*

The V-BLAST detection algorithm [1] relies on nulling and interference cancellation. The nulling step uses the ZF or the MMSE criterion. The interference of previously detected symbols is subtracted. Nulling and interference cancellation improve the overall performance when the order of detection is carefully chosen. For instance, in the kth iteration, the symbol with the maximum postdetection SNR among the remaining $n-k+1$ symbols is detected. This ordering scheme is known to be the optimal detection order. The whole algorithm is described as follows.

- Initialization:

$$\mathbf{r}_1 = \mathbf{r} \tag{26a}$$

$$\mathbf{G}_1 = \mathbf{H}^{\dagger} \tag{26b}$$

$$k_1 = \arg\min_{j} \|(\mathbf{G}_1)_j\|^2 \tag{26c}$$

- Recursion: for $i = 1$ to n

$$\mathbf{w}_{k_i} = (\mathbf{G}_i)_{k_i} \tag{26d}$$

$$\hat{x}_{k_i} = \arg\min_{x \in \mathcal{Q}} \left| x - \mathbf{w}_{k_i}^H \mathbf{r}_i \right|^2 \tag{26e}$$

$$\mathbf{r}_{i+1} = \mathbf{r}_i - \hat{x}_{k_i}(\mathbf{H})_{k_i} \tag{26f}$$

$$\mathbf{G}_{i+1} = \mathbf{H}^{\dagger}_{\bar{k}_i} \tag{26g}$$

$$k_{i+1} = \arg\min_{j \notin \{k_1, \ldots, k_i\}} \|(\mathbf{G}_{i+1})_j\|^2 \tag{26h}$$

where $(\mathbf{A})_i$ is the ith column of matrix $\mathbf{A}$, and $\mathbf{H}_{\bar{k}_i}$ is obtained by zeroing the $k_1, \ldots, k_i$th columns of $\mathbf{H}$.

As suggested in [2], given an optimum order $k_1, \ldots, k_n$, V-BLAST detection is equivalent to the zero-forcing decision-feedback detection (ZF-DFD). Assuming $\mathbf{\Pi}$ is the column permutation matrix obtained from the optimum order, we apply $\mathbf{\Pi}$ to $\mathbf{H}$.[2] Let the QR factorization of $\tilde{\mathbf{H}} = \mathbf{H}\mathbf{\Pi}$ be $\mathbf{QR}$, where $\mathbf{Q}$ is a unitary matrix, and $\mathbf{R}$ is an upper triangular one. Equation (1) is equivalent to

$$\mathbf{y} = \mathbf{R}\mathbf{x} + \mathbf{v} \tag{27}$$

where $\mathbf{y} = \mathbf{Q}^H\mathbf{r}$, and $\mathbf{v} = \mathbf{Q}^H\mathbf{n}$ is a noise vector whose entries are i.i.d. complex Gaussian with mean zero and variance σ_n^2.

[2] As in [2], the filtering matrices in constrained detectors and the corresponding constrained ordering can be similarly applied.

The second description of the V-BLAST algorithm is given as follows:

- for $i = n$ to 1

$$\hat{x}_i = \arg\min_{x \in \mathcal{Q}} |y_i - R_{i,i}x|^2 \tag{28a}$$

$$\mathbf{y} = \mathbf{y} - (\mathbf{R})_i \hat{x}_i \tag{28b}$$

- end where $R_{i,i}$ is the (i,i)th entry of $\mathbf{R}$, and $(\mathbf{R})_i$ is the ith column of $\mathbf{R}$.

B. Real DFDs

For real-valued constellations, we perform the V-BLAST algorithm (26a)–(26c) and (26d)–(26h) on the real system (6), which automatically takes the real constraint into account. We denote the V-BLAST for (6) as R-V-BLAST. By using the same methods as those used in Section III-B, R-V-BLAST was found to perform better than the original V-BLAST by suppressing the imaginary interference. More precisely, if $n = m$, and no permutations are used, the squared norm of the entries of $\mathbf{R}$ is known to be χ^2 distributed [27], specifically, $|R_{i,i}|^2 \sim \chi^2(2i)$, for $i = 1, \ldots, n$, and $|R_{i,j}|^2 \sim \chi^2(2)$, for $j > i$, where $\chi^2(k)$ denotes the χ^2 distribution with k degrees of freedom. Since the performance of V-BLAST is limited by the first detected symbol [14], the diversity order of V-BLAST detection is only one [28], [29]. However, if QR decomposition is performed on the $2n \times n$ real matrix $\tilde{\mathbf{H}}$ in (6), we first construct a $2n \times 2n$ real matrix $\mathbf{H}_1$ with each entry zero mean and variance 1, and the first n columns are equal to $\tilde{\mathbf{H}}$. Let the QR decomposition of $\tilde{\mathbf{H}}$ and $\mathbf{H}_1$ be $\tilde{\mathbf{H}} = \tilde{\mathbf{Q}}\tilde{\mathbf{R}}$ and $\mathbf{H}_1 = \mathbf{Q}_1\mathbf{R}_1$, respectively, and $|R_{i,i}|^2 \sim \chi^2(i)$, for $i = 1, \ldots, 2n$, and $|R_{i,j}|^2 \sim \chi^2(1)$, for $j > i$. We have $\mathbf{R}_2 = \tilde{\mathbf{H}}\mathbf{Q}_1^H$, which consists of the first n columns of $\mathbf{R}_1$. Therefore, the squared norms of the entries of $\tilde{\mathbf{R}}$ are also χ^2 distributed; however, $|\tilde{R}_{i,i}|^2 \sim \chi^2(i+n)$, for $i = 1, \ldots, n$, and $|\tilde{R}_{i,j}|^2 \sim \chi^2(1)$, for $j > i$. Therefore, by using the analysis approach in [14] and [26], it can be readily verified that the diversity order of R-V-BLAST increases to $(n+1)/2$. For real constellations, if the decoupled system is used with a real constraint (6), the diversity order increases from 1 to $(n+1)/2$. The result is a significant performance gain over the original V-BLAST. This also shows a diversity rate tradeoff.

For decoupleable complex constellations such as QAM, (1) can be rewritten as

$$\begin{bmatrix} \mathrm{Re}\{\mathbf{r}\} \\ \mathrm{Im}\{\mathbf{r}\} \end{bmatrix} = \begin{bmatrix} \mathrm{Re}\{\mathbf{H}\} & -\mathrm{Im}\{\mathbf{H}\} \\ \mathrm{Im}\{\mathbf{H}\} & \mathrm{Re}\{\mathbf{H}\} \end{bmatrix} \begin{bmatrix} \mathrm{Re}\{\mathbf{x}\} \\ \mathrm{Im}\{\mathbf{x}\} \end{bmatrix} + \begin{bmatrix} \mathrm{Re}\{\mathbf{n}\} \\ \mathrm{Im}\{\mathbf{n}\} \end{bmatrix} \tag{29}$$

or

$$\tilde{\mathbf{r}} = \tilde{\mathbf{H}}\tilde{\mathbf{x}} + \tilde{\mathbf{n}}. \tag{30}$$

In [30], it has been shown that applying V-BLAST to the equivalent real system (29) yields an additional performance gain. However, the diversity order is not increased. These findings suggest that performing the V-BLAST algorithm on the equivalent real system is always beneficial if the constellation is decoupleable or real.

C. CODFDs

The ZF nulling vector $\mathbf{w}_{k_i}$ [see (26d)] in V-BLAST completely removes the interference from the other antennas and also amplifies the additive noise. Instead of using $\mathbf{w}_i$ in (26d) to completely remove the interference, we use an incomplete nulling vector for a better tradeoff between noise enhancement and interference suppression. We propose using the filtering matrix in our proposed constrained detectors in Section III as the nulling vector instead of the ZF nulling vector in V-BLAST. We replace (26b) and (26g) with

$$\mathbf{G}_1 = (\mathbf{H}^H\mathbf{H} + \mathbf{\Lambda})^{-1}\mathbf{H}^H \tag{31}$$

and

$$\mathbf{G}_{i+1} = \left(\mathbf{H}_{\bar{k}_i}^H \mathbf{H}_{\bar{k}_i} + \mathbf{\Lambda}_i\right)^{-1} \mathbf{H}_{\bar{k}_i}^H \tag{32}$$

where $\mathbf{\Lambda}$ and $\mathbf{\Lambda}_i$ can be calculated by using (14) and (21) for constant unitary and nonunitary constellations, respectively.

Since interference cannot be completely removed when nulling is performed by using the CML, we propose to determine the detection order at each iteration by maximizing the SINR, which is defined as

$$\mathrm{SINR}_j = \frac{\left|(\mathbf{G}_{i+1}\mathbf{H}_{\bar{k}_i})_{j,j}\right|^2 E\left\{|x_j|^2\right\}}{\sum_{k=1,k\neq j}^{n} \left|(\mathbf{G}_{i+1}\mathbf{H}_{\bar{k}_i})_{j,k}\right|^2 E\left\{|x_k|^2\right\} + \sigma_n^2 \left\|(\mathbf{G}_{i+1})_j\right\|^2} \tag{33}$$

where $(\mathbf{A})_{i,j}$ is the (i,j)th entry of matrix $\mathbf{A}$, and $(\mathbf{G}_{i+1})_j$ denotes the jth row of matrix $\mathbf{G}_{i+1}$. In the V-BLAST detection algorithm, (26h) is replaced by

$$k_{i+1} = \arg\max_{j \notin \{k_1, \ldots, k_i\}} \mathrm{SINR}_j. \tag{34}$$

This modified V-BLAST detection is denoted as the constrained ordering DFD (CODFD).

Note that if $\mathbf{\Lambda} = \sigma_n^2\mathbf{I}_n$, the CODFD reduces to the MMSE decision-feedback detector (MMSE-DFD) in [31]. If $\mathbf{\Lambda} = \mathbf{0}$, our CODFD becomes the original V-BLAST. Different ordering schemes with CML detectors and MMSE detectors can be combined to form hybrid ordering schemes. From [14] and [26], the diversity order of the first few detected symbols is less than that of the later detected symbols. Since the CML detector with $g = n$ performs better than the other CML detectors and MMSE detectors, but with higher complexity, we can perform ordering with this detector in the first k symbols for a better tradeoff between noise enhancement and interference suppression. In the last $n - k$ stages, the original V-BLAST or the MMSE-DFD ordering with low complexity can be applied since the diversity order in these stages is high. This hybrid ordering scheme gives a tradeoff between complexity and performance.

D. Combined Constrained Detectors and DFDs

The performance of the ZF-DFDs, or the equivalent V-BLAST, is limited by the error propagation of decision feedback. Even with the V-BLAST optimal ordering, the diversity order of V-BLAST detection is just one [28], [29] because V-BLAST is a greedy algorithm. That is, a hard decision is based only on the "local" metric (28a) without taking the subsequent symbol decisions into account. We, thus, combine the constrained detectors in Section III and the ZF-DFDs to make hard decisions less greedily. At each iteration, a "global" metric, which is obtained by using the constrained detectors, is used to make a decision on each symbol.

In the ith iteration, we define $\mathbf{R}_i = \mathbf{R}(1:i-1, 1:i-1)$, $\mathbf{r}_i = \mathbf{R}(1:i-1, i)$, and $\mathbf{y}_i = \mathbf{y}(1:i-1)$. For each $x \in \mathcal{Q}$, after canceling x from $\mathbf{y}$, the soft decisions for the remaining $n-i$ symbols can be obtained by using the constrained detectors as

$$\hat{\mathbf{x}}_i = \left(\mathbf{R}_i^H \mathbf{R}_i + \mathbf{\Lambda}_i\right)^{-1} \mathbf{R}_i^H (\mathbf{y}_i - \mathbf{r}_i x) \tag{35}$$

where $\mathbf{x}_i = [x_1, \ldots, x_{i-1}]^T$, and $\mathbf{\Lambda}_i$ is defined in (12). Since the solution to (10) or (18) gives a low bound on $\|\mathbf{r} - \mathbf{H}\mathbf{x}\|^2$, the effect of x on the decision metric for the remaining $n-i$ symbols can be measured by using $\|\mathbf{y}_i - \mathbf{r}_i x - \mathbf{R}_i \hat{\mathbf{x}}_i\|^2$. The global metric for x is defined as

$$\begin{aligned} M_i(x) &= \|\mathbf{y}_i - \mathbf{r}_i x - \mathbf{R}_i \hat{\mathbf{x}}_i\|^2 + |y_i - R_{i,i} x|^2 \\ &= \left\| \left(\mathbf{I}_{n-i} - \mathbf{R}_i \left(\mathbf{R}_i^H \mathbf{R}_i + \mathbf{\Lambda}_i\right)^{-1} \mathbf{R}_i^H\right) (\mathbf{y}_i - \mathbf{r}_i x) \right\|^2 \\ &\quad + |y_i - R_{i,i} x|^2 \\ &= |a_i - b_i x|^2 \end{aligned} \tag{36}$$

where

$$\begin{aligned} a_i &= \sqrt{\left\| \left(\mathbf{I}_{n-i} - \mathbf{R}_i \left(\mathbf{R}_i^H \mathbf{R}_i + \mathbf{\Lambda}_i\right)^{-1} \mathbf{R}_i^H\right) \mathbf{r}_i \right\|^2 + |R_{i,i}|^2} \\ b_i &= \left(\mathbf{y}_i^H \left(\mathbf{I}_{n-i} - \mathbf{R}_i \left(\mathbf{R}_i^H \mathbf{R}_i + \mathbf{\Lambda}_i\right)^{-1} \mathbf{R}_i^H\right)^2 \mathbf{r}_i + y_i^* R_{i,i} \right) \Big/ a_i. \end{aligned} \tag{37}$$

In the ZF-DFDs, (28a) is simply replaced by

$$\hat{x}_i = D\left[\arg\min_{x \in \mathcal{Q}} M_i(x)\right]. \tag{38}$$

The resulting detector is denoted by CML-DFD. With a precomputed a_i and b_i, the total complexity of the CML-DFD is still $O(n^3)$.

Remarks:

1) If the CML detector is used, from the duality theory [21], $\|\mathbf{y}_i - \mathbf{r}_i x - \mathbf{R}_i \hat{\mathbf{x}}_i\|^2$ gives a lower bound and measures the effect of x on the remaining symbols.
2) If $\mathbf{\Lambda}_i = \sigma_n^2 \mathbf{I}_{n-i}$, (35) reduces to the MMSE. Although it does not give a lower bound on $\|\mathbf{y}_i - \mathbf{r}_i x - \mathbf{R}_i \hat{\mathbf{x}}_i\|^2$, the metric (36) also measures the effect of x on the overall metric. Therefore, the combined MMSE and DFD (CMMSE-DFD) enhances the performance.
3) The terms $\|\mathbf{y}_i - \mathbf{r}_i x - \mathbf{R}_i \hat{\mathbf{x}}_i\|^2$ and $|y_i - R_{i,i} x|^2$ are equally weighted in (36). However, we may differently weigh the two terms, and (36) can then be written as

$$M_i(x) = w_i \|\mathbf{y}_i - \mathbf{r}_i x - \mathbf{R}_i \hat{\mathbf{x}}_i\|^2 + (1 - w_i) |y_i - R_{i,i} x|^2 \tag{39}$$

where $0 \le w_i \le 1$ is the weight coefficient. If $w_i = 0.5$, (39) is equivalent to (36). If $w_i = 0$, (39) reduces to (28b), and the CML-DFD becomes the ZF-DFD. The coefficient w_i can be optimized by minimizing the MSE for x_i. In practice, w_i may be found by simulation.
4) When the channel rapidly varies, the coefficients a_i and b_i must be frequently updated, thereby increasing the average complexity of the detector. To alleviate this computation burden, we propose to use the global metric (36) to detect the first k symbols and the original V-BLAST local metric (28b) for the remaining $n-k$ symbols. The parameter k offers a complexity-performance tradeoff.

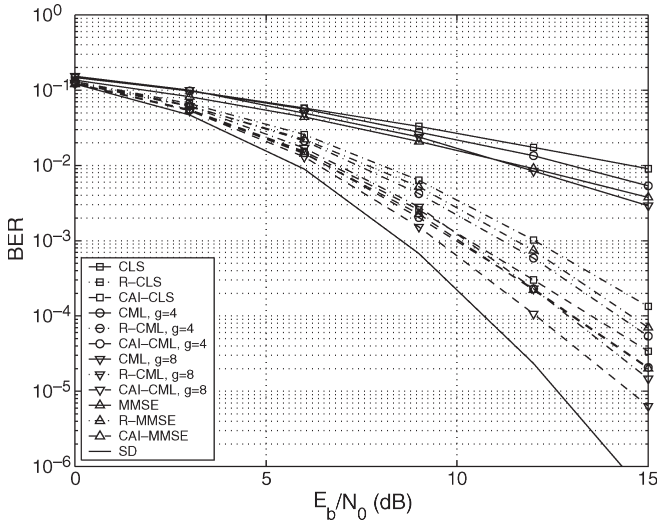


Fig. 1. Performance comparison of constrained detectors in an 8 × 8 MIMO system with BPSK.

V. Simulation Results

We test our proposed constrained detectors for a MIMO system with eight transmit and eight receive antennas over a flat Rayleigh fading channel. The receiver knows the channel state information and the noise variance. The notation CAI-X denotes the combination of the detector X and the coordinate ascent iterative correction described in Section III-D. The system is simulated by using MATLAB V7.0.4 on a workstation with an Intel Xeon processor at 3.2 GHz. The average CPU computation time is used as the measure of complexity. The SNR per bit is defined as

$$\frac{E_b}{N_0} = \frac{E\left\{\|\mathbf{H}\mathbf{x}\|^2\right\}}{m \log_2 |\mathcal{Q}| N_0} \tag{40}$$

where N_0 is the spectral noise density. The SD is implemented by using the algorithm in [4].

Fig. 1 shows the BER performance of different constrained detectors in a BPSK modulated system. We compare our detectors with the MLD and the CLS detector [8]. When all the detectors are applied to the complex system (1), the CLS and

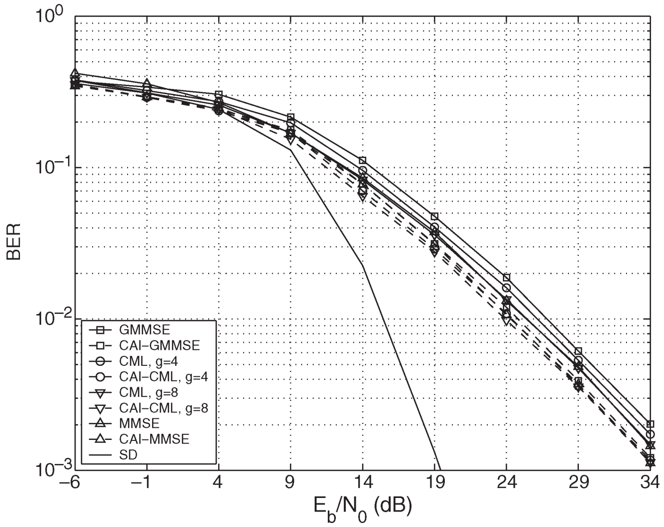


Fig. 2. Performance comparison of constrained detectors in an 8 × 8 MIMO system with 16QAM.

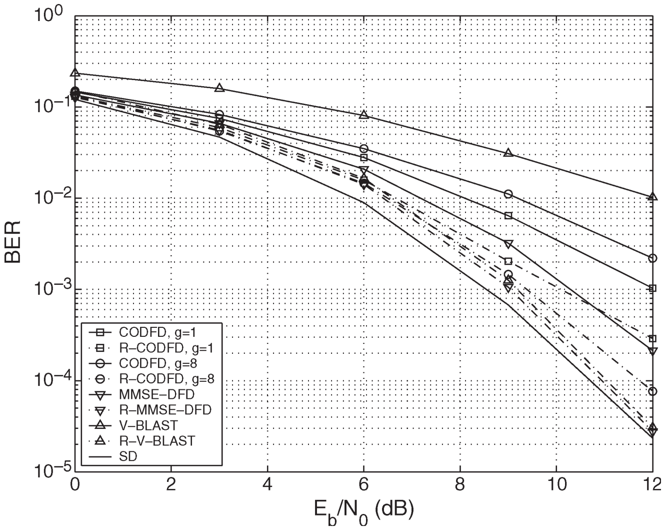


Fig. 3. Performance comparison of constrained ordering DFDs in an 8 × 8 MIMO system with BPSK.

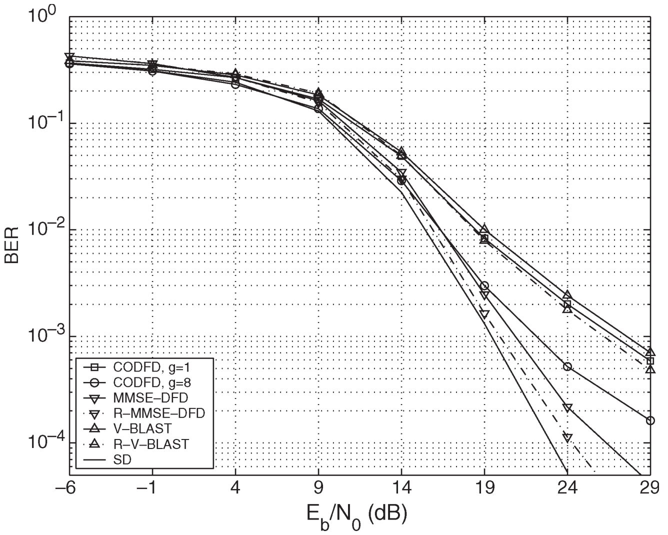


Fig. 4. Performance comparison of constrained ordering DFDs in an 8 × 8 MIMO system with 16QAM.

the CML perform close to the MMSE. In a high SNR, the CML with eight groups $(g = 8)$ performs better than the MMSE; however, all these detectors perform worse than the SD (the optimal detector). When they are applied to the real system (6), all the detectors perform better. At a BER of 10^{-3}, the R-MMSE has a 0.5-dB gain over the R-CLS. Both the R-CML with $g = 4$ and $g = 8$ perform better than the R-MMSE. They have 0.3- and 2-dB gain over the R-MMSE, respectively. After the iterative improvement is applied to all the detectors, the R-MMSE, the R-CLS, and the R-CML with four groups $(g = 4)$ have 2-, 1.8-, and 1.5-dB gains at a BER of 10^{-3}. The detector R-CML with $g = 8$ improves by 1 dB at a BER of 10^{-4}.

The BER of the GMMSE [7] and the different constrained detectors for 16QAM is shown in Fig. 2. The GMMSE performs worst among all the detectors. The CML with $g = 4$ has a 0.8-dB loss over the MMSE at a BER of 10^{-3}. In a low SNR, the CML with $g = 8$ performs better than the MMSE; however, they identically perform in a high SNR. With the coordinate ascent improvement, the R-MMSE, the R-CLS, the R-CML with $g = 4$, and the R-CML with $g = 8$ have 2-, 1.8-, 1-, and 1.2-dB gains at a BER of 10^{-2}, respectively. Since the groupwise hypersphere constraint (18) is loose, the resulting performance improvement is marginal. Tighter constraints are needed for high-order QAM constellations.

Fig. 3 compares the BER of a DFD and a real DFD with different constrained ordering schemes. BPSK modulation is used. The performance of V-BLAST and the SD is also shown in Fig. 3. We observe a dramatic performance improvement even for the constrained DFDs on the complex model. At a BER of 10^{-2}, the CODFD and the MMSE-DFD have more than 3-dB gain over V-BLAST. Therefore, the CODFD and the MMSE-DFD have smaller noise enhancement compared to the ZF-DFD. When the real constraint is imposed, R-V-BLAST and the R-MMSE-DFD perform close to the SD at a high SNR. They both perform only about 0.2-dB worse than the SD at a BER of 10^{-4}. The gap between the R-CODFD with $g = 8$ and R-V-BLAST is 0.7 dB at a BER of 10^{-4}. Since the diversity order of R-V-BLAST is $(n + 1)/2$, it performs well, and the performance improvement that is gained by using the R-MMSE-DFD is small.

Fig. 4 shows the BER of a DFD with different constrained ordering schemes for 16QAM. V-BLAST and the SD are used as benchmarks. The performance of R-V-BLAST that is achieved by using (30) is also presented. The CODFD with $g = 1$ has only a 0.7-dB gain over V-BLAST at a BER of 10^{-3}; however, it has a 0.8-dB loss over R-V-BLAST, perhaps due to the loose hypercube relaxation (18). When $g = 8$, the CODFD has a 2.7-dB loss over the SD at a BER of 10^{-3}. The MMSE-DFD performs better than the CODFD. At a BER of 10^{-3}, the MMSE-DFD has only a 1.5-dB loss over the SD. The gap reduces to 0.5 dB when the R-MMSE-DFD is used. Therefore, the R-MMSE-DFD is a preferable ordering scheme. However, all the order schemes achieve a diversity order of only one, which may be caused by the greedy nature of the DFD.

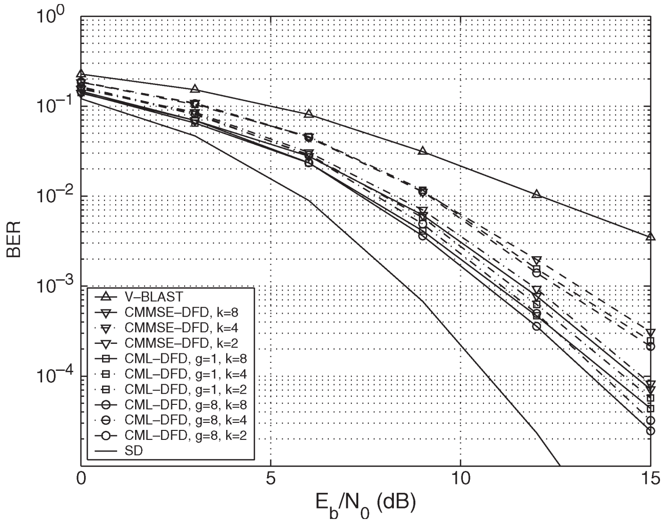


Fig. 5. Performance comparison of combined detectors and DFDs in an 8 × 8 BPSK MIMO system.

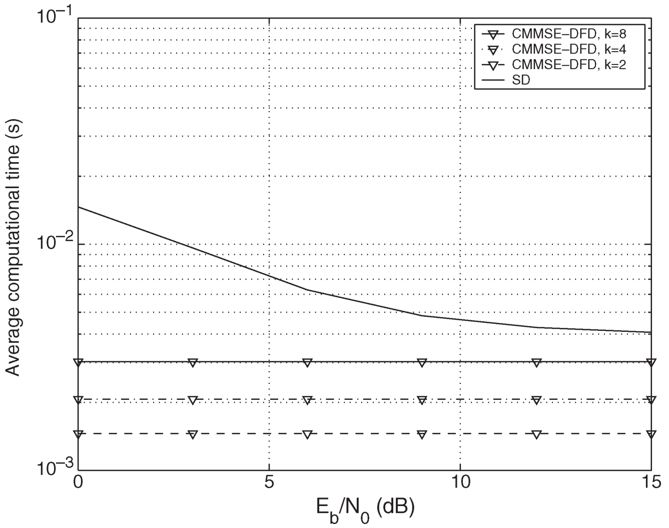


Fig. 6. Average computational time of combined constrained detectors and DFDs in an 8 × 8 BPSK MIMO system.

We present the results for the combined constrained detectors and DFDs for a BPSK system for a different k in Fig. 5, where k is the number of symbols that are detected by using the global metric (36), and the remaining symbols are detected by using the original V-BLAST local metric (28b). Our proposed CML-DFDs significantly improve the performance, indicating their ability to mitigate error propagation. The performance of all the combined detectors improves with k. At a BER of 10^{-2}, the CML-DFD with $g = 1$ and $k = 8$ has a 4-dB gain over V-BLAST. The CML-DFD performs better than the CMMSE-DFD with the same k. At a BER of 10^{-4}, the CML-DFD with $g = 8$ performs 0.5-dB better than the CMMSE-DFD when $k = 8$. The performance gain achieved by increasing k diminishes with the increase in k. For the CMMSE-DFD, a 1.2-dB gain is achieved by increasing k from 2 to 4 at a BER of 10^{-3}. However, the performance gain reduces to 0.2 dB when k increases from 4 to 8. The CML-DFD with $g = 8$ and $k = 8$ has the best performance among all the combined detectors, and it performs only 2.2-dB worse than the SD at a BER of 10^{-3}. Fig. 6 shows the average computational time of the SD and the CMMSE-DFD for a different k. The computational time of the SD does not include the preprocessing. For the CMMSE-DFD, we compute the coefficients a_i and b_i in each block, and this computation is included in the computational time. The CMMSE-DFD has constant complexity with the same k over all the SNRs, and its complexity increases with k. The CMMSE-DFD is faster than the SD in the observed SNR region. At SNR = 0 dB, the CMMSE-DFD with $k = 2$ is 13 times faster than the SD. In practice, the choice $k = n/2$ achieves a good performance-complexity tradeoff.

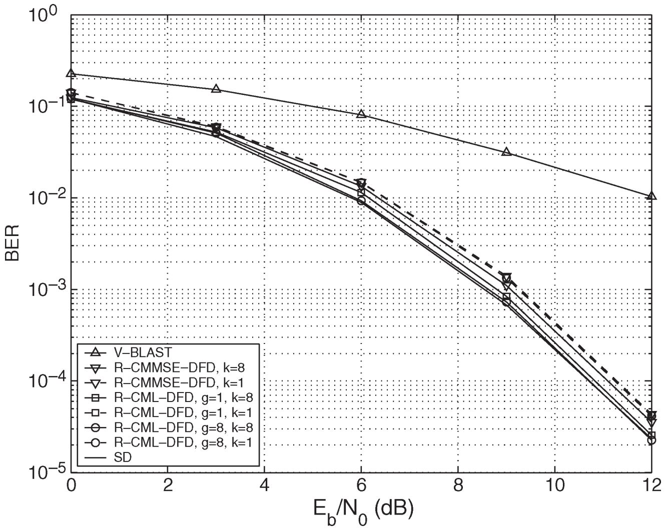


Fig. 7. Performance comparison of combined constrained detectors and DFDs in an 8 × 8 BPSK MIMO system using real constraints.

Fig. 7 shows the performance of the combined detectors for a BPSK system when the real constraint is applied. All the detectors perform close to the SD. The R-CMMSE-DFD with $k = 1$ has only a 0.6-dB loss over the SD at a BER of 10^{-4}. The R-CML-DFD with $g = 8$ and $k = 8$ almost achieves the maximum-likelihood performance. However, the performance gain that is achieved by increasing k decreases compared to the gain in the complex case in Fig. 7. Our combined detectors perform better than the V-BLAST detector. Fig. 8 shows the average computational time of the SD and the R-CMMSE-DFD. The R-CMMSE-DFD is less complex than the SD for a different SNR. Since only real operations are performed when solving the real system (29), the R-CMMSE-DFD is less complex than the corresponding CMMSE-DFD, whereas the former performs better for the same k. The R-CMMSE-DFD with $k = 1$ saves 16 times more complexity than the SD at SNR = 0 dB. When real constellations are used in practice, the R-CMMSE-DFD with $k = 1$ is enough to achieve good performance with low complexity.

Fig. 9 compares the combined constrained detectors and DFDs for a 16QAM system for different values of k. Although the performance of all the combined detectors improves by increasing k, the performance improvement is not as significant as that for a BPSK system. At a BER of 3×10^{-3}, the CMMSE-DFD with $k = 8$ has an about 4-dB gain over V-BLAST. The CML-DFD with $g = 1$ performs worse than the

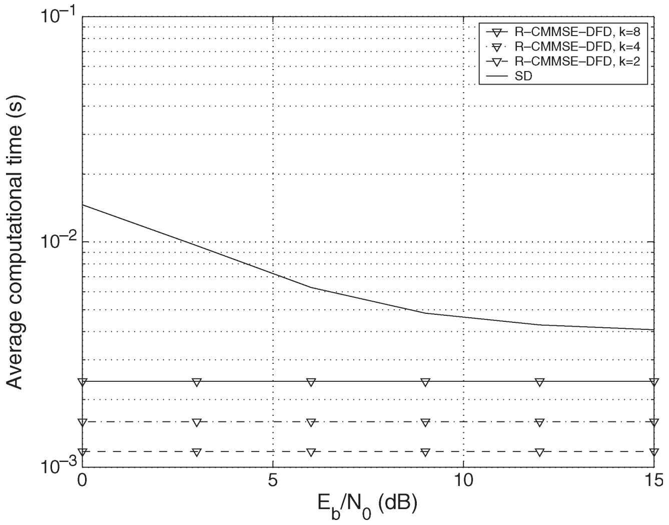


Fig. 8. Average computational time of combined constrained detectors and DFDs in an 8 × 8 BPSK MIMO system using real constraints.

CMMSE-DFD, and it has a 1-dB loss over the CMMSE-DFD at a BER of 10^{-3}. The CML-DFD with $g = 8$ and $k = 8$ achieves the best performance, and it almost reaches the benchmark set by the SD. Therefore, the global metric given by the CML-DFD with $g = 1$ is loose compared to the metrics of the other two detectors. Fig. 10 shows the average computational time of the SD and the CMMSE-DFD for a different k for a 16QAM system. The computational time of the SD does not include the preprocessing. For the CMMSE-DFD, we compute the coefficients a_i and b_i in each block, and this computation is included in the computational time. Similarly, the CMMSE-DFD has constant complexity for the same k over all the SNRs, and its complexity increases with k. In a low SNR, the CMMSE-DFD is less complex than the SD for all values of k. The SD has less complexity than the CMMSE-DFD with $k = 8$ in a high SNR (SNR > 23 dB). In a low SNR (SNR < 15 dB), which is usually the case in practice, our CMMSE-DFD has two to three orders of magnitude of complexity reduction over the SD. For example, at SNR $= 5$ dB, our CMMSE-DFD is 538 times faster than the SD. When the channel is static over several blocks, and a_i and b_i can be precomputed, the resulting computational saving is even more significant.

VI. Conclusion

In this paper, we have proposed several constrained detectors and constrained DFDs. Real constrained detectors are proposed to exploit the real-valued property of the real constellations such as BPSK. This proposal is found to increase the diversity order to $(n + 1)/2$. The previous CLS detector for OFDM/SDMA and the GMMSE detector for CDMA were generalized as MIMO detectors for unitary and nonunitary constellations. A coordinate ascent iterative technique has also been proposed to improve the performance of the proposed detectors. A constrained ordering scheme for DFDs has been derived to alleviate noise enhancement and to improve interference suppression. We also proposed combined constrained detectors and DFDs, defining a global metric to mitigate the error propagation. The complexity of these combined detectors is reasonably low. For future work, the diversity-multiplexing tradeoff of the proposed detectors in MIMO systems could be analyzed.

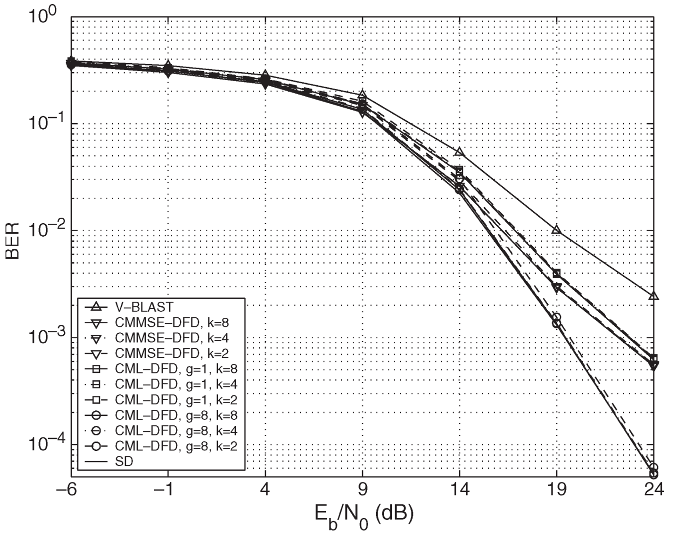


Fig. 9. Performance comparison of combined constrained detectors and DFDs in an 8 × 8 16QAM MIMO system.

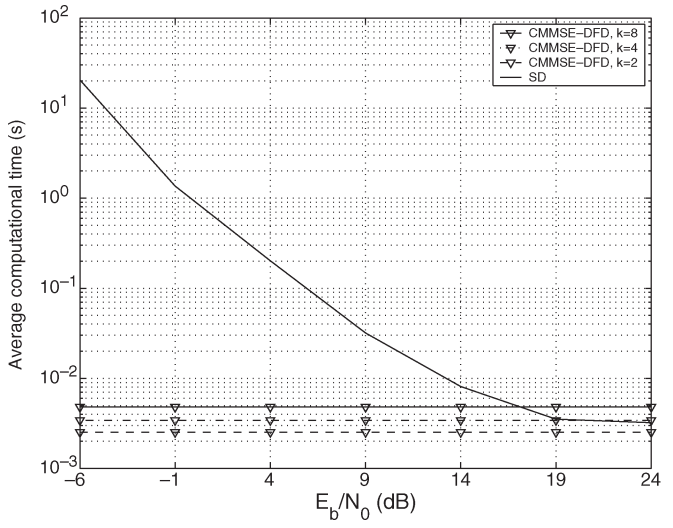


Fig. 10. Average computational time of combined constrained detectors and DFDs in an 8 × 8 16QAM MIMO system.

Acknowledgment

The authors would like to thank the anonymous reviewers for their critical comments, which greatly improved this paper, particularly two of these anonymous reviewers for the comments on the global minimum of (12) and the analytical partial derivative given in (15), (16), and (22).

References

[1] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 35, no. 1, pp. 14–15, Jan. 1999.

[2] G. Ginis and J. M. Cioffi, "On the relation between V-BLAST and the GDFE," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 364–366, Sep. 2001.

[3] U. Fincke and M. Pohst, “Improved methods for calculating vectors of short length in a lattice, including a complexity analysis,” *Math. Comput.*, vol. 44, no. 170, pp. 463–471, Apr. 1985.

[4] E. Viterbo and J. Boutros, “A universal lattice code decoder for fading channels,” *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1639–1642, Jul. 1999.

[5] M. O. Damen, A. Chkeif, and J. C. Belfiore, “Lattice code decoder for space-time codes,” *IEEE Commun. Lett.*, vol. 4, no. 5, pp. 161–163, May 2000.

[6] J. Jalden and B. Ottersten, “On the complexity of sphere decoding in digital communications,” *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.

[7] A. Yener, R. D. Yates, and S. Ulukus, “CDMA multiuser detection: A nonlinear programming approach,” *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 1016–1024, Jun. 2002.

[8] S. Thoen, L. Deneire, L. Van der Perre, M. Engels, and H. D. Man, “Constrained least squares detector for OFDM/SDMA-based wireless networks,” *IEEE Trans. Wireless Commun.*, vol. 2, no. 1, pp. 129–140, Jan. 2003.

[9] W.-K. Ma, T. Davidson, K. M. Wong, Z.-Q. Luo, and P.-C. Ching, “Quasi-maximum-likelihood multiuser detection using semi-definite relaxation with application to synchronous CDMA,” *IEEE Trans. Signal Process.*, vol. 50, no. 4, pp. 912–922, Apr. 2002.

[10] W.-K. Ma, P.-C. Ching, and Z. Ding, “Semidefinite relaxation based multiuser detection for M-ary PSK multiuser systems,” *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2862–2872, Oct. 2004.

[11] P. H. Tan and L. Rasmussen, “The application of semidefinite programming for detection in CDMA,” *IEEE J. Sel. Areas Commun.*, vol. 19, no. 8, pp. 1442–1449, Aug. 2001.

[12] A. Wiesel, Y. C. Eldar, and S. Shamai, “Semidefinite relaxation for detection of 16-QAM signaling in MIMO channels,” *IEEE Signal Process. Lett.*, vol. 12, no. 9, pp. 653–656, Sep. 2005.

[13] N. D. Sidiropoulos and Z.-Q. Luo, “A semidefinite relaxation approach to MIMO detection for high-order QAM constellations,” *IEEE Signal Process. Lett.*, vol. 13, no. 9, pp. 525–528, Sep. 2006.

[14] W.-J. Choi, R. Negi, and J. M. Cioffi, “Combined ML and DFE decoding for the V-BLAST system,” in *Proc. ICC*, Jun. 2000, vol. 3, pp. 1243–1248.

[15] T. Cui, C. Tellambura, and Y. Wu, “Constrained detection for multiple-input multiple-output channels,” in *Proc. GLOBECOM*, Nov. 2005, vol. 1, pp. 199–203.

[16] T. Cui and C. Tellambura, “An efficient generalized sphere decoder for rank-deficient MIMO systems,” *IEEE Commun. Lett.*, vol. 9, no. 5, pp. 423–425, May 2005.

[17] S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, NJ: Prentice-Hall, Sep. 2001.

[18] P. H. Tan and L. Rasmussen, “Multiuser detection in CDMA—A comparison of relaxations, exact, and heuristic search methods,” *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1802–1809, Sep. 2004.

[19] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2002.

[20] G. E. Forsythe and G. H. Golub, “On the stationary values of a second-degree polynomial on the unit sphere,” *J. Soc. Ind. Appl. Math.*, vol. 13, no. 4, pp. 1050–1068, Dec. 1965.

[21] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, Mar. 2004.

[22] N. Z. Shor, *Minimization Methods for Non-Differentiable Functions*. New York: Springer-Verlag, 1985.

[23] D. Chase, “Class of algorithms for decoding block codes with channel measurement information,” *IEEE Trans. Inf. Theory*, vol. IT-18, no. 1, pp. 170–182, Jan. 1972.

[24] D. J. Love, S. Hosur, A. Batra, and R. Heath, “Space-time chase decoding,” *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2035–2039, Sep. 2005.

[25] D. W. Waters and J. R. Barry, “The chase family of detection algorithms for multiple-input multiple-output channels,” in *Proc. GLOBECOM*, Nov. 2004, vol. 4, pp. 2635–2639.

[26] T. Cui and C. Tellambura, “Generalized feedback detection for MIMO systems,” in *Proc. GLOBECOM*, 2005, pp. 3077–3081.

[27] A. Papoulis and S. Pillai, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, Dec. 2001.

[28] S. Loyka and F. Gagnon, “Performance analysis of the V-BLAST algorithm: An analytical approach,” *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1326–1337, Jul. 2004.

[29] Y. Jiang, X. Zheng, and J. Li, “Asymptotic performance analysis of V-BLAST,” in *Proc. IEEE GLOBECOM*, Nov. 2005, pp. 3882–3886.

[30] R. Fischer and C. Windpassinger, “Real versus complex-valued equalisation in V-BLAST systems,” *Electron. Lett.*, vol. 39, no. 5, pp. 470–471, Mar. 6, 2003.

[31] A. Benjebbour, H. Murata, and S. Yoshida, “Comparison of ordered successive receivers for space-time transmission,” in *Proc. VTC—Fall*, Oct. 2001, vol. 4, pp. 2053–2057.

Tao Cui (S’04) received the M.Sc. degree from the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada, in 2005, and the M.S. degree from the Department of Electrical Engineering, California Institute of Technology, Pasadena, in 2006, where he is currently working toward the Ph.D. degree.

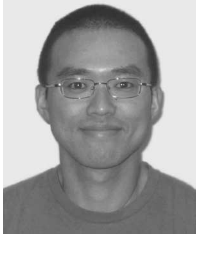
His research interests are in the interactions between networking theory, communication theory, and information theory.

Mr. Cui received the Best Paper Award at the IEEE International Conference on Mobile *Ad hoc* and Sensor Systems (MASS) in 2007 and Second Place in the ACM Student Research Competition at the 2007 Richard Tapia Celebration of Diversity in Computing Conference. He was a recipient of postgraduate scholarships from the Alberta Ingenuity Fund and the Alberta Informatics Circle of Research Excellence.

Chintha Tellambura (SM’02) received the B.Sc. degree (with first-class honors) from the University of Moratuwa, Moratuwa, Sri Lanka, in 1986, the M.Sc. degree in electronics from the University of London, London, U.K., in 1988, and the Ph.D. degree in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 1993.

From 1993 to 1994, he was with the University of Victoria and, from 1995 to 1996, with the University of Bradford, Bradford, U.K., as a Postdoctoral Research Fellow. From 1997 to 2002, he was with Monash University, Melbourne, Australia. He is currently with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada, as a Professor. His research interests include coding, communication theory, modulation, equalization, and wireless communications.

Prof. Tellambura is an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the Area Editor on Wireless Communications Theory and Systems for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He was the Chair of the Communication Theory Symposium at the 2005 IEEE Global Communications Conference, St. Louis, MO.



Yue Wu received the B.Eng. and M.Eng. degrees in information engineering from Xi’an Jiaotong University, Shaanxi, China, in 2000 and 2003, respectively. He is currently working toward the M.Sc. degree with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada.

His research interests include communication theory, broadband wireless communications, and MIMO systems.