

Noise-Predictive Decision-Feedback Detection for Multiple-Input Multiple-Output Channels

Deric W. Waters, *Student Member, IEEE*, and John R. Barry, *Senior Member, IEEE*

Abstract—The decision-feedback (DF) detector is a nonlinear detection strategy for multiple-input multiple-output (MIMO) channels that can significantly outperform a linear detector, especially when the order in which the inputs are detected is optimized according to the so-called Bell Labs Layered Space-Time (BLAST) ordering. The DF detector may be implemented as the cascade of a linear detector, which mitigates interference at the expense of correlating the noise, followed by a noise predictor, which exploits the correlation in the noise to reduce its variance. With this architecture, existing linear detectors can be easily upgraded to DF detectors. We propose a low-complexity algorithm for determining the BLAST ordering that is facilitated by the noise-predictive architecture. The resulting ordered noise-predictive DF detector requires fewer computations than previously reported ordered-DF algorithms. We also propose and derive the ordered noise-predictive minimum-mean-squared-error DF detector and show how to determine its BLAST ordering with low complexity.

Index Terms—Decision feedback, MIMO detection, noise prediction, ordering, reduced-complexity detection, successive interference cancellation, V-BLAST.

I. INTRODUCTION

IN multiple-input multiple-output (MIMO) communications, the detector that minimizes the joint error probability is the maximum-likelihood (ML) detector. Unfortunately, the complexity of the ML detector increases exponentially with the number of channel inputs and is often prohibitively complex. The decision-feedback (DF) detector trades performance for reduced complexity; it is outperformed by the ML detector but requires fewer computations. The DF detector emerges as a popular detection strategy in a wide range of MIMO applications. For example, in the context of a wireless point-to-point link with antenna arrays at both the transmitter and receiver, the DF detector is known as the Bell Labs Layered Space-Time (BLAST) nulling and cancelling detector [1]; in code division multiple access (CDMA) applications, it is known as the DF multiuser detector [2], and in packet transmission, it is known as a generalized DF equalizer (DFE) [3].

Manuscript received July 23, 2003; revised May 25, 2004. This work was supported in part by the National Science Foundation under Grants CCR-0082329 and CCR-0121565. Portions of this work were presented at the IEEE International Symposium on Advances in Wireless Communications (ISWC02), Invited Paper, MA1.4, Victoria, BC, Canada, September 23–24, 2002. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Constantinos B. Papadias.

The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250 USA (e-mail: deric@ece.gatech.edu; barry@ece.gatech.edu).

Digital Object Identifier 10.1109/TSP.2005.845474

The performance of the DF detector is strongly impacted by the order in which the inputs are detected. Unfortunately, optimizing the detection order is a difficult problem that often dominates the overall receiver complexity. It is common and practical to define as optimal the detection order that maximizes the worst-case post-detection SNR. This ordering, which is known as the BLAST ordering, approximately minimizes the joint error probability of the DF detector. The BLAST ordering algorithm of [4] uses repeated computations of a matrix pseudoinverse to find this ordering with a complexity of $O(N^4)$, where N is the number of channel inputs. Two $O(N^3)$ reduced-complexity ordering algorithms have also been proposed: the decorrelating algorithm of [5] and the square-root algorithm of [6]. Other algorithms settle for a suboptimal ordering in order to reduce complexity [7]–[9].

In [2], an architecture for implementing the DF detector based on linear prediction of the noise was presented. The noise-predictive DF detector consists of a linear detector followed by a linear prediction mechanism that reduces the noise variance before making a decision. In this paper, we propose a low-complexity technique for determining the BLAST symbol ordering that is facilitated by the noise-predictive DF detector. The resulting ordered noise-predictive DF (O-NP-DF) detector is mathematically equivalent to the ordered DF detectors of [4]–[6]. However, the O-NP-DF detector is less complex than the lowest complexity BLAST-ordered DF detector previously reported [5]. In fact, if the linear detection filter is already known, the O-NP-DF detector requires less than half the computations required by other BLAST-ordered DF detectors. The noise-predictive approach allows *existing* systems that use linear detection to be transformed (upgraded) into BLAST-ordered DF detectors with the addition of relatively simple processing.

In this paper, we also derive the minimum-mean-squared error (MMSE) version of the noise-predictive DF detector for MIMO channels. We show that our novel ordering algorithm is easily modified to find the MMSE BLAST ordering.

We begin by establishing the channel model and reviewing the conventional DF detector in Section II. Section III describes the noise-predictive zero-forcing DF (NP-ZF-DF) detector of [2]. Section IV describes a low-complexity implementation of the BLAST-ordered NP-ZF-DF detector. Section V derives the noise-predictive MMSE DF (NP-MMSE-DF) detector and describes how to find the corresponding BLAST ordering. Finally, Section VI compares the complexities of both versions of the O-NP-DF detector with previously proposed implementations of the optimally ordered ZF-DF detector.

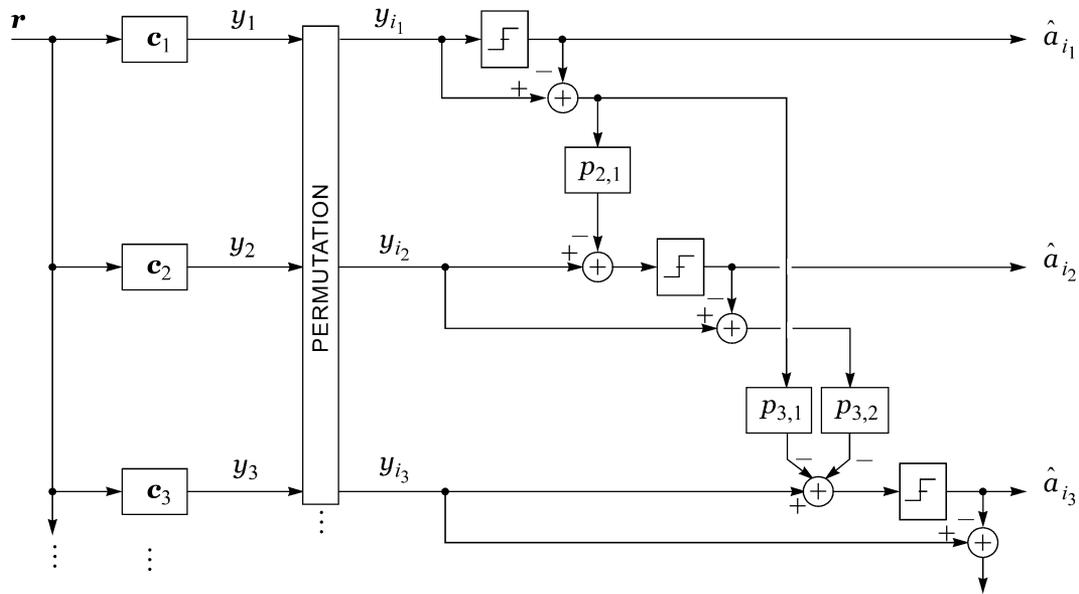


Fig. 1. Noise-predictive DF detector.

II. CONVENTIONAL DF DETECTION

In this paper, we consider the following model of a MIMO channel with N inputs $\mathbf{a} = [a_1, \dots, a_N]^T$ and M outputs $\mathbf{r} = [r_1, \dots, r_M]^T$:

$$\mathbf{r} = \mathbf{H}\mathbf{a} + \mathbf{w} \tag{1}$$

where \mathbf{H} is a complex $M \times N$ channel matrix, and where $\mathbf{w} = [w_1, \dots, w_M]^T$ is additive noise. We assume that the columns of \mathbf{H} are linearly independent, which implies that there are at least as many outputs as inputs $M \geq N$. We assume that the noise components are uncorrelated with complex variance σ^2 so that $E[\mathbf{w}\mathbf{w}^*] = \sigma^2\mathbf{I}$, where \mathbf{w}^* denotes the conjugate transpose of \mathbf{w} . Further, we assume that the inputs are chosen from the same unit-energy alphabet A and are uncorrelated so that $E[\mathbf{a}\mathbf{a}^*] = \mathbf{I}$.

We now summarize the ZF-DF detector: a well-known detection strategy for MIMO channels that was first proposed in [2]. Consider the unique QR decomposition of the channel

$$\mathbf{H} = \mathbf{Q}\mathbf{D}\mathbf{M} \tag{2}$$

where $\mathbf{Q} = [\mathbf{q}_1, \dots, \mathbf{q}_M]$ is an $M \times N$ matrix with orthonormal columns, where \mathbf{D} is an $N \times N$ real diagonal matrix whose diagonal elements are positive, and where \mathbf{M} is an $N \times N$ lower triangular matrix with ones on the diagonal. The DF detector first applies a forward filter $\mathbf{D}^{-1}\mathbf{Q}^*$ (sometimes referred to as the whitened-matched filter) to the received vector, yielding

$$\begin{aligned} \mathbf{y} &= \mathbf{D}^{-1}\mathbf{Q}^*\mathbf{r} \\ &= \mathbf{M}\mathbf{a} + \mathbf{D}^{-1}\mathbf{Q}^*\mathbf{w}. \end{aligned} \tag{3}$$

The i th element of \mathbf{y} is thus

$$y_i = a_i + \sum_{j < i} m_{i,j}a_j + \frac{\mathbf{q}_i^*\mathbf{w}}{d_{i,i}} \tag{4}$$

where $m_{i,j}$ and $d_{i,j}$ are the elements from the i th row and j th column of the matrices \mathbf{M} and \mathbf{D} , respectively. Since \mathbf{M} is lower

triangular, y_1 is free of interference. As a result, the decision \hat{a}_1 can be found directly by quantizing y_1 to the nearest element in A . Using this decision, the interfering term can be subtracted from y_2 . Proceeding iteratively, the ZF-DF detector is succinctly defined by the following recursion:

$$\hat{a}_i = \text{dec} \left\{ y_i - \sum_{j < i} m_{i,j}\hat{a}_j \right\} \tag{5}$$

where $\text{dec}\{x\}$ represents the quantization of x to the nearest symbol in the alphabet A .

III. ZF NOISE-PREDICTIVE DF DETECTION

We now derive an alternative implementation of the ZF-DF detector based on linear prediction of the noise, as first proposed in [2]. Fig. 1 shows the block diagram of the *noise-predictive zero-forcing DF* (NP-ZF-DF) detector that employs this linear-prediction strategy; the filters \mathbf{c}_i and $p_{i,j}$ will be defined shortly. The notion of ordering (the permutation block) is neglected momentarily by assuming an identity permutation.

The starting point for the noise-predictive ZF-DF detector is the ZF *linear* detector [11], which essentially inverts the channel by computing $\mathbf{y} = \mathbf{C}\mathbf{r}$, where \mathbf{C} is the channel pseudoinverse:

$$\mathbf{C} = (\mathbf{H}^*\mathbf{H})^{-1}\mathbf{H}^*. \tag{6}$$

In Fig. 1, \mathbf{c}_i denotes the i th row of \mathbf{C} . From (1), the output of this filter is free of interference:

$$\mathbf{y} = \mathbf{a} + \mathbf{n} \tag{7}$$

where the noise $\mathbf{n} = [n_1, \dots, n_N]^T = \mathbf{C}\mathbf{w}$ is no longer white; its autocorrelation matrix is $\mathbf{R}_{nn} = E[\mathbf{n}\mathbf{n}^*] = \sigma^2(\mathbf{H}^*\mathbf{H})^{-1}$.

The correlation of the noise can be exploited using linear prediction to reduce its variance. If the first $i - 1$ elements of the noise vector were known, we could form an estimate \hat{n}_i of the i th element n_i and subtract this estimate from y_i to reduce its

variance. Specifically, given $\{n_1, \dots, n_{i-1}\}$, a linear predictor estimates n_i according to

$$\hat{n}_i = \sum_{j<i} p_{i,j} n_j \quad (8)$$

or, equivalently, $\hat{\mathbf{n}} = \mathbf{P}\mathbf{n}$, where \mathbf{P} is a strictly lower triangular *prediction filter* whose element at the i th row and j th column is $p_{i,j}$. This process is complicated by the fact that the receiver does not have access to n_i directly but rather to the sum $y_i = a_i + n_i$. However, as shown in Fig. 1, the decision about a_i can be subtracted from y_i to yield n_i as long as the decision is correct.

Let us define the *total MSE* as $E[\|\hat{\mathbf{n}} - \mathbf{n}\|^2] = \sum_i E[|\hat{n}_i - n_i|^2]$, which measures the quality of the prediction. As shown in [12], this total MSE is minimized by the following prediction filter:

$$\mathbf{P} = \mathbf{I} - \mathbf{M} \quad (9)$$

where \mathbf{M} is defined by the QR decomposition of (2). Having thus defined the prediction coefficients, the NP-ZF-DF detector of Fig. 1 can be summarized succinctly by the following recursion:

$$\hat{a}_i = \text{dec} \left\{ y_i - \sum_{j<i} p_{i,j} (y_j - \hat{a}_j) \right\}. \quad (10)$$

We now show that the ZF-DF detector (5) and the NP-ZF-DF detector (10) are equivalent. Substituting (7) and (9) into (10) yields the following for the noise-predictive implementation:

$$\hat{a}_i = \text{dec} \left\{ a_i + \sum_{j \leq i} m_{i,j} \mathbf{c}_j \mathbf{w} - \sum_{j<i} m_{i,j} (\hat{a}_j - a_j) \right\} \quad (11)$$

where we exploited the fact that $m_{i,i} = 1$ and $p_{i,j} = -m_{i,j}$ when $j < i$, and where we substituted $n_j = \mathbf{c}_j \mathbf{w}$. On the other hand, for the conventional implementation, substituting (4) into (5) gives

$$\hat{a}_i = \text{dec} \left\{ a_i + \frac{\mathbf{q}_i^* \mathbf{w}}{d_{i,i}} - \sum_{j<i} m_{i,j} (\hat{a}_j - a_j) \right\}. \quad (12)$$

The conventional and noise-predictive detectors are equivalent when (11) and (12) are identical or when

$$\sum_{j \leq i} m_{i,j} \mathbf{c}_j = \frac{\mathbf{q}_i^*}{d_{i,i}}. \quad (13)$$

In matrix form, (13) simplifies to

$$\mathbf{M}\mathbf{C} = \mathbf{D}^{-1}\mathbf{Q}^* \quad (14)$$

but since

$$\begin{aligned} \mathbf{M}\mathbf{C} &= \mathbf{M}(\mathbf{H}^*\mathbf{H})^{-1}\mathbf{H}^* \\ &= \mathbf{M}(\mathbf{M}^{-1}\mathbf{D}^{-2}\mathbf{M}^{-*})\mathbf{H}^* \\ &= \mathbf{D}^{-1}\mathbf{Q}^* \end{aligned} \quad (15)$$

we conclude that the conventional ZF-DF detector and the NP-ZF-DF detector are indeed equivalent.

IV. OPTIMALLY ORDERED ZF NOISE-PREDICTIVE DF DETECTION

To implement the ordered NP-ZF-DF detector of Fig. 1, the receiver must first determine the channel pseudoinverse \mathbf{C} , the symbol detection order, and the linear prediction filter \mathbf{P} . In this section, we show how to calculate both the optimal detection order and the prediction filter given knowledge of the channel pseudoinverse.

We first describe two low-complexity algorithms for finding the best (BLAST) detection order. As implied by Fig. 1, these sorting algorithms occur after $\mathbf{y} = \mathbf{C}\mathbf{r}$ has been calculated. The permutation in the block diagram of Fig. 1 gives the detector the flexibility to use any symbol detection order, but in this paper, we assume that the BLAST ordering is used. Let $\{i_1, i_2, \dots, i_N\}$ denote the BLAST ordering, which is a permutation of the integers $\{1, 2, \dots, N\}$ such that i_k denotes the index of the k th symbol to be detected.

The noise-predictive view of the DF detector leads to a simple algorithm for finding the BLAST ordering. As proven in [1], the BLAST ordering can be found in a recursive fashion by choosing each i_k to maximize the post-detection SNR of the k th symbol or equivalently minimize its MSE. Specifically, because the MSE for the first detected symbol is $\sigma^2 \|\mathbf{c}_{i_1}\|^2$, we have

$$i_1 = \arg \min_{j \in \{1, \dots, N\}} \|\mathbf{c}_j\|^2. \quad (16)$$

In other words, the channel pseudoinverse row with the smallest norm determines which symbol to detect first. Once i_1 is chosen, and assuming \hat{a}_{i_1} is correct, the MSE for the second symbol is

$$\begin{aligned} E[|n_{i_2} - \hat{n}_{i_2}|^2] &= E[|\mathbf{c}_{i_2} \mathbf{w} - p_{2,1} \mathbf{c}_{i_1} \mathbf{w}|^2] \\ &= \sigma^2 \|\mathbf{c}_{i_2} - p_{2,1} \mathbf{c}_{i_1}\|^2. \end{aligned} \quad (17)$$

When the prediction coefficient $p_{2,1}$ is chosen to minimize the above MSE, the term $p_{2,1} \mathbf{c}_{i_1}$ reduces to the projection of \mathbf{c}_{i_2} onto the subspace spanned by \mathbf{c}_{i_1} , which we denote as $\hat{\mathbf{c}}_{i_2}$. Hence, the optimal i_2 satisfies

$$i_2 = \arg \min_{j \neq i_1} \|\mathbf{c}_j - \hat{\mathbf{c}}_j\|^2. \quad (18)$$

Repeating the above procedure recursively leads to the following simple and succinct procedure for finding the BLAST ordering:

$$i_k = \arg \min_{j \notin \{i_1, \dots, i_{k-1}\}} \|\mathbf{c}_j - \hat{\mathbf{c}}_j\|^2 \quad (19)$$

where $\hat{\mathbf{c}}_j$ denotes the projection of \mathbf{c}_j onto the span of $\{\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_{k-1}}\}$. This is a key result that is the basis of the noise-predictive implementation of the BLAST-ordered DF detector. In words, *finding the BLAST ordering amounts to choosing the rows of the channel pseudoinverse, where the best choice for the k th row is the unchosen row that is closest to the subspace spanned by the rows already chosen.*

Function A. Input: \mathbf{C} . Output: $\{i_1, i_2, \dots, i_N\}$ and \mathbf{F}

- (A-1) $\mathcal{U} = \{1, 2, \dots, N\}$ = the set of unchosen rows.
 (A-2) $E_j = \|\mathbf{c}_j\|^2, j \in \mathcal{U}$.
 (A-3) for $k = 1$ to $N - 1$,
 (A-4) $i_k = \underset{j \in \mathcal{U}}{\operatorname{argmin}} E_j$
 (A-5) $\mathcal{U} = \mathcal{U} - i_k$; remove chosen row from \mathcal{U} .
 (A-6) $f_{i_k, k} = \sqrt{E_{i_k}}$
 (A-7) $\mathbf{c}_{i_k} = \mathbf{c}_{i_k} / f_{i_k, k}$
 (A-8) for $j \in \mathcal{U}$
 (A-9) $f_{j, k} = \mathbf{c}_j \mathbf{c}_{i_k}^*$
 (A-10) $\mathbf{c}_j = \mathbf{c}_j - f_{j, k} \mathbf{c}_{i_k}$
 (A-11) $E_j = E_j - |f_{j, k}|^2$
 (A-12) end
 (A-13) end
 (A-14) $i_N = \mathcal{U}$.

Fig. 2. Noise-predictive sorting algorithm using MGS orthogonalization.

A computationally efficient implementation of the sorting algorithm of (19) is given in Fig. 2. It is based on an adaptation of the modified Gram–Schmidt (MGS) QR decomposition [13]. The algorithm accepts the channel pseudoinverse \mathbf{C} as an input, and it produces two outputs: the optimal ordering $\{i_1, \dots, i_N\}$ and an intermediate matrix \mathbf{F} that can be used to determine the linear prediction filter \mathbf{P} . The MGS procedure of the sorting algorithm operates on the rows of \mathbf{C} , $\{\mathbf{c}_{i_1}, \dots, \mathbf{c}_{i_N}\}$. During the first iteration ($k = 1$), line A-4 chooses the row nearest to the null space. Then, line A-10 removes the portions from the remaining rows of \mathbf{C} that are parallel to \mathbf{c}_{i_1} . Therefore, in the next iteration ($k = 2$), each of the candidate rows of \mathbf{C} is orthogonal to \mathbf{c}_{i_1} . Consequently, the remaining row closest to the subspace spanned by the previously chosen row is simply the row with minimum norm. As before, line A-10 ensures that the remaining rows of \mathbf{C} are orthogonal to \mathbf{c}_{i_2} . The iterations continue until $k = N - 1$, when the BLAST ordering is determined.

We now present an alternative algorithm for finding the optimal ordering based on the Householder QR decomposition [13], which is less complex than Function A for large N . The pseudocode is given in Fig. 3. It is similar to the MGS algorithm, but it rotates the subspace represented by the unchosen rows of \mathbf{C} such that making them orthogonal to the row just chosen is done by deleting the first column of \mathbf{C} . Line B-6 creates a Householder vector \mathbf{v} that yields a matrix \mathbf{C} whose effective dimensions are $(N - k) \times (M - k)$ after the subspace rotation performed in line B-8.

Given the intermediate output \mathbf{F} of either of the sorting algorithms just described, calculating the linear prediction filter \mathbf{P} is straightforward. To avoid confusion, let $\mathbf{\Pi}$ denote an $N \times N$ permutation matrix whose j th column is the i_j th column of the identity matrix. In Fig. 1, the ordering is accounted for by permuting the rows of the linear detector so that the cascade of the

Function B. Input: \mathbf{C} . Output: $\{i_1, i_2, \dots, i_N\}$ and \mathbf{F}

- (B-1) $\mathcal{U} = \{1, 2, \dots, N\}$ = the set of unchosen rows.
 (B-2) $E_j = \|\mathbf{c}_j\|^2, j \in \mathcal{U}$.
 (B-3) for $k = 1$ to $N - 1$,
 (B-4) $i_k = \underset{j \in \mathcal{U}}{\operatorname{argmin}} E_j$
 (B-5) $\mathcal{U} = \mathcal{U} - i_k$; remove chosen row from \mathcal{U} .
 (B-6) $\mathbf{v} = \mathbf{c}_{i_k}; d = \sqrt{E_{i_k}}; v_1 = v_1 + dv_1 / |v_1|$
 (B-7) $\mathbf{x} = \mathbf{C}\mathbf{v}^*$
 (B-8) $\mathbf{C} = \mathbf{C} - \mathbf{x}\mathbf{v}^* / x_{i_k}$
 (B-9) $E_j = E_j - |\mathbf{c}_j \cdot \mathbf{v}|^2, j \in \mathcal{U}$
 (B-10) Delete first column from \mathbf{C} ; store it as k -th column of \mathbf{F}
 (B-11) end
 (B-12) $i_N = \mathcal{U}$.

Fig. 3. Noise-predictive sorting algorithm using Householder orthogonalization.

Function C. Input: \mathbf{F} . Output: \mathbf{P}

- (C-1) $\mathbf{P} = \mathbf{0}_{N \times N}; \mathbf{T} = \mathbf{I}_{N \times N}$
 (C-2) for $k = 2$ to N ,
 (C-3) for $j = (k - 1)$ down to 1,
 (C-4) $t_{k, j} = f_{i_k, j} / f_{i_j, j}$
 (C-5) $p_{k, j} = t_{k, j} - \sum_{m=j+1}^{k-1} t_{k, m} p_{m, j}$
 (C-6) end
 (C-7) end

Fig. 4. Calculation of the prediction filter \mathbf{P} from the output of the noise-predictive sorting functions A and B.

channel pseudoinverse and the permutation leads to an effective front-end filter of

$$\mathbf{C}' = \mathbf{\Pi}^* \mathbf{C}. \quad (20)$$

This ordered linear detection filter is the pseudoinverse of an ordered channel matrix

$$\mathbf{H}' = \mathbf{H} \mathbf{\Pi}. \quad (21)$$

When performed on \mathbf{H}' , the decomposition of (2) yields the matrices \mathbf{D}' and \mathbf{M}' . From (9) and in these new terms, the ordered prediction filter is

$$\mathbf{P} = \mathbf{I} - \mathbf{M}'. \quad (22)$$

Unfortunately, unless $\mathbf{\Pi}$ is the identity matrix, \mathbf{M}' is not equal to \mathbf{M} . However, the matrix $(\mathbf{M}')^{-1}$ is easily calculated from the output matrix \mathbf{F} of the sorting algorithm (Function A or B). First, permute the rows of \mathbf{F} such that it is a lower triangular matrix. The columns of this permuted \mathbf{F} need only be divided by the corresponding diagonal elements $f_{i_j, j}$ for the j th column to arrive at $(\mathbf{M}')^{-1}$. Next, simply invert $(\mathbf{M}')^{-1}$ using back substitution, and use (22) to get \mathbf{P} . Fig. 4 gives the pseudocode for this procedure.

In summary, the optimally ordered noise-predictive zero-forcing DF (O-NP-ZF-DF) detector implementation has four steps. First, the channel pseudoinverse is applied to the received vector. Next, the optimal symbol order is calculated from the channel pseudoinverse using either the MGS or Householder sorting algorithm. Then, the linear prediction filter is calculated from the output of the sorting algorithm. After these calculations, the detector can be implemented using (10), as illustrated in Fig. 1.

V. NOISE-PREDICTIVE MMSE DF DETECTION

The ZF-DF detector cancels the interference completely without regard to noise amplification. The MMSE-DF detector improves on this strategy by finding the optimal balance between interference cancellation and noise reduction that minimizes the total MSE [14]. In this section, we derive a noise-predictive implementation of the optimally ordered MMSE-DF detector.

Like the ZF-DF detector, the MMSE-DF detector can also be implemented as a cascade of a linear filter and a noise-predictive mechanism so that the basic architecture of Fig. 1 applies to both the zero-forcing and MMSE versions of the DF detector. However, instead of the channel pseudoinverse, the noise-predictive MMSE DF (NP-MMSE-DF) detector begins with the MMSE linear detection filter [11]:

$$\mathbf{C}_{\text{MMSE}} = \mathbf{R}^{-1}\mathbf{H}^* \quad (23)$$

where

$$\mathbf{R} = \mathbf{H}^*\mathbf{H} + \sigma^2\mathbf{I}. \quad (24)$$

This choice for \mathbf{C}_{MMSE} minimizes the total MSE $E[\|\boldsymbol{\varepsilon}\|^2]$, where $\boldsymbol{\varepsilon} = \mathbf{C}_{\text{MMSE}}\mathbf{r} - \mathbf{a}$ is the vector of errors after the linear filter. Unlike the ZF case, this error vector contains residual intersymbol interference (ISI) as well as noise:

$$\begin{aligned} \boldsymbol{\varepsilon} &= \mathbf{R}^{-1}(\mathbf{H}^*\mathbf{H} + \sigma^2\mathbf{I} - \sigma^2\mathbf{I})\mathbf{a} + \mathbf{C}_{\text{MMSE}}\mathbf{w} - \mathbf{a} \\ &= -\sigma^2\mathbf{R}^{-1}\mathbf{a} + \mathbf{C}_{\text{MMSE}}\mathbf{w}. \end{aligned} \quad (25)$$

In the following, we continue to use our “noise”-predictive terminology, even though strictly speaking, the “noise” being predicted is $\boldsymbol{\varepsilon}$, which contains residual ISI as well as noise.

Let $\hat{\boldsymbol{\varepsilon}} = \mathbf{P}\boldsymbol{\varepsilon}$ denote an estimate of $\boldsymbol{\varepsilon}$ based on linear prediction. Let $\mathbf{e} = (\mathbf{I} - \mathbf{P})\boldsymbol{\varepsilon}$ denote the error in this estimate. We now derive the strictly lower triangular linear prediction filter \mathbf{P} that minimizes the total MSE $E[\|\mathbf{e}\|^2]$. From (25), the autocorrelation matrix $\mathbf{R}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} = E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^*]$ of $\boldsymbol{\varepsilon}$ can be written as

$$\begin{aligned} \mathbf{R}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} &= \sigma^4\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{R}}^{-1} + \sigma^2\mathbf{C}_{\text{MMSE}}\mathbf{C}_{\text{MMSE}}^* \\ &= \sigma^4\tilde{\mathbf{R}}^{-2} + \sigma^2\tilde{\mathbf{R}}^{-1}(\mathbf{H}^*\mathbf{H} + \sigma^2\mathbf{I} - \sigma^2\mathbf{I})\tilde{\mathbf{R}}^{-1} \\ &= \sigma^2\tilde{\mathbf{R}}^{-1}. \end{aligned} \quad (26)$$

Since $\tilde{\mathbf{R}}$ is Hermitian and positive definite, $\mathbf{R}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}$ has the following Cholesky factorization:

$$\mathbf{R}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} = \sigma^2\tilde{\mathbf{M}}^{-1}\tilde{\mathbf{D}}^{-2}\tilde{\mathbf{M}}^{-*} \quad (27)$$

where $\tilde{\mathbf{M}}^{-1}$ is a lower triangular matrix with diagonal elements of one, and where $\tilde{\mathbf{D}}^{-2}$ is a real diagonal matrix with positive diagonal elements. The total MSE after linear prediction is related to $\mathbf{R}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}$ by

$$E[\|\mathbf{e}\|^2] = \text{trace}\{(\mathbf{I} - \mathbf{P})\mathbf{R}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}(\mathbf{I} - \mathbf{P}^*)\}. \quad (28)$$

It is easy to show [12] that the best choice for $(\mathbf{I} - \mathbf{P})$ cancels $\tilde{\mathbf{M}}^{-1}$:

$$\mathbf{I} - \mathbf{P} = \tilde{\mathbf{M}}. \quad (29)$$

Therefore, the effective front-end filter of the noise-predictive MMSE-DF (NP-MMSE-DF) detector is given by

$$\begin{aligned} (\mathbf{I} - \mathbf{P})\mathbf{C}_{\text{MMSE}} &= \tilde{\mathbf{M}}\tilde{\mathbf{R}}^{-1}\mathbf{H}^* \\ &= \tilde{\mathbf{D}}^{-2}\tilde{\mathbf{M}}^{-*}\mathbf{H}^*. \end{aligned} \quad (30)$$

This forward filter is identical to the forward filter of the conventional MMSE-DF detector defined in [15]. With this forward filter, the corresponding feedback filter is $-\mathbf{P}$, which is identical to the feedback filter of the conventional MMSE-DF detector defined in [15]. Therefore, we conclude that the NP-MMSE-DF detector is equivalent to the conventional MMSE-DF detector.

Just as for the ZF-DF detector, the performance of the MMSE-DF detector is improved if the detection order of the symbols is chosen to minimize the maximum MSE. However, the ordering problem for the NP-MMSE-DF detector is complicated by the fact that the “noise” includes residual ISI. For convenience, we define an augmented matrix \mathbf{B} :

$$\mathbf{B} = [\mathbf{C}_{\text{MMSE}} \quad \sigma\tilde{\mathbf{R}}^{-1}] \quad (31)$$

so that $\mathbf{R}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} = \sigma^2\mathbf{B}\mathbf{B}^*$. Let \mathbf{b}_j be the j th row of the matrix \mathbf{B} . From (26), the MSE for the first detected symbol is equal to $[\mathbf{R}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}]_{i_1, i_1} = \sigma^2\|\mathbf{b}_{i_1}\|^2$. Therefore, we choose the symbol with minimum MSE by

$$i_1 = \arg \min_{j \in \{1, \dots, N\}} \|\mathbf{b}_j\|^2. \quad (32)$$

After i_1 is chosen, and assuming \hat{a}_{i_1} is correct, the MSE for the second symbol is

$$E[|\varepsilon_{i_2} - \hat{\varepsilon}_{i_2}|^2] = E[|\varepsilon_{i_2} - p_{2,1}\varepsilon_{i_1}|^2]. \quad (33)$$

Let $\tilde{\mathbf{r}}_j$ be the j th row of the matrix $\tilde{\mathbf{R}}^{-1}$. Then, by substituting from (25), the MSE for the second symbol becomes

$$\begin{aligned} &E[|\boldsymbol{\varepsilon}_{i_2} - p_{2,1}\boldsymbol{\varepsilon}_{i_1}|^2] \\ &= E\left[|\mathbf{c}_{i_2}\mathbf{w} - \sigma^2\mathbf{r}_{i_2}\mathbf{a} - p_{2,1}\mathbf{c}_{i_1}\mathbf{w} + \sigma^2p_{2,1}\mathbf{r}_{i_1}\mathbf{a}|^2\right] \\ &= \sigma^2\|\mathbf{c}_{i_2} - p_{2,1}\mathbf{c}_{i_1}\|^2 + \sigma^4\|\mathbf{r}_{i_2} - p_{2,1}\mathbf{r}_{i_1}\|^2 \\ &= \sigma^2\|\mathbf{b}_{i_2} - p_{2,1}\mathbf{b}_{i_1}\|^2 \end{aligned} \quad (34)$$

where the last equality in (34) follows from straightforward algebraic manipulation. When the prediction coefficient $p_{2,1}$ is chosen to minimize the MSE, the term $p_{2,1}\mathbf{b}_{i_1}$ reduces to the

TABLE I
COMPLEXITY OF FUNCTION A: MGS SORTING ALGORITHM

Line	ZF Operations
(A-2)	$2MN - N$
(A-6)-(A-7)	$MN - M + N - 1$
(A-9)	$MN^2 - MN - N^2/2 + N/2$
(A-10)	$MN^2 - MN$
(A-11)	$N^2 - N - 2$
ZF total complexity	$2MN^2 + MN + N^2/2 - M - N/2 - 3$
MMSE total complexity	$2MN^2 + 2N^3 + MN + 3N^2/2 - M - 3N/2 - 3$

projection of \mathbf{b}_{i_2} onto the subspace spanned by \mathbf{b}_{i_1} , which we denote as $\hat{\mathbf{b}}_{i_2}$. Hence, the optimal i_2 satisfies

$$i_2 = \arg \min_{j \neq i_1} \|\mathbf{b}_j - \hat{\mathbf{b}}_j\|^2. \quad (35)$$

The above procedure can be repeated recursively to determine the BLAST ordering. In a fashion reminiscent of the ZF sorting algorithm (19), the above procedure is succinctly described by the following recursive sorting algorithm:

$$i_k = \arg \min_{j \notin \{i_1, \dots, i_{k-1}\}} \|\mathbf{b}_j - \hat{\mathbf{b}}_j\|^2 \quad (36)$$

where $\hat{\mathbf{b}}_j$ denotes the projection of \mathbf{b}_j onto the span of $\{\mathbf{b}_{i_1}, \dots, \mathbf{b}_{i_{k-1}}\}$.

The MMSE sorting algorithm (36) just described is identical to the ZF sorting algorithm (19), except that \mathbf{c}_j and $\hat{\mathbf{c}}_j$ have been replaced by \mathbf{b}_j and $\hat{\mathbf{b}}_j$, respectively. As a result, we need not derive an implementation of the MMSE ordering algorithm from scratch; it is realized by both Functions A and B when their inputs are the augmented matrix \mathbf{B} instead of the channel pseudoinverse.

When $\sigma = 0$, \mathbf{C}_{MMSE} reduces to the channel pseudoinverse, and the $\sigma \hat{\mathbf{R}}^{-1}$ matrix has no impact on the sorting algorithm. Therefore, as expected, the MMSE sorting algorithm reduces to the ZF sorting algorithm when the noise is zero.

VI. COMPLEXITY COMPARISON

In this section, we compare the complexity of the optimally ordered noise-predictive zero-forcing DF (O-NP-ZF-DF) detector to the complexity of three previously reported optimally ordered ZF-DF detector implementations: the BLAST detector [4], the modified decorrelation DF (MDDF) detector, and the square-root (SQRT) detector. The BLAST detector has $O(N^4)$ complexity due to repeated pseudoinverse calculations. The MDDF detector [5] uses the symmetry of the Cholesky decomposition of the channel to reduce the complexity to $O(N^3)$. Finally, the SQRT detector [6] exploits the symmetry in the QR decomposition of the channel to achieve $O(N^3)$ complexity. The SQRT detector of [6] uses the MMSE criterion, but for the sake of a fair complexity comparison, we consider its ZF version here.

Performance does not affect this comparison since all of the detectors are mathematically identical. The distinguishing criteria among these detectors is the number of computations they require. None of the detectors addresses estimation directly; therefore, in this comparison, the complexity of estimation is neglected. However, it is possible to estimate \mathbf{C} directly [10], [11], in which case, the O-NP-DF detector would have very low complexity. Since many systems estimate \mathbf{H} rather than \mathbf{C} , we also examine the complexity when the channel pseudoinverse is calculated from \mathbf{H} .

Several notes are appropriate regarding the complexity comparison. First, we measure complexity as the total number of complex additions, subtractions, multiplications, divisions, and square-roots required each time the detector is calculated. Second, in the context of DF detectors, MIMO systems with N and M as low as two are of interest. As a result, lower order complexity terms are not always negligible. Third, the SQRT algorithm proposed using a more complex QR decomposition that improves numerical stability by avoiding the inversion of a lower triangular matrix. However, to make a fair complexity comparison, we assume that the MGS implementation of the QR decomposition [13] is used by all detectors. Thus, the SQRT algorithm we consider here is less complex and less stable than that originally proposed in [6]. The complexity of the SQRT and MDDF algorithms depends on the ordering of the channel. We consider their worst-case complexity since a practical implementation must be designed to handle the maximum possible complexity. Finally, the complexity of the quantization operation is ignored since it is the same for all the detectors and depends on the symbol constellation.

The complexity analysis begins with the proposed optimally ordered noise-predictive ZF DF (O-NP-ZF-DF) detector. Tables I and II itemize the complexities of the MGS and Householder-based sorting algorithms, respectively. The complexity of their respective MMSE counterparts can be found by substituting $(M + N)$ for M . The MGS sorting algorithm (Function A) is less complex than the Householder sorting algorithm (Function B) for small M and N . However, depending on the condition number of the input matrix, the MGS algorithm may have significantly worse numerical accuracy [13, p. 232]. The calculation of the linear prediction filter (see Fig. 4) has a complexity of $N^3/3 - N^2/2 + N/6$ for both the ZF and MMSE ver-

TABLE II
COMPLEXITY OF FUNCTION B: HOUSEHOLDER SORTING ALGORITHM

Line	ZF Operations
(B-2)	$2MN - N$
(B-6)	$5N - 5$
(B-7)	$MN^2 - N^3/3 + MN - N^2/2 - 2M + 11N/6 - 1$
(B-8)	$MN^2 - N^3/3 + 2MN - N^2/2 - 3M + 23N/6 - 3$
(B-9)	$N^2 - N - 2$
ZF total complexity	$2MN^2 - 2N^3/3 + 5MN - 5M + 26N/3 - 11$
MMSE total complexity	$2MN^2 + 4N^3/3 + 5MN + 5N^2 - 5M + 11N/3 - 11$

TABLE III
COMPARISON OF DF DETECTORS

Algorithm	Number of Computations	Required Estimation
O-NP-ZF-DF MGS, Function A	$2MN^2 + N^3/3 + MN - M - N/3 - 3$ $+ (2MN + N^2 - N - 1)L$	C
O-NP-ZF-DF Householder, Function B	$2MN^2 - N^3/3 + 5MN - N^2/2 - 5M + 27N/6 -$ $11 + (2MN + N^2 - N - 1)L$	C
O-NP-ZF-DF $M=N$ Function A	$2MN^2 + 2N^3 + MN - M - N - 3$ $+ (2MN + N^2 - N - 1)L$	H
O-NP-ZF-DF $M>N$ Function A	$5MN^2 + 2N^3/3 - 2MN - M + 17N/6 - 3$ $+ (2MN + N^2 - N - 1)L$	H
ZF-MDDF	$2MN^2 + 11N^3/3 + 2N^2 + 10N/3 - 1$ $+ (2MN + N^2 - N)L$	H
ZF-SQRT	$4MN^2 + 5N^3/3 + 2MN + 3N^2 + 10N/3$ $+ (4MN - 2M - N)L$	H
ZF-BLAST	$MN^3 + N^4/12 + 5MN^2/2 + 3MN/2$ $+ 2N^2/3 + N/2 - 2M + 1$ $+ (4MN - 2M - N)L$	H

sions of the O-NP-DF detector. After $\{\mathbf{c}_i\}$ and $\{p_{i,j}\}$ are known, the detector of (10) has complexity $2MN + N^2 - N1$.

If the channel matrix is known, the least-complex option for calculating the channel pseudoinverse depends on the relative values of M and N . If $M = N$, we propose to use the lower-upper triangular (LU) decomposition [13, p. 98], which requires $5N^3/3 - 2N/3$ operations to calculate the pseudoinverse. Otherwise, we propose to use the MGS QR decomposition, which requires $3MN^2 + N^3/3 - 3MN + 19N/6$ operations to calculate the pseudoinverse.

Table III gives the complexity of the O-NP-ZF-DF detector when the channel pseudoinverse is given and when the channel matrix is given, where it is assumed that the detection filters are recalculated every L symbol periods. For comparison, Table III also gives the complexity of the ZF-MDDF, ZF-SQRT, and

ZF-BLAST algorithms, assuming that MGS QR decompositions are used. Fig. 5 illustrates the complexities of Table III as a function of N , where $M = N$ and $L = 1$ are assumed. For $N > 6$, Function B is used to implement the O-NP-ZF-DF detector (given **C**), which is shown to be roughly 56% less complex than the MDDF detector (given **H**). Fig. 5 also shows that the O-NP-ZF-DF detector (given **H**) requires approximately 32% fewer computations than the ZF-MDDF detector (given **H**).

The complexity curves generally follow the dominant terms in the complexity expressions. However, the BLAST complexity curve in Fig. 5 demonstrates the importance of including lower order terms. Even though the BLAST detector's asymptotic complexity is an order of magnitude greater, it is less complex than the SQRT and MDDF detectors when

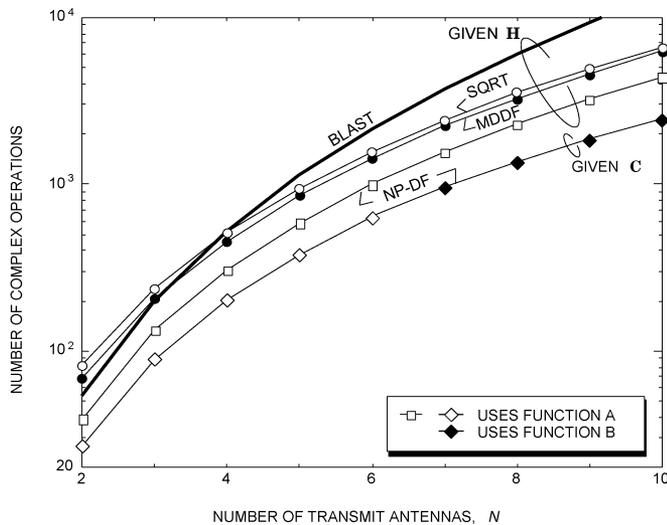


Fig. 5. Complexity comparison for various implementations of the ordered ZF-DF detector, assuming the ZF criterion with $M = N$ and $L = 1$.

$M = N < 4$. The complexity of the BLAST algorithm could be further reduced by using the LU decomposition to calculate first the channel pseudoinverse and, then, the MGS QR decomposition for the other pseudoinverse calculations.

VII. CONCLUSION

The noise-predictive DF detector consists of a linear detector and a linear prediction mechanism that reduces noise variance. We showed that the noise-predictive view of the DF detector leads to a simple and computationally efficient way of calculating the BLAST detection ordering for both the MMSE and ZF versions of the DF detector. The noise-predictive implementation makes it easy to upgrade an existing linear detector by appending relatively simple additional processing. Furthermore, despite the fact that the linear detector and this add-on processing may have been designed independently, the overall complexity of the resulting noise-predictive DF detector is lower than previously reported ordered DF detectors.

REFERENCES

- [1] G. J. Foschini, G. Golden, R. Valenzuela, and P. Wolniansky, "Simplified processing for wireless communication at high spectral efficiency," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 11, pp. 1841–1852, Nov. 1999.
- [2] A. Duel-Hallen, "Decorrelating decision-feedback multiuser detector for synchronous code-division multiple access channel," *IEEE Trans. Commun.*, vol. 41, no. 2, pp. 285–290, Feb. 1993.
- [3] J. M. Cioffi and G. D. Forney *et al.*, "Generalized decision-feedback equalization for packet transmission with ISI and Gaussian noise," in *Communication, Computation, Control and Signal Processing*, A. Paulraj *et al.*, Eds. Boston, MA: Kluwer, 1997, ch. 4, pp. 79–127.

- [4] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over rich-scattering wireless channel," in *Proc. Int. Symp. Signals, Syst., Electron.*, Oct. 1998, pp. 295–300.
- [5] W. Zha and S. Blostein, "Modified decorrelating decision-feedback detection of BLAST space-time system," in *Proc. Int. Conf. Commun.*, vol. 1, May 2002, pp. 335–339.
- [6] B. B. Hassibi, "An efficient square-root algorithm for BLAST," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 2, Jun. 2000, pp. II737–II740.
- [7] D. Wubben, R. Bohnke, J. Rinas, V. Kugn, and K. Kammeyer, "Efficient algorithm for decoding layered space-time codes," *Electron. Lett.*, vol. 37, no. 22, pp. 1348–1350, Oct. 25, 2001.
- [8] W. Wai, C. Tsui, and R. Cheng, "A low complexity architecture of the V-BLAST System," in *Proc. IEEE Wireless Commun. Networking Conf.*, vol. 1, 2000, pp. 310–314.
- [9] X. Tao, Z. Yu, H. Qin, P. Zhang, H. Haas, and E. Costa, "New sub-optimal detection algorithm of layered space-time code," in *Proc. IEEE Veh. Technol. Conf.*, vol. 4, May 2002, pp. 1791–1794.
- [10] A. Benjebbour and S. Yoshida, "Novel semi-adaptive ordered successive receivers for MIMO wireless systems," in *Proc. IEEE Int. Symp. Pers. Indoor Mobile Radio Commun.*, vol. 2, Sept. 2002, pp. 582–586.
- [11] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
- [12] R. T. Causey and J. R. Barry, "Blind multiuser detection using linear prediction," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1702–1710, Dec. 1998.
- [13] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996, pp. 203–233.
- [14] C. Belfiore and J. Park, "Decision-feedback equalization," *Proc. IEEE*, vol. 67, no. 8, pp. 1143–1156, Aug. 1979.
- [15] A. Duel-Hallen, "Equalizers for multiple input/multiple output channels and PAM systems with cyclostationary input sequences," *IEEE J. Sel. Areas Commun.*, vol. 10, no. 3, pp. 630–639, Apr. 1992.



communication systems.

Deric W. Waters (S'02) received the B.S. degrees in electrical engineering and computer science from Texas Tech University, Lubbock, in 1999. He studied at Georgia Tech Lorraine, Metz, France, and l'Ecole Supérieure d'Ingénieurs de Marseille, Marseille, France, and received the M.S. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta. Since 2002, he has been pursuing the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology in the area of signal processing for



John R. Barry (M'93–SM'04) received the B.S. degree in electrical engineering from the State University of New York, Buffalo, in 1986 and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1987 and 1992, respectively.

Since 1992, he has been with the Georgia Institute of Technology, Atlanta, where he is an Associate Professor with the School of Electrical and Computer Engineering. His research interests include wireless communications, equalization, and multiuser communications. He is a coauthor with E. A. Lee and D. G. Messerschmitt of *Digital Communications* (Norwell, MA: Kluwer, 2004, Third ed.) and the author of *Wireless Infrared Communications* (Norwell, MA: Kluwer, 1994).