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

Deric W. Waters and John R. Barry

School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, Georgia, 30332-0250

***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, provided that the order in which the inputs are detected is chosen carefully. We show that 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 using linear prediction to reduce its variance. A key advantage of the noise-predictive detector is that it leads to a simple algorithm for optimizing the detection order that is 28% lower in complexity than the lowest-complexity algorithm previously reported.**

I. Introduction

This paper considers the following MIMO channel with N inputs $\boldsymbol{a} = [a_1, \ldots a_N]^T$ and M outputs $\boldsymbol{r} = [r_1, \ldots r_M]^T$:

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

where $\mathbf{H}$ is a complex $M \times N$ channel matrix with linearly independent columns, and where $\boldsymbol{w} = [w_1, \ldots w_M]^T$ is additive white noise satisfying $\mathrm{E}[\boldsymbol{w}\boldsymbol{w}^*] = \sigma^2\mathbf{I}$.

The DF detector emerges as a popular strategy in a wide range of applications for which (1) applies. In the context of a wireless point-to-point link with N transmit antennas and M receive antennas, the DF detector is known as the BLAST nulling and cancelling detector [1]; in the context of a CDMA system with N users and M chips per symbol, it is known as the decision-feedback multiuser detector [2]; and in the context of packet transmission, it is known as a generalized DFE [3].

The performance of the DF detector is strongly impacted by the order in which the symbols are detected. Unfortunately, optimizing the detection order is a difficult problem that often dominates the overall receiver complexity. When the aim is to minimize the joint error probability, the BLAST ordering algorithm of [1] can be used to find the optimal ordering; it is a recursive scheme that suffers from high $O(N^4)$ complexity because it involves repeated computations of a matrix pseudoinverse. Several $O(N^3)$ reduced-complexity ordering algorithms have been proposed. The square-root algorithm of [4] is both numerically stable and low in complexity, while the decorrelating algorithm of [5] is somewhat lower in both complexity and stability. Other algorithms sacrifice optimality in order to reduce complexity [6–8].

In this paper we present a new architecture for implementing the DF detector based on linear prediction. The detector consists of a cascade of a linear MIMO detector followed by a linear prediction mechanism that reduces the noise variance. The noise-predictive detector is functionally equivalent to the conventional DF detector, but it offers two implementation advantages. First, the noise-predictive DF detector can be adapted directly using low-complexity adaptive algorithms, without the need for an intermediate channel estimation step. Second, the noise-predictive DF detector leads to a simple $O(N^3)$ algorithm for determining the optimal detection order that is 28% lower in complexity than the lowest-complexity ordering algorithm previously reported [5].

II. Noise-Predictive DF Detection

To simplify our presentation we focus on the zero-forcing (ZF) DF detector, although the results apply to the minimum-mean-squared-error (MMSE) DF detector as well. Consider first a ZF *linear* detector [9], which computes $\boldsymbol{y} = \mathbf{C}\boldsymbol{r}$ where $\mathbf{C} = (\mathbf{H}^*\mathbf{H})^{-1}\mathbf{H}^*$ is the channel pseudoinverse. From (1), the output of this filter is free of interference:

$$\boldsymbol{y} = \boldsymbol{a} + \boldsymbol{n}\,, \tag{2}$$

where the noise $\boldsymbol{n}$ is no longer white; its autocorrelation matrix is $\mathbf{R}_{\boldsymbol{nn}} = \mathrm{E}[\boldsymbol{n}\boldsymbol{n}^*] = \sigma^2(\mathbf{H}^*\mathbf{H})^{-1}$.

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 are 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 $\{n_1, \ldots n_{i-1}\}$, a linear predictor estimates n_i according to:

$$\hat{n}_i = \sum_{j=1}^{i-1} p_{ij} n_j\,, \tag{3}$$

or equivalently $\hat{\boldsymbol{n}} = \mathbf{P}\boldsymbol{n}$, where $\mathbf{P}$ is a strictly lower triangular *prediction filter* with zeros on the diagonal. 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, once a decision about a_i has been made, it need only be subtracted from y_i to yield n_i, assuming the decision is correct.

The above linear prediction strategy leads to the *noise-predictive DF detector* shown in Fig. 1, where $\boldsymbol{c}_i$ denotes the i-th row of the channel pseudoinverse $\mathbf{C} = (\mathbf{H}^*\mathbf{H})^{-1}\mathbf{H}^*$, and where $\{p_{ij}\}$ are the prediction coefficients. The outputs of the pseudoinverse are permuted according to the ordering algorithm of Section III. Since ordering and detection may be

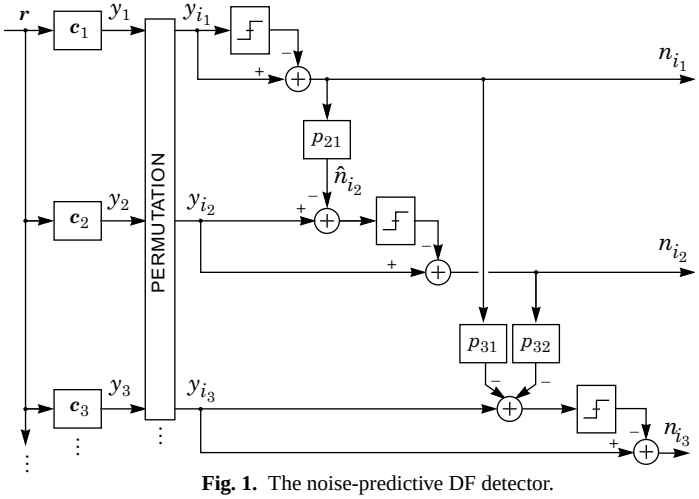

Fig. 1. The noise-predictive DF detector.

separated without loss of generality, we assume an identity permutation in the remainder of this section.

The prediction filter that minimizes the MSE $E[\|\hat{\boldsymbol{n}} - \boldsymbol{n}\|^2]$ is easily expressed in terms of the following QR decomposition:

$$\mathbf{H} = \mathbf{QDM}\,, \tag{4}$$

where $\mathbf{Q}$ is an $M \times N$ matrix with orthonormal columns, where $\mathbf{D}$ is a diagonal matrix with real and nonnegative diagonal entries, and where $\mathbf{M}$ is lower triangular matrix with ones on the diagonal. It is straightforward to show that [10], in terms of this decomposition, the optimal prediction filter is $\mathbf{P} = \mathbf{I} - \mathbf{M}$.

It is easy to show that the noise-predictive DF detector of Fig. 1 is equivalent to the conventional DF detector of [1–3]. The cascade of the pseudoinverse $\mathbf{C}$ and the prediction error filter $\mathbf{E} = \mathbf{I} - \mathbf{P}$ reduces to an effective overall filter $\mathbf{F}$ of:

$$\begin{aligned}\mathbf{F} &= (\mathbf{I} - \mathbf{P})\mathbf{C} = \mathbf{M}(\mathbf{H}^*\mathbf{H})^{-1}\mathbf{H}^* \\ &= \mathbf{M}(\mathbf{M}^{-1}\mathbf{D}^{-2}\mathbf{M}^{-*})\mathbf{H}^* = \mathbf{D}^{-2}\mathbf{M}^{-*}\mathbf{H}^* = \mathbf{D}^{-1}\mathbf{Q}^*.\end{aligned} \tag{5}$$

The output of the prediction error filter is:

$$(\mathbf{I} - \mathbf{P})\boldsymbol{y} = \boldsymbol{a} - \mathbf{P}\boldsymbol{a} + \boldsymbol{e}\,, \tag{6}$$

where $\boldsymbol{e} = \boldsymbol{n} - \hat{\boldsymbol{n}}$ is the effective noise with reduced variance after prediction. From (6), we see that the prediction process has introduced an interference term $-\mathbf{P}\boldsymbol{a}$. The fact that $\mathbf{P}$ is strictly lower triangular allows this interference to be canceled using decision feedback, namely, by feeding past decisions through a feedback filter $\mathbf{B} = -\mathbf{P} = \mathbf{M} - \mathbf{I}$. The filters $\mathbf{F}$ and $\mathbf{B}$ just derived are precisely the forward and feedback filters of the ZF-DF detector, respectively; thus, we conclude that the proposed noise-predictive DF detector of Fig. 1 is functionally equivalent to the conventional DF detector of [1–3].

III. A Low-Complexity Ordering Algorithm

Let i_k denote the index of the k-th symbol to be detected, so that $\{i_1, i_2, \ldots i_N\}$ is a permutation of $\{1, 2, \ldots N\}$. The noise-predictive view leads to a very simple algorithm for finding the optimal ordering that is lower in complexity than previously reported algorithms, making it an attractive choice even if linear prediction is not used for implementation.

As proven in [1], the optimal ordering can be found in a greedy and recursive fashion by choosing each i_k so as to maximize the post-detection SNR, or equivalently minimize the MSE. Specifically, because the MSE for the first detected symbol is $\sigma^2\|\boldsymbol{c}_{i_1}\|^2$, we have $i_1 = \text{argmin}\|\boldsymbol{c}_i\|$. Once i_1 is chosen, the MSE for the second symbol is $E[|n_{i_2} - \hat{n}_{i_2}|^2]$, which reduces to $\sigma^2\|\boldsymbol{c}_{i_2} - p_{21}\boldsymbol{c}_{i_1}\|^2$. When the prediction coefficient p_{21} is optimized, the term $p_{21}\boldsymbol{c}_{i_1}$ reduces to the projection of $\boldsymbol{c}_{i_2}$ onto the subspace spanned by $\boldsymbol{c}_{i_1}$. Hence, the optimal i_2 satisfies $i_2 = \text{argmin}_{i \neq i_1}\|\boldsymbol{c}_i - \hat{\boldsymbol{c}}_i\|^2$, where $\hat{\boldsymbol{c}}_i$ is the projection of $\boldsymbol{c}_i$ onto the subspace spanned by $\boldsymbol{c}_{i_1}$. Repeating this procedure recursively leads to the following simple and succinct description of an optimal ordering algorithm:

$$i_k = \underset{i \notin \{i_1, \ldots i_{k-1}\}}{\text{argmin}} \|\boldsymbol{c}_i - \hat{\boldsymbol{c}}_i\|^2\,, \tag{7}$$

where $\hat{\boldsymbol{c}}_i$ is the projection of $\boldsymbol{c}_i$ onto the span of $\{\boldsymbol{c}_{i_1}, \ldots \boldsymbol{c}_{i_{k-1}}\}$. In words, the best choice for the k-th row is the unchosen row that is closest to the subspace spanned by the rows already chosen.

```
(1)  Perform QR decomposition H = QG
(2)  Initialize C = G^{-1}Q* and E = I
(3)  for k = 1 to N,
(4)      i_k = argmin_{i ∉ {i_1, ... i_{k-1}}} ||c_i||
(5)      v = c_{i_k}   ;   d = ||v||   ;   r = |v_1|
(6)      v_1 = v_1 + d v_1 / r
(7)      C = C − C v* v / (d^2 + dr)
(8)      Delete first column from C; store it as k-th column of B
(9)      for j = (k − 1) downto 1,
(10)         E_{k,j} = −( Σ_{m=j+1}^{k} E_{k,m} B_{i_m,j} ) / B_{i_j,j}
(11)     end
(12) end
```

Fig. 2. Proposed sorting algorithm.

The sorting algorithm could be implemented by applying the Gram-Schmidt procedure to the pseudoinverse rows $\{\boldsymbol{c}_i\}$, but we propose a more efficient implementation as described by the pseudocode of Fig. 2. The algorithm accepts the channel $\mathbf{H}$ as an input, and it produces the optimal ordering $\{i_1, \ldots i_N\}$ as well as the prediction error filter $\mathbf{E} = \mathbf{I} - \mathbf{P}$ as outputs. Lines (9) – (11), which use backsubstitution to find the prediction coefficients, can be omitted if only the detection order is needed. Lines (5) and (6) create a Householder vector $\boldsymbol{v}$ that leads to low complexity by allowing $\mathbf{C}$ to decrease in dimension with each iteration; see line (8). The number of operations required to initialize $\mathbf{C}$ to the pseudoinverse of $\mathbf{H}$ in lines (1) and (2) is $3MN^2 - N^3/3$, while that for line (7) is $2MN^2 - 2N^3/3$, and that for line (10) is $N^3/3$ [11]. The norms in line (4) need not be calculated anew for each k, but can be calculated recursively according to $\|\boldsymbol{c}_i\|_{k+1}^2 = \|\boldsymbol{c}_i\|_k^2 - |B_{i,k}|^2$. The total complexity is thus $5MN^2 - 2N^3/3$, which reduces to $13N^3/3$ when $M = N$. This is significantly lower in complexity than the $27N^4/4$ complexity of BLAST [1], and is 55% less complex than the $29N^3/3$ algorithm of [4], and is 28% less complex than the[1] $6N^3$ algorithm of [5].

References

[1] G. J. Foschini, G. Golden, R. Valenzuela, and P. Wolniansky, "Simplified Processing for Wireless Communication at High Spectral Efficiency," *IEEE JSAC*, vol. 17, No. 11, pp. 1841-1852, 1999.

[2] A. Duel-Hallen, "Decorrelating Decision-Feedback Multiuser Detector for Synchronous Code-Division Multiple Access Channel," *IEEE Trans. on Communications*, Vol. 41, No. 2, pp. 285-290, February 1993.

[3] J. M. Cioffi and G. D. Forney, "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, Chapter 4, pp. 79–127.

[4] B. B. Hassibi, "An Efficient Square-Root Algorithm for BLAST," submitted, *IEEE Trans. Sig. Proc.*, available at http://mars.bell-labs.com.

[5] W. Zha and S. Blostein, "Modified Decorrelating Decision-Feedback Detection of BLAST Space-Time System," *ICC*, paper B04-3, May 2002.

[6] D. Wubben, R. Bohnke, J. Rinas, V. Kugn and K. Kammeyer, "Efficient Algorithm for Decoding Layered Space-Time Codes," *Electronic Letters*, vol. 37, no. 22, pp. 1348-1350, Oct. 25, 2001.

[7] W. Wai, C. Tsui, and R. Cheng, "A Low Complexity Architecture of the V-BLAST System," *IEEE WCNC*, vol. 1, pp. 310-314, 2000.

[8] M. Damen, K. Abed-Meraim and S. Burykh, "Iterative QR Detection for BLAST," submitted to *Wireless Personal Comm.*, Kluwer, August 2000.

[9] S. Verdú, *Multiuser Detection*, Cambridge University Press, 1998.

[10] R. T. Causey and J. R. Barry, "Blind Multiuser Detection using Linear Prediction," *IEEE JSAC*, vol. 16, no. 8, pp. 1702-1710, December 1998.

[11] G. H. Golub, C.F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, 3rd edition, (1996), pp. 203-233.

1. The complexity of [5] breaks down as $5N^3/3$ to initialize $\mathbf{L}_N$ and $\mathbf{L}_N^{-1}$, $10N^3/3$ for the N iterations (i^2 to compute norms for the i-th iteration, and $6Ni$ to triangularize $\mathbf{L}_i$ and $\mathbf{L}_i^{-1}$), and N^3 to multiply $\mathbf{H}^*$ by $\mathbf{L}^{-1}$.