Partial Decision-Feedback Detection for Multiple-Input Multiple-Output Channels

Deric W. Waters and John R. Barry

School of ECE
Georgia Institute of Technology
Atlanta, GA 30332-0250 USA
{deric, barry}@ece.gatech.edu

***Abstract* — The BLAST ordered decision-feedback (ODF) detector is a nonlinear detection strategy for multiple-input multiple-output channels that can significantly outperform a linear detector, at the expense of increased computational complexity. We propose the *partial decision-feedback (PDF)* detector, a simplified version of the ODF detector that only feeds back one decision. The PDF detector reduces complexity significantly compared to the ODF detector while suffering limited performance loss. For example, over a 5×5 Rayleigh fading channel with 64-QAM inputs, the PDF detector is one-third as complex as the ODF detector, yet it requires only 0.5 dB more average signal energy to reach a symbol-error rate of 10^{-3}.**

I. Introduction

Multiple-input multiple-output (MIMO) communications systems have generated a flurry of research recently because of their promise of high spectral efficiency and spatial diversity [1]. The maximum-likelihood (ML) detector minimizes the word-error probability for the MIMO channel, but its complexity increases exponentially with the number of channel inputs and is often prohibitively complex.

The BLAST ordered decision-feedback (ODF) detector [2] achieves only a fraction of the diversity available in the MIMO channel. However, as demonstrated in [3], the ODF detector can still achieve high capacity with low complexity. This motivated the development of various algorithms that reduce the complexity of the ODF detector by an order of magnitude [4–7], as well as algorithms that sacrifice performance in order to reduce complexity further [8,9].

The linear detector [10] is implemented with a single matrix-vector multiplication followed by a slicer. When the channel pseudoinverse is estimated directly rather than computed from an estimate of the channel [10, 11], the linear detector requires an order of magnitude fewer computations than the least complex ODF detector [4]. In fact, the ODF detector can be viewed as a two stage process, where the first stage *is* the linear detection filter [12]. In the second stage, a decision-feedback mechanism improves performance, but also increases complexity.

This research was supported in part by National Science Foundation grants CCR-0082329 and CCR-0121565.

We propose the *partial decision-feedback* (PDF) detector, which functions like the ODF detector, except that it cancels interference from only one symbol decision. This allows the PDF detector to attain an attractive balance between the performance of the ODF detector and the low complexity of the linear detector. Using the noise-predictive implementation proposed in this paper, the PDF detector can achieve nearly the same performance as the ODF detector with significantly fewer computations. In fact, for sufficiently high signal-to-noise ratio (SNR), the word-error rate of the PDF detector approaches that of the ODF detector.

The PDF detector is related to the group detector [13–15], which divides symbols into two groups, and then detects the first group using ML detection. After cancelling the interference due to the first group of symbols, the second group of symbols is detected using a suboptimal technique. The PDF detector can be viewed as a special case of the group detector where the first and second groups are both detected using linear detection. In this paper we focus specifically on the case where the first group contains only a single symbol.

Like the PDF detector, the multiuser detector of [16] also cancels the interference of only a subset of available decisions. It first divides the users into groups according to their signal energies. Then, the detection strategy for each user group is different, but a given user always uses every decision from stronger users for interference cancellation. The PDF detector not only differs in how it orders the users, but it also removes the interference from only a subset of the stronger users.

The remainder of this paper is organized as follows. The PDF detector is presented in Section II. In Section III we show how the word-error probability of the PDF detector approaches that of the ODF detector at high SNR. Finally, in Section IV we use simulations to compare the performance and complexity of the PDF, ODF, and linear detectors.

II. Partial Noise-Predictive DF

This paper considers a 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 $\mathbf{H} = [\boldsymbol{h}_1, \ldots \boldsymbol{h}_N]$ is a complex $M \times N$ channel matrix, and where $\boldsymbol{w} = [\,w_1, \ldots w_M]^T$ is additive white 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 $\mathrm{E}[\,\boldsymbol{w}\boldsymbol{w}^*] = \sigma^2\mathbf{I}$, where $\boldsymbol{w}^*$ denotes the conjugate transpose of $\boldsymbol{w}$. Further, we assume that the inputs are chosen from the same unit-energy alphabet $\mathcal{A}$ and are uncorrelated, so that $\mathrm{E}[\boldsymbol{a}\boldsymbol{a}^*] = \mathbf{I}$.

The PDF detector can significantly reduce complexity relative to the ODF detector while suffering limited performance loss. Before comparing the two detectors, we first describe the PDF detector. Following a noise-predictive implementation [12,4], the PDF detector begins with a permuted version of the zero-forcing linear detection filter:

$$\boldsymbol{y} = \Pi\mathbf{C}\boldsymbol{r}\,, \tag{2}$$

where $\mathbf{C} = (\mathbf{H}^*\mathbf{H})^{-1}\mathbf{H}^*$ is the channel pseudoinverse, and where Π is a permutation matrix that moves the first symbol to be detected into the first row of $\boldsymbol{y}$. The effective front-end filter is $\mathbf{G} = \Pi\mathbf{C}$, which removes intersymbol interference, yielding:

$$\boldsymbol{y} = \tilde{\boldsymbol{a}} + \boldsymbol{n}\,, \tag{3}$$

where $\boldsymbol{y} = [y_1, \ldots y_N]^T$, $\tilde{\boldsymbol{a}} = \Pi\boldsymbol{a}$ is a reordered version of the channel inputs, and where the noise $\boldsymbol{n} = [n_1, \ldots n_N]^T = \mathbf{G}\boldsymbol{w}$ is no longer white; instead, it has autocorrelation matrix $\mathrm{E}[\boldsymbol{n}\boldsymbol{n}^*] = \sigma^2\Pi(\mathbf{H}^*\mathbf{H})^{-1}\Pi^*$.

The first step in the PDF detector is to decide which symbol to detect first, and define Π accordingly. To minimize error propagation, we propose that the symbol with the smallest noise variance be detected first. Since the noise variance of the first symbol is proportional to the squared norm of the corresponding row of the channel pseudoinverse, the index of the first symbol is:

$$i = \underset{j \in \{1, 2, \ldots N\}}{\operatorname{argmin}} \| \boldsymbol{c}_j \|^2, \tag{4}$$

where $\boldsymbol{c}_j$ is the j-th row of $\mathbf{C}$. The permutation matrix Π is then defined by swapping the first and i-th rows of the identity matrix.

The first decision, $\hat{a}_1$, is found by quantizing y_1 to the nearest element of $\mathcal{A}$. Observe that whenever the first decision is correct, $\hat{a}_1 = \tilde{a}_1$, the receiver can recover the first noise sample by subtracting the quantizer output from its input, according to:

$$n_1 = y_1 - \hat{a}_1. \tag{5}$$

Since the other noise samples $\{n_k\}$ are correlated with n_1, we can exploit knowledge of n_1 to predict $\{n_k\}$ for $k > 1$. Let $p_k(\,y_1 - \hat{a}_1)$ denote the predicted value of the noise n_k, where p_k is the prediction coefficient. The PDF detector subtracts this estimate from y_k before making a decision, yielding:

$$\hat{a}_k = \mathrm{dec}\{y_k - p_k(\,y_1 - \hat{a}_1)\}, \tag{6}$$

where $\mathrm{dec}\{x\}$ rounds x to the nearest element of $\mathcal{A}$, and where $p_1 = 0$. Finally, in order to deliver its estimate of $\boldsymbol{a}$, the PDF detector must swap the 1-st and i-th elements of $\hat{\boldsymbol{a}}$.

Just as i was chosen to minimize the noise variance of the first symbol, the best prediction coefficients also minimize the noise variance of the remaining symbols. This criterion leads to a simple equation for calculating $\{p_k\}$. When $\hat{a}_1$ is correct, the noise variance for the k-th symbol reduces to:

$$\begin{aligned}\mathrm{E}[|\, n_k - p_k n_1 \,|^2] &= \mathbf{E}[|\, \boldsymbol{g}_k\boldsymbol{w} - p_k\boldsymbol{g}_1\boldsymbol{w} \,|^2] \\ &= \sigma^2\| \boldsymbol{g}_k - p_k\boldsymbol{g}_1\|^2,\end{aligned} \tag{7}$$

where $\boldsymbol{g}_k$ is the k-th row of $\mathbf{G}$. The noise variance is minimized when the term $p_k\boldsymbol{g}_1$ is the projection of $\boldsymbol{g}_k$ onto the subspace spanned by $\boldsymbol{g}_1$, so the k-th prediction coefficient is given by:

$$p_k = \boldsymbol{g}_k\,\boldsymbol{g}_1{}^*/\|\,\boldsymbol{g}_1\|^2. \tag{8}$$

The noise-predictive ODF detector proceeds in a similar fashion, but it improves performance by using $\{n_1, \ldots, n_{k-1}\}$ along with $k-1$ prediction coefficients to estimate n_k more accurately [4]. Calculating the extra prediction coefficients to achieve this improved noise estimate requires significantly more complexity. We will see later that this extra complexity does not always buy a significant gain in performance.

An efficient implementation of the PDF detector is given in Fig. 1. Assuming that the detector knows the channel pseudoinverse, the total number of complex operations required by the PDF detector per detected word is the sum of the computations in Fig. 1, namely $6MN - 2M - N + 1$. Table 1 compares this complexity to that of the ODF and linear detectors, where we see that the complexity of the PDF and linear detectors increases at a slower rate, $O(MN)$, than that of the ODF detector, $O(MN^2)$.

PDF algorithm: Input: $\mathbf{C}$, $\boldsymbol{r}$; Output: $\hat{\boldsymbol{a}}$	Complexity
(A-1) $\mathbf{G} = \mathbf{C}$	
(A-2) $E_j = \|\boldsymbol{c}_j\|^2,\ j = 1, 2, \ldots N$	$(2M-1)N$
(A-3) $i = \underset{j \in \{1, 2, \ldots N\}}{\operatorname{argmin}} E_j$	
(A-4) swap 1-st and i-th rows of $\mathbf{G}$	
(A-5) $\boldsymbol{y} = \mathbf{G}\boldsymbol{r}$	$(2M-1)N$
(A-6) $\hat{a}_1 = \mathrm{dec}\{y_1\}$	
(A-7) $n = (\,y_1 - \hat{a}_1)/E_i$	2
(A-8) for $k = 2, \ldots N$,	
(A-9) $p = \boldsymbol{g}_k\boldsymbol{g}_1{}^*$	$(2M-1)(N-1)$
(A-10) $\hat{a}_k = \mathrm{dec}\{y_k - pn\}$	$2N-2$
(A-11) end	
(A-12) swap $\hat{a}_i$ and $\hat{a}_1$	

Fig. 1. The partial DF detector algorithm and its complexity.

Table 1: Number of operations per detected word.

Detector	Complexity
ODF	$2MN^2 + N^3/3 + 3MN + N^2 - M - 4N/3 - 4$
PDF	$6MN - 2M - N + 1$
Linear	$2MN - N$

III. Performance Analysis

The word-error rate (WER) of the PDF detector converges to that of the ODF detector at high SNR because the error rate of the first symbol detected dominates the WER of both detectors. In order to see this, let us consider the probability of error on the first symbol compared to the probability of error on the remaining symbols. Let E_j represent the event of an error on the j-th symbol detected, so that $\mathbf{E} = \bigcup_{j=1}^{N} E_j$ represents the occurrence of a word error. For the two detectors, the probabilities of word error are given by the following expressions:

$$Pr[\mathbf{E}\,|\,\mathrm{PDF}] = Pr[E_1|\,\mathrm{ODF}] + Pr[\bar{E}_1|\,\mathrm{ODF}]\,Pr[\mathbf{E}\,|\,\bar{E}_1, \mathrm{ODF}], \quad (9)$$

$$Pr[\mathbf{E}\,|\,\mathrm{PDF}] = Pr[E_1|\,\mathrm{ODF}] + Pr[\bar{E}_1|\,\mathrm{ODF}]\,Pr[\mathbf{E}\,|\,\bar{E}_1, \mathrm{PDF}], \quad (10)$$

where $\bar{E}_1$ is the complement of E_1, and we used the fact that $Pr[E_1\,|\,\mathrm{PDF}] = Pr[E_1\,|\,\mathrm{ODF}]$. In the absence of error propagation, the symbol-error rate of the j-th symbol of the ODF detector has diversity order $M - N + j$ [17], meaning that it decays asymptotically as $\mathrm{SNR}^{-(M-N+j)}$. In (9), this means that $Pr[E_1\,|\,\mathrm{ODF}]$ decays as $\mathrm{SNR}^{-(M-N+1)}$, and further that $Pr[\mathbf{E}\,|\,\bar{E}_1, \mathrm{ODF}]$ decays as $\mathrm{SNR}^{-(M-N+2)}$, as argued in Theorem 1 of [17]. Similarly, since $Pr[\mathbf{E}\,|\,\bar{E}_1, \mathrm{PDF}]$ behaves like the WER of a linear detector applied to an $M \times (N-1)$ channel, it also decays asymptotically as $\mathrm{SNR}^{-(M-N+2)}$. Therefore, the second terms in (9) and (10) converge to zero faster than the first terms:

$$\lim_{\mathrm{SNR}\to\infty} \frac{Pr[\bar{E}_1|\,\mathrm{ODF}]\,Pr[\mathbf{E}\,|\,\bar{E}_1\mathrm{ODF}]}{Pr[E_1|\,\mathrm{ODF}]} = 0\ , \quad (11)$$

$$\lim_{\mathrm{SNR}\to\infty} \frac{Pr[\bar{E}_1|\,\mathrm{ODF}]\,Pr[\mathbf{E}\,|\,\bar{E}_1,\mathrm{PDF}]}{Pr[E_1|\,\mathrm{ODF}]} = 0\ . \quad (12)$$

In other words, the error rate of the first symbol dominates at high SNR. It follows that the WER of the PDF detector converges to that of the ODF detector at high SNR:

$$\lim_{\mathrm{SNR}\to\infty} \frac{Pr[\mathbf{E}\,|\,\mathrm{ODF}]}{Pr[\mathbf{E}\,|\,\mathrm{PDF}]} = 1\ . \quad (13)$$

IV. Numerical Results

In this section, we compare the performance and complexity of the PDF, ODF and linear detectors. We will show that the performance-complexity trade-off depends on the dimensions of the channel, as well as the size of the input alphabet. Although the previous section predicts identical performance for the PDF and ODF detectors at high SNR, we will see that there can be a significant gap at low SNR.

We consider noise-predictive implementations of the PDF and ODF detectors that append add-on processing after the channel pseudoinverse has been applied to the channel output. Therefore, in our comparison we assume that the channel pseudoinverse is known to both detectors. In the simulations shown here, the SNR is taken as the average energy per bit on each receive antenna divided by the noise power: $\mathrm{SNR} = N\,(2\,\sigma^2 \log_2|\mathcal{A}|)^{-1}$.

A. Performance Comparison

In order to compare the performance of the ODF and PDF detectors, we simulated 10^6 Rayleigh fading 6×6 and 6×5 channels with 64-QAM inputs. Fig. 2 shows the average symbol-error rate (SER) curves of the ODF, PDF, and linear detectors as measured on these channels. For the 6×6 channel, the SER of the PDF detector approaches that of the ODF detector at SER = 10^{-3}, as predicted in Section III, while for the 6×5 channel the SER of both detectors fall well below 10^{-4} before converging.

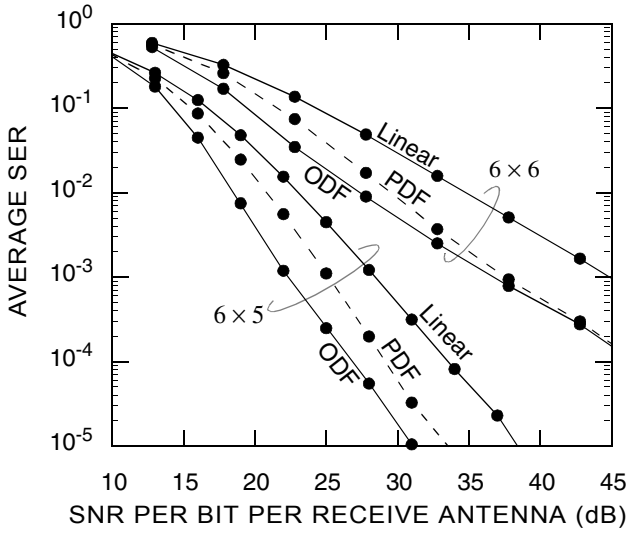

Fig. 2. Overall SER curves averaged over 10^6 different 6×6 and 6×5 Rayleigh fading channels with 64-QAM inputs.

The reason the SER of the PDF and ODF detectors do not converge sooner can be clearly demonstrated by extracting the SER of the i-th symbol. To do so, Fig. 3 shows the SER of the PDF detector for the i-th symbol, $P_1 = Pr[\,E_1\,|\,\mathrm{PDF}\,]$, and the remaining symbols when the i-th symbol was correctly

detected, $P_R = \sum_{j=1}^{N} Pr[\, E_j \,|\, \bar{E}_1, \text{PDF}\,] / N$, as measured over the same 6×6 and 6×5 channels as before. Observe that P_R decreases faster than P_1 for both the 6×6 and 6×5 channels, therefore P_1 will eventually dominate P_R. For the 6×5 channel, the overall SER of the ODF and PDF detectors fall below 10^{-6} before the SNR is sufficiently high for convergence.

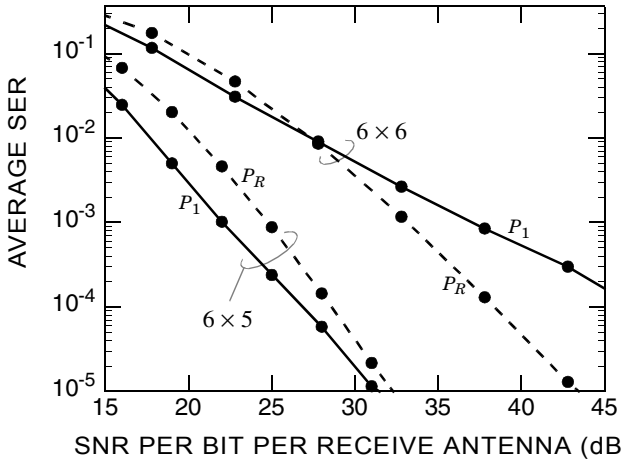

Fig. 3. The SER P_1 of the i-th symbol and the SER P_R of the remaining symbols, both for the PDF detector, the latter assuming no error propagation from the i-th symbol.

B. Performance Versus Complexity

Fig. 4 shows the complexity of the linear, PDF, and ODF detectors for $M \times M$, $M \times (M-1)$, and $M \times (M-2)$ channels, where complexity is taken from Table 1. We see that the PDF detector complexity increases at the same rate as that of the linear detector as M increases, but it is approximately three times as large. On the other hand, the ODF detector is significantly more complex than the PDF detector, even for small M, and its complexity increases at a faster rate.

In order to see how much performance improvement the additional processing of the ODF and PDF detectors delivers, we compare the SNR they require to reach a target SER to that of the linear detector. Fig. 5 shows SNR improvement as averaged over 10^5 realizations of Rayleigh fading channels with the same dimensions considered in Fig. 4, with 4 QAM inputs. We see that the SNR improvement of the PDF detector decreases as the *diversity* $M - N + 1$ of the channel increases. While the SNR improvement of the ODF detector is increasing with M for every channel dimension, the SNR improvement of the PDF detector is increasing with M only for square channels.

In order to see the trade-off between performance and complexity, we can combine the information presented in Fig. 4 and Fig. 5. Fig. 6 shows this performance-complexity trade-off for the same channel dimensions considered before with 4 and 64 QAM inputs, where the performance and

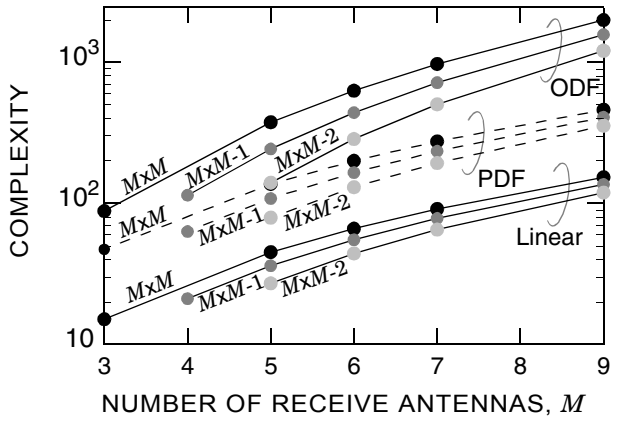

Fig. 4. Complexity of the ODF, PDF, and linear detectors for $M \times M$, $M \times (M-1)$, and $M \times (M-2)$ channels.

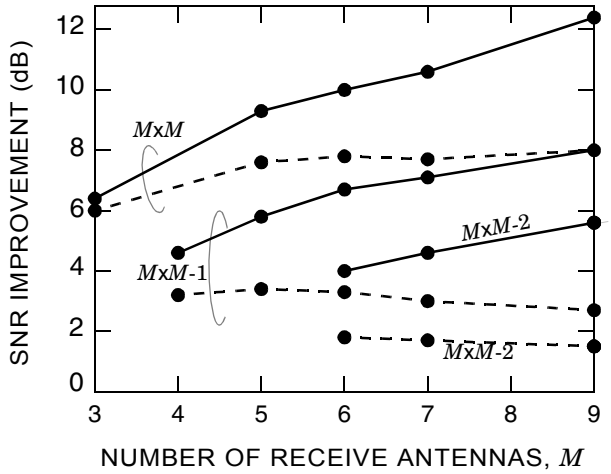

Fig. 5. SNR improvement of the ODF and PDF detectors over the linear detectors for $M \times M$, $M \times (M-1)$, and $M \times (M-2)$ channels with 4-QAM inputs.

complexity are measured relative to the ODF detector. The vertical axis shows the SNR penalty, how much more SNR the linear and PDF detector require than the ODF detector. The horizontal axis shows the complexity of the linear and PDF detectors normalized by the complexity of the ODF detector. This graph allows us to easily see the performance-complexity trade-off between the ODF, PDF, and linear detectors. For example, consider the 5×5 channel with 64-QAM inputs, the graph shows that the PDF detector is about one third as complex as the ODF detector, yet suffers only 0.5 dB of penalty in SNR.

The PDF detector always gives the designer the ability to trade performance for reduced complexity, but in some cases it has a better return. For example, the size of the alphabet affects performance but not complexity. Specifically, for the 6×6 channel, a 4-QAM alphabet incurs 1.5 dB more performance loss than a 64-QAM alphabet but their

complexities are the same. Also, for the PDF detector, the $M \times M$ channels incur less performance loss and decrease complexity more than the $M \times (M-1)$ channels. Specifically, for the 6×6 channel with 64-QAM inputs, the SNR penalty is 0.7 dB and the normalized complexity is 32% for the PDF detector. Meanwhile, for a channel with one less input, the PDF detector suffers an SNR penalty of 2.7 dB and has a normalized complexity of 38%.

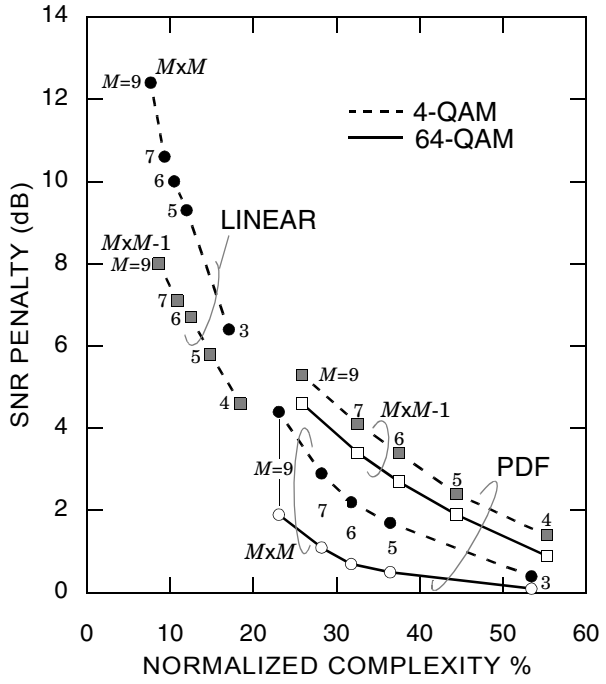

Fig. 6. Average SNR penalty (relative to ODF) versus normalized complexity (relative to the ODF) for the PDF and linear detectors. Averaged over 10^5 Rayleigh fading channels.

V. Conclusion

The partial decision-feedback detector combines the strategies of the BLAST ordered decision-feedback detector and the linear detector. We have shown that, by feeding back only one decision, the PDF detector can significantly reduce complexity while incurring minimal performance loss relative to the ODF detector. This leads to an impressive performance-complexity trade-off. For example, simulations of a 5×5 Rayleigh fading channel with 64-QAM inputs show that the PDF detector is one third as complex as the ODF detector, yet suffers only 0.5 dB of penalty in SNR.

References

[1] G. Foschini, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment When Using Multi-Element Antennas," *Bell Labs Tech. J.*, pp. 41-59, Autumn 1996.

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

[3] P. Wolniansky, G. Foschini, G. Golden, R. Valenzuela, "V-BLAST: An Architecture for Realizing Very High Data Rates Over Rich-Scattering Wirelss Channel", *Int. Symp. on Sig., Sys., and Elec.,* pp. 295 -300, Oct. 1998.

[4] D. Waters, J. Barry, "Noise-Predictive Decision-Feedback Detection for Multiple-Input Multiple-Output Channels," *IEEE Trans. Sig. Proc.*, July 2003, in press.

[5] W. Zha, S. Blostein, "Modified Decorrelating Decision-Feedback Detection of BLAST Space-Time System," *IEEE Int. Conf. Comm.*, vol. 1, pp. 335-339, May 2002.

[6] J. Benesty, Y. Huang, J. Chen, "A Fast Recursive Algorithm for Optimum Sequential Signal Detection in a BLAST System," *IEEE Trans. Sig. Proc.*, vol. 51, no. 7, pp. 1722-1730, July 2003.

[7] B. Hassibi, "An Efficient Square-Root Algorithm for BLAST", *IEEE Int. Conf. Acoust., Sp., Sig. Proc.*, vol. 2, pp. II737-II740, June 2000.

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

[9] W. Wai, C. Tsui, R. Cheng, "A Low Complexity Architecture of the V-BLAST System," *IEEE Wireless Comm. Net. Conf.*, vol. 1, pp. 310-314, 2000.

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

[11] A. Benjebbour, S. Yoshida, "Novel Semi-Adaptive Ordered Successive Receivers for MIMO Wireless Systems," *Proc. IEEE Int. Symp. Pers. Ind. Mob. Radio Comm.*, vol. 2, pp. 582-586, Sept. 2002.

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

[13] Y. Li, Z. Luo, "Parallel Detection for V-BLAST System," *IEEE Int. Conf. Comm.*, vol. 1, pp. 340-344, May 2002.

[14] Varanasi, Aazhang, "Near-Optimum Detection in Synchronous Code-Division Multiple-Access Systems," *IEEE Trans. Comm*., vol. 39, No. 5, pp. 725-736, May 1991.

[15] W. Choi, R. Negi, J. Cioffi, "Combined ML and DFE Decoding for the V-BLAST System," *IEEE Int. Conf. Comm.*, vol. 3, pp. 1243-1248, June 2000.

[16] A. Duel-Hallen, "A Family of Multiuser Decision-Feedback Detectors for Asynchronous Code-Division Multiple-Access Channels," *IEEE Trans. Comm.*, vol. 43, no. 2/3/4, pp. 421-434, Feb./Mar./Apr. 1995.

[17] N. Prasad, and M. K. Varanasi, "Analysis of Decision-Feedback Detection for MIMO Rayleigh Fading Channels and Optimum Allocation of Transmitter Powers and QAM Constellations," *Proc. Allerton Conf. Comm., Control, and Comp*., Univ. of IL., Oct. 2001.