Convergence Behavior Analysis and Detection Switching for the Iterative Receiver of MIMO-BICM Systems

Tao Yang, *Student Member, IEEE*, Jinhong Yuan, *Member, IEEE* , Zhenning Shi, *Member, IEEE* and Mark C. Reed, *Member, IEEE*

Abstract

The iterative receivers with a list sphere detector (LSD) and a parallel interference canceller (PIC) for MIMO-BICM systems are considered. The convergence behaviors of these iterative detection and decoding schemes are analyzed via variance transfer (VT) functions. For ergodic channels, we show that the water fall region of an iterative receiver can be predicted by using a variance exchange graph (VEG). For slow fading channels, we show that the FER of the iterative receiver is essentially limited by the *early interception rate* of the iterative receiver. Moreover, we show that the VT function of an LSD of a small list size is superior to that of PIC at a high variance region whereas it is worse than that of PIC at a low variance region. Then, we propose a detection switching (DSW) approach for the iterative receiver and present a detection switching criterion based on cross entropy. By applying the DSW, we show that the near-capacity performance can be achieved with a significantly reduced complexity.

I. INTRODUCTION

By exploiting multiple-input multiple-output (MIMO) systems as well as space–time coding (STC) techniques, spectral efficiency and diversity of wireless communication can be significantly improved [1]. To achieve the capacities of MIMO channels, the MIMO bit-interleaved-coded-modulation (BICM) systems are widely explored [2]. In order to reduce the complexity of the receiver for a MIMO-BICM system, various iterative detection and decoding schemes are under intensive investigations.

Though ~~it is~~ well-known ~~that a~~ maximum a posteriori probability (MAP) detector is able to achieve the optimal MIMO detection performance in an iterative receiver, the complexity of the MAP detector grows exponentially with the number of transmit antennas. In order to further reduce the complexity of the iterative receiver, some computationally efficient detectors are employed to substitute the MAP detector. Among those, some "list version" non-linear detection approaches, such as list sphere detection [3], list sequential detection [4], and ~~some~~ linear detectors, such as zero-forcing (ZF) detector [5], parallel interference canceller (PIC) [6][18] and minimum mean square error (MMSE) filter [13][14][16], are frequently cited in literature. To date, it is still an open question whether the linear detectors or the non-linear detectors are more appropriate for MIMO-BICM systems in terms of performance and complexity trade-off.

In this paper, we focus on an iterative receiver with a list sphere detector (LSD) and that with a PIC. We employ a semi-analytical approach, namely the *variance transfer* (VT) function, to analyze the convergence behaviors of the iterative receivers. The VT function of the PIC is derived and that of an LSD or a decoders is obtained via numerical evaluations. For an ergodic channel, we show that the "waterfall" region of an iterative receiver can be predicted via a *variance exchange graph* (VEG). We will see that the VEG approach is able to facilitate our understanding of the iterative receiver and it is simpler than the extrinsic information transfer (EXIT) chart [11]. For a slow fading channel, we show that the frame error rate (FER) performance of an iterative receiver is lower-bounded by the corresponding *early interception rate* (ECR).

In addition, we compare the variance transfer characteristic of the PIC and that of the LSD. We show that at a high variance region, the VT function of an LSD of a small list size is superior to that of PIC. At a low variance region, however, we show that the PIC can outperform the LSD with even a large list size. Motivated by the convergence behavior analysis, we propose an iterative receiver which employs an LSD of a small list size for the first few iterations and a PIC for the rest. We refer to this approach as an iterative receiver with *detection switching*. Furthermore, we present a switching criterion based on cross entropy (CE). We show that by employing the detection switching, the performance of the iterative receiver with an LSD of a small list size can be considerably improved. As compared to the scheme with an LSD of a large list size, the proposed receiver can achieve the near-capacity performance with a significantly reduced complexity.

The paper is organized as follows. Section II describes the MIMO-BICM system and an iterative receiver with LSD and that with a PIC. In section III, we analyze the convergence behaviors of the iterative receivers via VT functions and VEGs. In section IV, we present the detection switching and derive the switching criterion based on cross entropy. In section V, we show the simulation results and assess the complexities of various iterative receivers. Finally, section VI presents the conclusions.

II. System Model

A. Transmitter Architecture

The block diagram of a generic MIMO-BICM transmitter is shown in Fig. 1. The message sequence $\mathbf{u}$ is encoded by a channel code with code rate R_c. Then, the coded sequence is de-multiplexed into n_T streams where n_T is the number of transmitter antennas. A cyclic-shifter is employed at the de-multiplexer for spatial interleaving [2]. Then, each stream is independently inter-leaved, modulated (to a constellation of 2^μ points) and transmitted by a separate antenna. Neglecting the time index, we denote x_i the transmitted symbol from antenna i, $i \in \{1, 2, ..., n_T\}$, and use a symbol vector $\mathbf{x} = [x_1, ..., x_{n_T}]^T$ to represent the transmitted symbols from all the antennas. Each symbol vector carries $K = \mu \cdot n_T$ coded digits which are denoted as $\mathbf{c} = [c_1, c_2, ..., ..., c_l, ..., c_K]^T$, $c_l \in \{-1, 1\}$ [1].

Let $\mathbf{H}$ represent the channel matrix of size $n_R \times n_T$ where n_R is the number of receiver antennas. Its jith element, denoted as h_{ji}, is the complex fading coefficient from transmitter antenna i to receiver antenna j, $j \in \{1, 2, ..., n_R\}$ and it follows a Rayleigh distribution. Without losing generality, we assume that $E[|h_{ji}|^2] = 1$. Furthermore, we assume that there are no correlations between antennas so that all the elements in $\mathbf{H}$ are independent. After coherent detection, the received signal vector $\mathbf{r} = [r_1, ..., r_{n_R}]^T$ can be written as

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

where $\mathbf{n} = [n_1, ..., n_{n_R}]^T$ and each entry is a sample of a complex memory-less additive white Gaussian noise (AWGN) of a zero mean and a variance σ_n^2 in each real dimension.

B. Iterative Receivers for MIMO-BICM

Now, we briefly depict the iterative receiver with an LSD detector and that with a PIC detector for the convenience of the subsequent analysis. The details of these receivers can be found in [3][6].

1) An Iterative Receiver with a List Sphere Detector: The block diagram of an iterative receiver is shown in Fig. 2. In every iteration, the detector incorporates soft information provided by the channel decoder to perform processing in spatial domain. The channel decoder operates with the soft information given by the detector to carry out time-domain processing. The soft information is exchanged between the detector and the decoder in an iterative manner and the operation of the receiver is usually referred to as iterative detection and decoding (IDD).

Instead of performing full a posteriori probability (APP) detection [3], the LSD generates the extrinsic information from a portion (list) of all the 2^K candidates. In the kth receiver iteration, the *extrinsic LLRs* of the detector's output for the lth coded digits is denoted by $\Lambda_E^k(c_l|\mathbf{r})$ and it can

[1] l is the index of the coded digits contained in one symbol vector (not a time index).

be approximated as [3]

$$\Lambda_E^k(c_l|\mathbf{r}) \approx \frac{1}{2}\max_{\mathbf{c}\in\pounds\cap\mathbb{C}_{l,+1}}\left\{-\frac{1}{\sigma_n^2}||\mathbf{r}-\mathbf{H}\mathbf{x}||^2+\mathbf{c}_{[l]}^T\cdot\mathbf{\Lambda}_{A,[l]}^k\right\}$$
$$-\frac{1}{2}\max_{\mathbf{c}\in\pounds\cap\mathbb{C}_{l,-1}}\left\{-\frac{1}{\sigma_n^2}||\mathbf{r}-\mathbf{H}\mathbf{x}||^2+\mathbf{c}_{[l]}^T\cdot\mathbf{\Lambda}_{A,[l]}^k\right\} \tag{2}$$

where $\mathbb{C}_{l,b}$ is the set made up of 2^{K-1} bit vectors $\mathbf{c}$ having $c_l = b$, $b \in \{-1, 1\}$, and $\pounds$ is a list of candidates which contains the maximum likelihood (ML) estimate and its $N-1$ neighbors of smallest value of $||\mathbf{r}-\mathbf{H}\mathbf{x}||^2$. In (2), $\mathbf{c}_{[l]}$ is the sub-vector of $\mathbf{c}$ obtained by removing its lth element c_l. Also, $\mathbf{\Lambda}_{A,[l]}$ is the column vector made up of the *a priori* LLR values associated with $\mathbf{c}_{[l]}$, which are obtained from the decoder's *extrinsic* output. The list of candidates $\pounds$ is efficiently searched by employing the sphere decoding algorithm (SDA) [3].

In each receiver iteration, the extrinsic output of the LSD is demodulated, de-interleaved, multiplexed (spatially de-interleaved) and forwarded to the decoder. The decoder generates the extrinsic LLRs on the coded digits. This soft information is de-multiplexed (spatially interleaved) interleaved, modulated and feedback to the detector to carry out MIMO detection for the next iteration.

2) An Iterative Receiver with PIC: The iterative receiver with a PIC shares a similar structure as the iterative receiver with an LSD. In the kth iteration, the soft outputs of the PIC of the signal from the ith transmitter antenna can be written as

$$y_i^k = x_i^k\sum_{j=1}^{n_R}|h_{ji}|^2 + \sum_{\substack{m=1\\ m\neq i}}^{n_T}(x_m^k - \widetilde{x}_m^{k-1})\sum_{i=1}^{n_R}h_{ji}^*h_{jm} + \sum_{j=1}^{n_R}h_{ji}^*n_j \tag{3}$$

where $\widetilde{x}_m^{k-1}$ is an estimate on the transmitted symbol and this symbol-estimate is obtained by modulating the estimated coded digits from the decoder. The estimate on a coded digit is obtained from the decoder and it is computed by $\widetilde{c_l} = (+1)p(c_l = 1) + (-1)p(c_l = -1) = \tanh\left[\frac{\Lambda_{P,l}}{2}\right]$. Here, $p(c_l = \pm 1)$ is the a posteriori probability (APP) and $\Lambda_{P,l}$ is the corresponding a posteriori LLR from the decoder. Generally, APP feedback is utilized in a scheme with interference cancellation [6][18].

In (3), the first term shows that the total channel gain (in amplitude) for the signal from transmit antenna i is $\alpha_i = \sum_{j=1}^{n_R} |h_{ji}|^2$. The second term shows the residual interference and the last term is the output noise which has a variance $\alpha_i \sigma_n^2$ for each real dimension. According to the "central limit theorem" (here $n_T \geq 4$ will suffice), the residual interference plus noise follows a Gaussian distribution with variance σ^2. Then, the probabilities on the coded digits (at the PIC's output) can be computed by

$$p(y_i^k = b) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y_i^k - b\alpha_i)^2}{2\sigma^2}\right], \; b \in \{-1, 1\} \tag{4}$$

Since the soft estimates from the decoder are generated from APPs, the soft decision at the output of detector will be biased for $k > 1$. To combat the bias effect, a decision statistic combining (DSC) approach is employed. Please refer to reference [6] or [17] for the details of DSC.

III. Convergence Behavior of Iterative Detection and Decoding

A. *Introduction of Variance Transfer Functions*

In order to facilitate the design of an iterative receiver, we employ the variance transfer functions of the detector and decoder to analyze the convergence behavior. Depending on different type of detectors, there are two kinds of variances to be utilized.

The first type of variance under consideration is the *observation variance* (OV) which is the normalized variance of the noisy observation of the signal. This type of variance can be directly employed to analyze the linear detectors and it is exactly the inverse of the signal to interference-plus-noise ratio (SINR). Assuming that the Gaussian consistent condition [11] is met, the observation of the signal y can be viewed as the summation of symbol c of unitary power and a Gaussian noise sample n with variance σ_o^2, that is $y = c + n$. Given the OV, the relationship between signal y and its LLR value is given by

$$\Lambda(y) = \log\left[\frac{P(y|c=1)}{P(y|c=-1)}\right] = \frac{2y}{\sigma_o^2} \tag{6}$$

The second type of variance is the *bit variance* (BV) which is defined as

$$\sigma_b^2(y) = E\left(|y - \widetilde{y}|^2\right) = E\left[\left|y - \tanh\left(\frac{\Lambda(y)}{2}\right)\right|^2\right] \tag{7}$$

where $\widetilde{y}$ is ~~an~~ estimate on y ~~and it is obtained from its~~ LLR value $\Lambda(y)$, that is $\widetilde{y} = \tanh\left[\frac{\Lambda(y)}{2}\right]$. This type of variance is adopted to analyze a non-linear component such as an LSD or a decoder. The relation between BV and OV can be acquired by combining (6) and (7).

By computing the BVs (or OVs) of the input and output signals of a component, we are able to depict its *variance transfer (VT)* function. Next, we will use the above two type of variances to find the VT functions of the detectors where 4-QAM modulation is considered.

B. VT Function of the Detectors

In this paper, we consider that there is a single decoder and we assume perfect interleaving. Thus, the signals from all the transmit antennas are i.i.d in the iterative processing. In addition, the statistics for the real part and imaginary part of the signals are also i.i.d and we may only focus on the real part of the signal. Based on these conditions, we define the BV of the real part of the signal from any transmit antenna as $\sigma_b^2\left[\mathrm{Re}\left(x^k\right)\right]$ and we have

$$\sigma_b^2\left[\mathrm{Re}\left(x^k\right)\right] = E\left[\mathrm{Re}(x_1 - \widetilde{x}_1^{k-1})^2\right] = \ldots = E\left[\mathrm{Re}(x_{n_T} - \widetilde{x}_{n_T}^{k-1})^2\right] = \sigma_b^2\left[\mathrm{Im}\left(x^k\right)\right].$$

For an ergodic channel, in which the instantaneous channel matrix varies over symbol vectors, the OV of the real part of the PIC's output signal is

$$\begin{aligned} \sigma_o^2\left[\mathrm{Re}(y_i^k)\right] &= E\left\{\frac{\sigma_b^2\left[\mathrm{Re}\left(x^k\right)\right]\sum_{m=1,m\neq i}^{n_T}\left|\sum_{j=1}^{n_R} h_{ji}^* h_{jm}\right|^2 + \sigma_n^2\left|\sum_{j=1}^{n_R} h_{ji}^*\right|^2}{\left[\sum_{j=1}^{n_R}|h_{ji}|^2\right]^2}\right\} \\ &= \frac{(n_T-1)\sigma_b^2\left[\mathrm{Re}\left(x^k\right)\right] + \sigma_n^2}{n_R}. \end{aligned} \tag{8}$$

where the last step is valid since we assume the channel coefficients are i.i.d and $E[|h_{ji}|^2] = 1$. It is clear that the OV of PIC's output is a linear function of the PIC's input BV. Therefore, the close-form expression of (8) is referred to as an variance transfer function. For the ideal case with perfect feedback[2], we see from the VT function that the PIC should achieve the optimal interference-free performance.

[2]It means that the decoder output is completely reliable that $\sigma_b^2\left[\mathrm{Re}(x^k)\right] = \sigma_b^2\left[\mathrm{Im}(x^k)\right] = 0$.

The VT function of an LSD in close-form is hard to derive due to the non-linear operation of the detector. Therefore, we will obtain it from numerical evaluation, where extremely long block length is used in order to achieve a relatively high accuracy. Now, we consider the case with $n_R = n_T = 4$ and 4-QAM modulation. Fig. 3 presents the derived VT function of PIC, where the output OV is transformed into BV according to (6) and (7). The numerically evaluated VT curves of LSD of various list sizes ($N = 8, 16, 32, 64, 128, 256$) are also plotted. It is worth noting that the LSD with a full list size ($N = 256$) is equivalent to an MAP detector. Comparing the VT functions of LSDs and that of the PIC, we observe that the LSD of $N \geq 8$ has a better VT characteristic than PIC at a high variance region, whereas the PIC is of a preferable VT function than an LSD ($N \leq 256$) at a region of very low variance.

The different VT characteristics of LSD and PIC detectors can be explained from (2) and (3). From (2) we see that at a high input variance region, the soft information from the decoder is not so reliable that the a priori LLR values to the LSD are very small. Hence, the term $\mathbf{c}_{[l]}^T \cdot \mathbf{\Lambda}_{A,[l]}^k$ in (2) is not very significant (not sensitive to the reduction in the list size) and the detector is close to an optimal MAP detector. For PIC, on the other hand, the residual interference is very large at a high input variance region and the operation is close to a matched filtering process. Since the interference cancellation with matched filtering does not consider the interference structure of the signal [8][14], the performance of the matched filtering is far from that of the optimal MAP detector at a regime with high interference.

At a very low input variance region, the a priori LLR values to the LSD are so large that the summation of $\mathbf{c}_{[l]}^T \cdot \mathbf{\Lambda}_{A,[l]}^k$ dominates the computation of (3). As a result, the accuracy of the generated extrinsic output is very sensitive to the list size of LSD. Thus, the VT of an LSD with a small list size will stray away from that with a full list size. On the other hand, we have shown in the last section that the PIC is able to achieve the optimal performance as the input variance approaches zero. Therefore, we claim that at a low input variance region, the PIC has a better transfer property than the LSD.

The above analysis motivates a *detection switching* in the iterative receiver which will be presented in Section IV.

C. VEG of Iterative Detection and Decoding

By plotting the VT function (where BVs are used) of a detector and that of a decoder in a single graph where the axes of the VT curve for the decoder are swapped, we obtain the *variance exchange graph* (VEG) of the iterative receiver. The input variance and the output variance of decoder are denoted by $\sigma^2_{dec,in}$ and $\sigma^2_{dec,out}$. The input variance and the output variance of detector are denoted by $\sigma^2_{det,in}$ and $\sigma^2_{det,out}$.

The VEGs of the iterative receivers in an ergodic channel are shown in Fig. 4 where $E_b/N_0 = 2.25$ dB. The solid curves represent the VT functions of the decoders and the dashed curve is the VT of the LSD ($N = 128$). We observe that the curve of LSD intercepts with that of a turbo decoder (TC) of a generator polynomial $[37, 21]_8$ at $\sigma^2_{dec,in} = 0.85$ and the trajectory stops at this intersection point. Since this intersection happens before the IDD process converges to a very small variance, we refer to this intersection as an *early interception* (EC) between the VT curves of the detector and the decoder. When EC happens, the iterative receiver cannot converge to successful decoding. On the other hand, we observe that the VT curve of the LSD does not intercept with the VT curve of TC $[7, 5]_8$ at $E_b/N_0 = 2.25$ dB. Thus, the iterative receiver with LSD ($N = 128$) and turbo decoder $[7, 5]_8$ is likely to converge to successful decoding. Then, the minimum E_b/N_0 at which EC does not happen is found and we predict that the waterfall region of the iterative detection and decoding should be around 2.25 dB.

For a slow fading channel where the channel matrix remains the same for one transmission block while differing over blocks, there is a probability that the transfer function curve of the detector intercepts with that of the decoder before the IDD converges to a low variance (EC happens). We refer to this probability as *early interception ratio (ECR).* For an iterative receiver, the ECR dictates its performance at high-variance regions (which is relevant to the behavior in the first few iterations) and the achievable FER on a slow fading channel is lower-bounded by the ECR. Since the transfer functions of the decoders are not regular, it is not feasible to obtain the ECR in a close-form. In this paper, the ECR can be obtained by simulation with a very large number of frames.

The ECRs of iterative receivers with PIC-DSC and LSD (with $N = 8, 16, 32, 64, 128$) are shown in Fig. 5. We observe that the iterative receiver with the PIC-DSC has a larger ECR than that with an LSD of a list size greater than 16. Moreover, we observe that as the list size of LSD approaches the full list size, the ECR improvement becomes marginal. This suggests that in order to achieve a relatively small ECR, a large list size is not compulsory.

IV. Detection switching approach for the Iterative detection and decoding

A. Detection Switching Approach

Based on the analysis in the previous section, we introduce a detection switching approach which can improve the performance of the iterative LSD receiver with a reduced complexity. In section III, we notice that although the LSD has a better VT function at a high variance region, the PIC can outperform the LSD at a low variance region. Therefore, it is reasonable to employ an LSD at the first few iterations, where the variances are large. As the receiver iterates, the variance will be gradually reduced. Then, we can replace the LSD with a PIC at subsequent iterations in which the variance is low. In another word, we would like to switch the detector from an LSD to a PIC in the iterative receiver as long as the IDD process works in a low variance region, in which the EC is not likely to happen in the subsequent iterations. Moreover, we also notice from Fig. 5 that the ECR for LSD with $N = 32$ is almost the same as that with $N = 64, 128, 256$. As the ECR determines the behaviors of the iterative receiver in the first few iterations, in which the variance is relatively high, we can use an LSD with a small list size to further reduce the complexity in the scheme with detection switching. The small list size of LSD will not compromise the performance of the iterative receiver since in the subsequent iterations, the detector will be switched to the PIC which has a preferable VT function than the LSD, previously shown in Fig. 3. In this paper, the transition from an LSD to a PIC in the iterative receiver is referred to as detection switching (DSW).

B. Switching Criterion

Simply, the switching can be carried out as soon as the variance is smaller than a pre-defined threshold. However, measuring the BVs requires the knowledge of the transmitted signal which is not available in practice. In this paper, we use cross entropy (CE) which is able to reflect the convergence of iterative processing, to determine when the detector should be switched.

Let us denote the APPs of coded digits from the decoder at the kth receiver iteration by $q^k(c_t)$. The CE between the output APPs of coded digits from two consecutive receiver iterations is represented by

$$T(k) = E_{q^k(c_t)} \left\{ \log \frac{q^k(c_t)}{q^{k-1}(c_t)} \right\} \tag{11}$$

For the first few iterations, the cross entropy of the decoder output APPs can be approximated by [7]

$$T(k) \approx \sum_t \log \frac{q^k(c_t)}{q^{k-1}(c_t)} = \sum_t \frac{\left| \triangle \Lambda^k(c_t) \right|^2}{e^{\left(\left| \Lambda^{k-1}(c_t) \right| \right)}} \tag{12}$$

Although the above approximation is not strict, the threshold can be obtained empirically. Simulation shows that switching at $T(k) < 0.1$ is able to yield good results.

It is worthy noting that for the iterative receiver with an LSD, the extrinsic information from the decoder is fed back to the detector. In the first iteration with PIC, the extrinsic information from the decoder is used for cancellation and hence the output of PIC is not biased. At subsequent PIC iterations, the APPs from the decoder is used for cancellation and hence it is biased. Thus, the DSC is employed after the detection switching.

V. Simulations and Complexity Evaluation

To examine the effectiveness of the VEG approach, which has been visualized in Fig. 4, we plot in Fig. 6 the BERs of the iterative receivers with LSD on an ergodic channel. For the scheme with an LSD of list size $N = 128$, we observe that the waterfall region starts at around 2.2dB. This observation lines up with our prediction by using the VEG (Section III. B). At BER=10^{-5}, a 0.2

dB performance gain is observed by replacing the LSD of list size $N = 128$ with an MAP detector. The performance of the iterative receiver with the LSD ($N = 64$) is 0.15 dB worse than that with the LSD ($N = 128$). The performance of the iterative receiver with LSD ($N = 32$) is more than 1 dB away from that with a MAP detector.

Now, we consider the performance of the iterative receivers on slow fading channels. In the simulation, we start the iteration with an LSD and switch to PIC once the cross entropy $T(k)$ is smaller than 0.1. If the CE does not decrease for a number of consecutive iterations and the condition $T(k) < 0.1$ can not be met, employing the LSD for the subsequent iterations will not improve the performance. In such a case, we perform the detection switching at the end of the fifth receiver iteration since the PIC may have the possibility to outperform the LSD and it is less complicated. The total number of receiver iteration I_D in the simulations is set to ten.

The FERs of the iterative receivers are shown in Fig. 7. We observe that the FER of the iterative receiver with LSD ($N = 16$) is similar to that with a PIC-DSC at FER=10^{-2}. By employing the DSW, however, the iterative receiver is about 0.6 dB better than that with LSD ($N = 16$) or PIC-DSC. Moreover, we see that the iterative receiver with DSW (switching from LSD of $N = 16$ to PIC-DSC) even outperforms that with a LSD of a larger list size ($N = 32$) and its performance is very close to that with a MAP detector. Also, the performance of this scheme is only 1.2dB away from the outage probability of the channel. This approach can be directly extended to a scheme with a larger number of antennas or a higher-level modulation.

Now, we show the complexities of the iterative receivers with and without DSW. Without losing generality, we consider a system with $n_T = n_R$ and 4-QAM modulation. The computational efficiency of the LSD is largely determined by the complexity of generating the list set $\pounds$ as well as computing the soft information. In this paper, we borrow the results in [9] to estimate the complexity of finding the candidate list $\pounds$. In particular, 4 “parallel searchers” is the best choice [9] for a 4 by 4 MIMO system with 4-QAM.

Let I_D be the total number of receiver iterations and Γ be the block length. We use a computation unit (CU) to denote the complexity of one addition or one multiplication. Empirically, each exponential operation costs 5 CUs. Then, the complexity of LSD is

$$\underbrace{\frac{2N\Gamma(n_T M - 1)}{R_c} I_D}_{\text{generating the soft outputs}} + \underbrace{\frac{52 M n_T \Gamma}{R_c}}_{\text{generating the list}}$$

CUs per block and the complexity of PIC-DSC is $\left(\frac{8n_R}{R_c M} + \frac{16}{R_c}\right)\Gamma I_D$ CUs per block. It is clear that reducing the list size of LSD can greatly lower the complexity. Moreover, the complexity of PIC-DSC is almost negligible compared to that of an LSD with a large list size.

The reduction in the complexity by using the DSW is due to the reduced list size as well as the ~~decreased~~ number of iterations in which LSD operations are conducted. For an iterative receiver with DSW, we use I_{LSD} to denote its average number of iterations in which the LSD is conducted[3]. Here, we are interested in comparing the complexity of the scheme with LSD $(N = 32)$ and that with DSW between LSD $(N = 16)$ to PIC-DSC[4]. Numerical results show that compared to the iterative receiver with LSD $(N = 32)$, employing the iterative receiver with detection switching from LSD $(N = 16)$ to PIC-DSC is able to save up to 80% CUs.

VI. Conclusions

In this paper, the convergence behaviors of the iterative receivers for MIMO-BICM systems are analyzed. For ergodic channels, we showed that the water fall region of an iterative receiver can be predicted by using a variance exchange graph. For slow fading channels, we showed that the FER of the iterative receiver is essentially lower bounded by the ECR of the iterative receiver. After investigating the convergence behaviors of the schemes, we proposed a detection switching approach, which is the main contribution of this paper, and showed that the near-capacity performance can be achieved with a significantly reduced complexity.

[3] I_{LSD} is obtained from simulations and the average number of iterations with the PIC-DSC is thus $I_D - I_{LSD}$.

[4] Since the FER curves of these two scheme are close, although the scheme with DSW (N=16) is slightly better.

ACKNOWLEDGEMENT

The work is partly supported by ARC discovery project, UNSW FRG, and National ICT Australia affiliated with the ANU. National ICT Australia is funded through the Australian Government's Backing Australia's Ability initiative and in part through the Australian Research Council.

REFERENCES

[1] G. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas", *Bell Labs Technical Journal,* Autumn 1996, pp. 41-59.

[2] Stefan H.Muller-Weinfutner, "Coding Approaches for Multiple Antenna Transmission in Fast Fading and OFDM", *IEEE transactions on signal processing, Vol. 50, No.10,* October 2002

[3] Bertrand M.Hochwald and Stephan Ten Brink, "Achieving Near-Capacity on a Multiple-Antenna Channel", *Bell Laboratories Lucent Technologies,* Aug. 2001.

[4] S.Baero, J.Hagenauer, and M.Witzke, "Iterative Detection of MIMO transmission using a List-Sequential (LISS) Detector", *in International Conference on Communications (ICC)*, Anchorage, USA, May 2003.

[5] Matthew R.McKay and Iain B.Collings, "Capacity and Performance of MIMO-BICM With Zero-Forcing Receivers", *IEEE Transactions on communications,* Vol 53, No.1, Jan 2005.

[6] Branka Vucetic and Jinghong Yuan, "Space-time Coding", *WILLEY*.

[7] Joachim Hagenauer, Elke Offer and Lutz Papke, "Iterative Decoding of Binary Block and Convolutional Codes", *IEEE transactions on information therory, Vol.42, No.2,* March 1996.

[8] Zhenning Shi, Christian Schlegel, "Joint Iterative Decoding of Serially Concatenated Error Control Coded CDMA", *IEEE journal on selected areas in communications, VOL 19, No.8,* August 2001.

[9] Geoff Knagge, Graeme Woodward, Steven R. Weller, Brett Ninness, "A VLSI Optimised Parallel Tree Search for MIMO", *6th Australian Communications Theory Workshop,* 0-9580345-6-7/05/$20.00 2005.

[10] S. Verdu and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp 622-640, March 1999.

[11] S. T. Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun*, vol. 49, no. 10, pp. 1727-1737, Oct 2001.

[12] T. Yang, J. Yuan, Z. Shi and M. C. Reed, "Detection switching in the iterative receiver for MIMO-BICM systems", *Proc. IEEE ITW 2006,* Oct 2006.

[13] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 1046-1061, Jul. 1999.

[14] Z. Shi and C. Schlegel, "Iterative Multisuser Detection and error Control Code Decoding in Random CDMA", *IEEE Trans. Signal. Proc.,* vol 54, no. 5, pp. 1886-1895, May 2006.

[15] J. Choi, "Iterative receivers with bit-level cancellation and detection for MIMO-BICM systmes," *IEEE Trans. Signal. Proc*, vol. 53, no. 12, pp. 4568-4577, Dec 2005.

[16] U. Madhow and M. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA", *IEEE Trans. Comm.,*Dec 2004.

[17] S. Marinkovic, B. S. Vucetic and J. Evans, " Improved iterative parallel interference cancellation for coded CDMA systems," *Proc*. *IEEE ISIT'2001,* (Washington*, DC*, USA, Jun 24-29, 2001), pp. 34.

[18] R. Milner, L. Rasmussen and F.Brannstrom, "Recursive LLR combining in iterative multiuser decoding of coded CDMA," *Proc. Australian Communication Theory Workshop (AusCTW 2007)*, Feb. 2007.

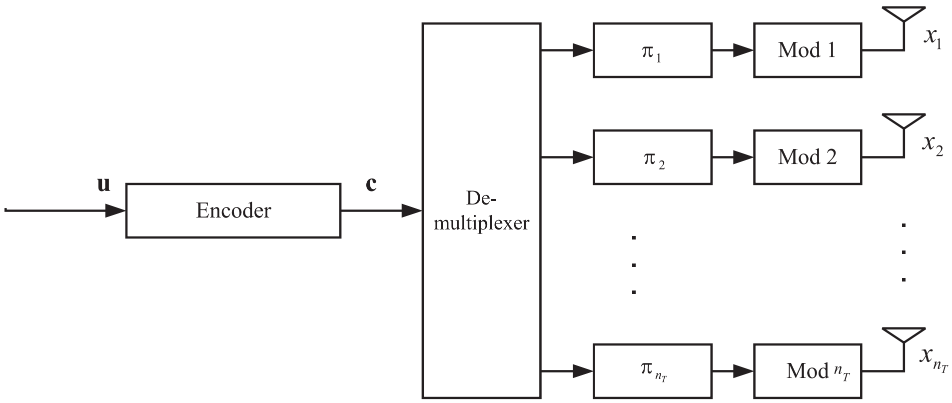

Fig. 1. Transmitter architecture for MIMO-BICM systems. $\pi_1, ..., \pi_{n_T}$ denote the interleavers and "Mod" implies a modulator.

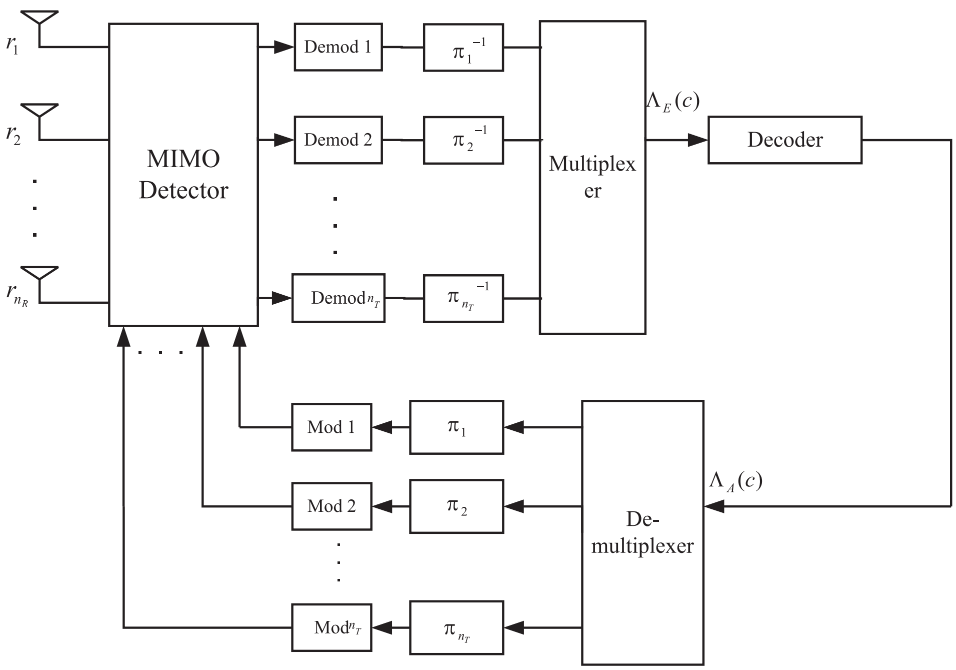

Fig. 2. The architecture of the iterative receivers for MIMO-BICM systems

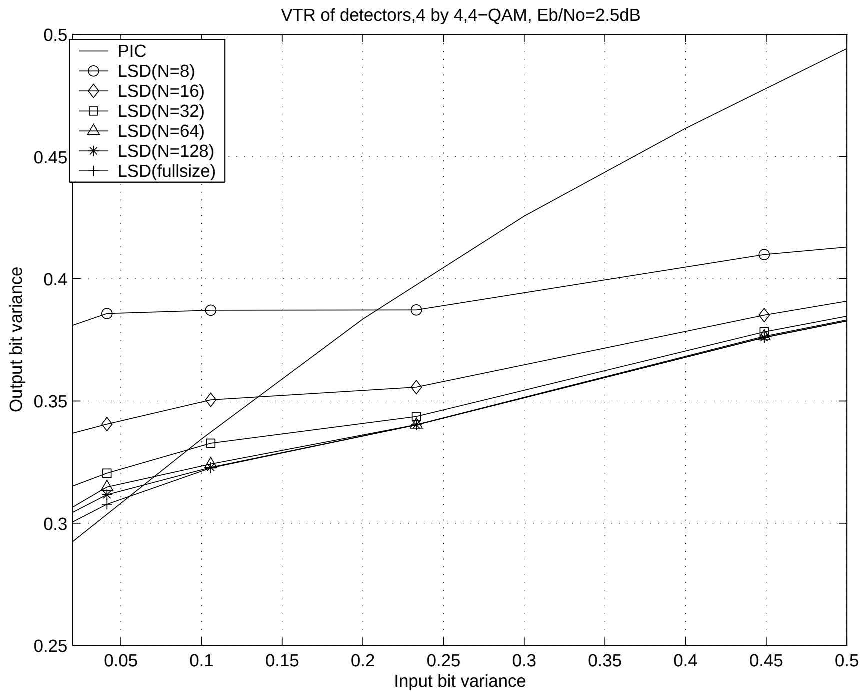

Fig. 3. The VT functions of detectors for a fast fading channel. The full list size of LSD is $2^{n_T \mu} = 256$. To ensure relatively smooth curves, the frame size is set to 16384.

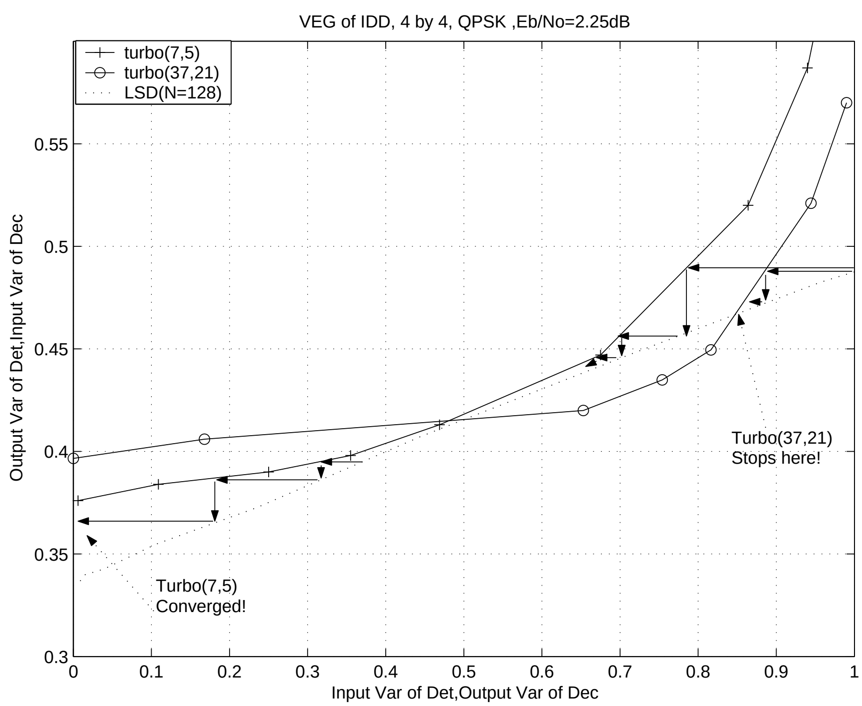

Fig. 4. VEG of IDD,Tx=4,Rx=4, 4-QAM, R_c=1/2. The horizontal axis is for $\sigma^2_{dec,out}$ and $\sigma^2_{det,in}$ and the vertical axis is for $\sigma^2_{dec,in}$ and $\sigma^2_{det,out}$.

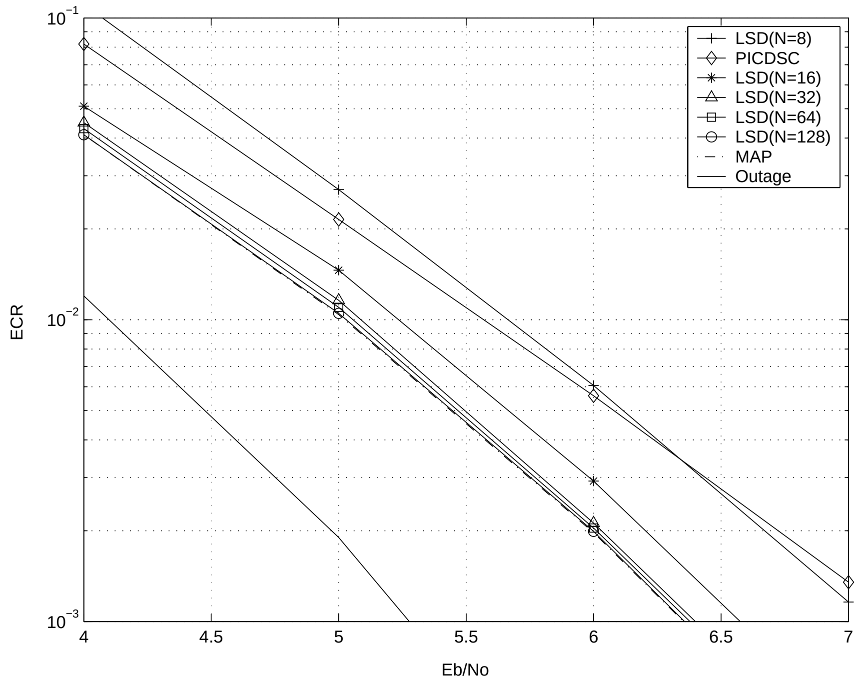

Fig. 5. Early interception ratio of iterative receivers with TC $[7, 5]_8$ and various detectors on a 4 by 4 slow fading channel. Block size=1024 information bits (256 information bits per antenna per block).

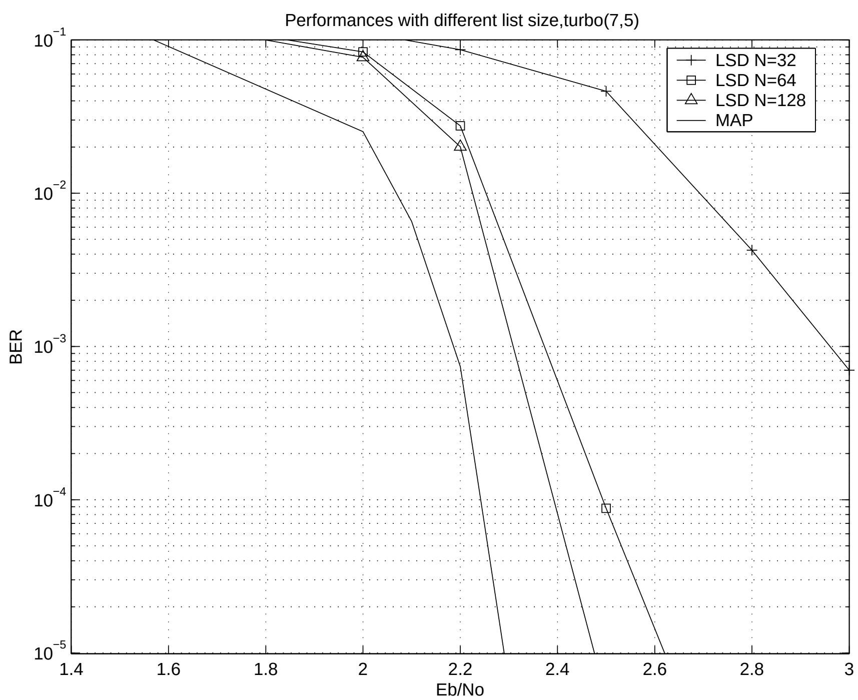

Fig. 6. Performance of MIMO-BICM iterative receiver with LSD of various list sizes on ergodic channels. Block length=8192 and $I_C = 10$ is the number of turbo decoding iterations.

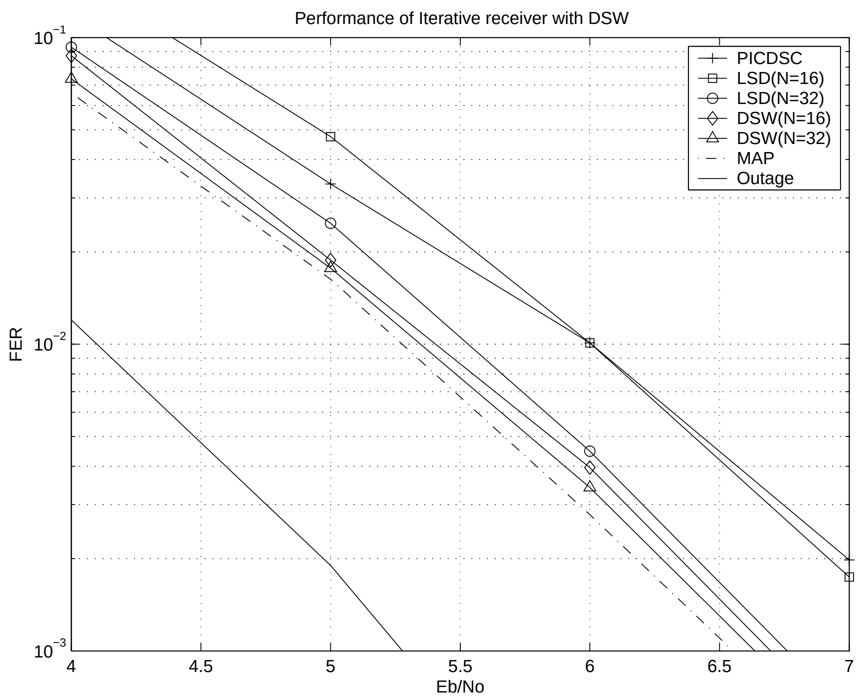

Fig. 7. FER of iterative receiver with DSW and TC $[7, 5]_8$ on a 4 by 4 slow fading channel. The block size is 1024 and 4-QAM with Grey mapping is used.