Detection Switching in the Iterative Receiver of MIMO-BICM

Tao Yang, Jinhong Yuan, *University of New South Wales, Australia*
Zhenning Shi, Mark Reed, *National ICT Australia, Australia*

***Abstract*— The iterative receivers with LSD and PIC on slow Rayleigh fading MIMO channels are considered. The convergence behavior of the iterative detection and decoding is analyzed via VTR. We present an iterative receiver with detection switching and show that this scheme is able to considerably reduce the complexity of receiver without performance degradation.**

I. Introduction

By exploiting multiple transmit and multiple receive (MIMO) antennas as well as space–time coding (STC) techniques, spectral efficiency and diversity of wireless communication can be significantly improved [1]. For fast fading channels, the MIMO-Bit-Interleaved-Coded-Modulation (BICM) systems are shown to outperform the space-time trellis codes provided that a Maximum-Likelihood (ML) receiver is employed [2]. However, the ML receiver is of a complexity exponential to the number of antennas and is hard to implement. Therefore, various sub-optimal iterative receiver structures are proposed in literature.

For coded MIMO-BICM systems, there are mainly two groups [3] of iterative receivers in literature, which are receivers with a posteriori probability (APP) based detection such as sphere detection [4], sequential detection [5] and receivers using interference suppression and cancellation (ISC) techniques such as zero-forcing (ZF) [6], paralell interference canceler (PIC) [7], minimum mean square error (MMSE) filtering[8], etc. In general, an APP based receiver has an advanced performance than a receiver with ISC but requires much greater complexity. To reduce the complexity, some “list version” detection approaches such as list sphere detection [3] and list sequential detection [4] are proposed. However, performance degradation is observed for those receivers. In this background, an iterative receiver with the best performance and complexity trade-off has yet to be found.

In this paper, we focus on an iterative receiver with list sphere detection (LSD) and that with PIC. The channel we considered is a slow Rayleigh fading channel such that the fading coefficient remains the same within each frame while varying over frames.

For analysis, we use a methodology named *variance transfer* (VTR) function to observe the evolution of the error variance [13]. The convergence behaviour of the iterative detection and decoding (IDD) can be exactly predicted in a corresponding *variance exchange graph* (VEG). We show that the frame error performance of an IDD scheme on a slow-fading MIMO channel is limited by its early-interception rate (ECR) which is defined in section III of this paper.

The main contribution of this paper is to present a detection switching (DSW) approach for the iterative receiver whereby the LSD and PIC cooperatively serve as the MIMO detector. We propose a switching criterion based on cross entropy (CE) [11][12] and show that significantly better performance and complexity trade-off is obtained by employing the DSW.

The paper is organized as follows. Section II describes the MIMO-BICM system and iterative receiver with LSD and PIC. The complexity issue is also aroused in this section. In section III, we analyze the convergence behavior of the IDD with the assistance of VTR. In section IV we present the DSW and derive the switching criterion based on the Cross Entropy (CE). In section V, we show the simulation results. Finally, section VI presents the conclusions.

II. System Model

A. Transmitter architecture

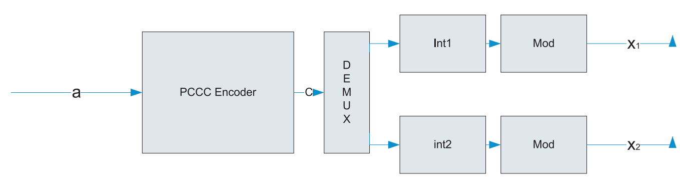

Fig. 1. MIMO-BICM transmitter architecture

The transmitter structure of MIMO-BICM is shown in Fig. 1. The information bits stream $A = [a_1, a_2, ..., a_t, ..., a_\Gamma]$ are encoded by a rate $1/2$ parallel concatennated convolutional code (PCCC) encoder and yield two codeword sequences $C^0 = [c_1^0, c_2^0, ..., c_t^0, ..., c_\Gamma^0]$ and $C^1 = [c_1^1, c_2^1, ..., c_t^1, ..., c_\Gamma^1]$, where Γ is the sequence length or frame size. The codeword sequences are de-multiplexed into n_T streams, where n_T is the number of transmitter antennas and a cyclic-shifter is employed at the demultiplexer for spatial interleaving [8]. Then, each stream is independently interleaved, modulated and transmitted by a separate antenna. We denote $\mathbf{x}_t^i$ the transmitted *symbol* of antenna i at time t. Let $\mathbf{H}$ represent the channel matrix of size $n_R \times n_T$ and its ji-th element $\mathbf{h}_{ji}$ is the complex fading factor from transmitter antenna i to receiver antenna j, where n_R is the number of receiver antennas.

The received complex signal from antenna j at time t is [10]

$$\mathbf{r}_t^j = \sum_{i=1}^{n_T} \mathbf{h}_{ji}\mathbf{x}_t^i + \mathbf{n}_t^j \tag{1}$$

where $\mathbf{n}_t^j$ is a sample of a complex additive-white-gaussian-noise (AWGN) of zero mean and variance σ_w^2 in each dimension. The noise of each receiver antenna is statistically independent and has the same distribution. The received symbol vector at time t is represented as $\mathbf{R}_t = [\mathbf{r}_t^1, ..., \mathbf{r}_t^{n_R}]^T$.

B. Iterative receiver with LSD

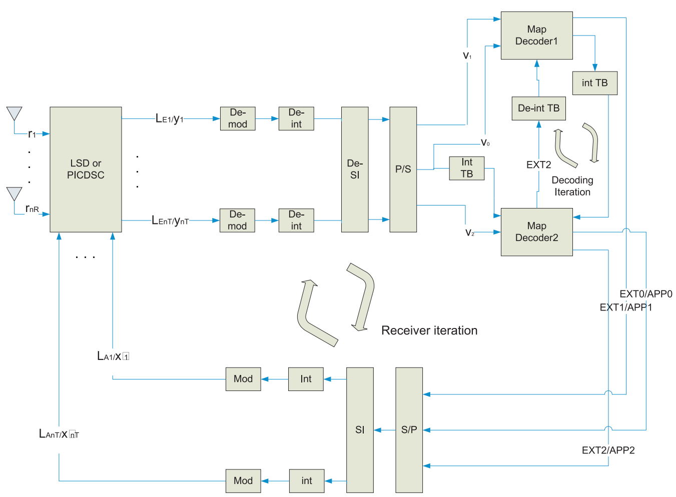

Fig. 2. MIMO-BICM Iterative PICDSC receiver

1) Architecture: The iterative receiver with LSD is shown in Fig. 2. Let us denote $\mathbf{X}_t = [\mathbf{x}_t^1, \mathbf{x}_t^2, ..., \mathbf{x}_t^{n_T}]^T$ where $\mathbf{x}_t^i$ the transmitted symbol of antenna i at time t. The sphere detector (SD) finds the maximum likelihood estimate as

$$\begin{aligned}\widehat{\mathbf{X}}_{\mathbf{t,ml}} &= \arg\min_{\widehat{\mathbf{X}}_{\mathbf{t,ml}}\in\Lambda} ||\mathbf{R}_t - \mathbf{H}\mathbf{X}||^2 \\ &= \arg\min_{\widehat{\mathbf{X}}_{\mathbf{t,ml}}\in\Lambda} (\mathbf{X}_t - \widehat{\mathbf{X}}_t)^T\mathbf{H}^T\mathbf{H}(\mathbf{X_t} - \widehat{\mathbf{X}}_\mathbf{t})\end{aligned} \tag{2}$$

where $\widehat{\mathbf{X}}_t$ is the center of the search sphere that $\widehat{\mathbf{X}}_t = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}$. Λ is the lattice defined by having each entry of the n_T dimensional vector $\mathbf{X}_t$ from a constellation of 2^M consecutive integers [4]. M is the level of the modulation. SD only examines the candidates that lie within a sphere

$$(\mathbf{X}_t - \widehat{\mathbf{X}}_t)^T\mathbf{H}^T\mathbf{H}(\mathbf{X_t} - \widehat{\mathbf{X}}_\mathbf{t}) \leq r^2 \tag{3}$$

with the radius r large enough to contain the solution. By using the Cholesky factorization, we obtain $\mathbf{U}$ which is an upper triangular $n_T \times n_T$ matrix satisfing that $\mathbf{U}^T\mathbf{U} = \mathbf{H}^T\mathbf{H}$. Let us denote entries of the matrix $\mathbf{U}$ as $\mathbf{u}_{ij}$, $i \leq j = 1, ..., n_T$, then equation (3) becomes

$$(\mathbf{X}_t - \widehat{\mathbf{X}}_t)^T\mathbf{U}^T\mathbf{U}(\mathbf{X_t} - \widehat{\mathbf{X}}_\mathbf{t}) = \sum_{i=1}^{n_T} \mathbf{u}_{ii}^2 \left[\mathbf{x}_t^i - \widehat{\mathbf{x}}_t^i + \sum_{j=i+1}^{n_T} \frac{\mathbf{u}_{ij}}{\mathbf{u}_{ii}}(\mathbf{x}_t^j - \widehat{\mathbf{x}}_t^j)\right]^2 \leq r^2 \tag{4}$$

Starting from $i = n_T$, the SD choose a candidate for $\mathbf{x}_t^{n_T}$ satisfying (4) and proceeds to $\mathbf{x}_t^{n_T-1}$ and so forth. If SD reaches the $\mathbf{x}_t^1$, it chooses a value within the corresponding range. Otherwise it goes back to the previous layer and continue the search with another candidate. If all the candidates are searched but no candidate is within the range, a larger radius is chosen and the same operation is repeated.

In order to generate the soft information, SD is revised to generate a list of the candidates $\pounds$ which contains the ML estimate and its $N-1$ neighbors according to (2). This is the list sphere detection (LSD) and N is the list size. The Log Likelihood Ratio (LLR) of the LSD output at the kth iteration is approximated as

$$L_E^k(x_\gamma|\mathbf{R}) \approx \frac{1}{2}\max_{\mathbf{X}\in\pounds\cap\mathbb{Z}_{\gamma,1}} \left\{-\frac{1}{\sigma^2}||\mathbf{R}-\mathbf{H}\mathbf{X}||^2 + X_{[\gamma]}^T\mathbf{L}_{A,[\gamma]}^{k-1}\right\} - \frac{1}{2}\max_{\mathbf{X}\in\pounds\cap\mathbb{Z}_{\gamma,-1}} \left\{-\frac{1}{\sigma^2}||\mathbf{R}-\mathbf{H}\mathbf{X}||^2 + X_{[\gamma]}^T\mathbf{L}_{A,[\gamma]}^{k-1}\right\} \tag{5}$$

where x_γ is the bit of interest and $0 < \gamma \leq n_T M$. $\mathbb{Z}_\gamma, \pm 1$ is the set of $2^{n_T M-1}$ bit vectors X having $x_\gamma = \pm 1$. $X_{[\gamma]}$ denotes the subvector of X except for the γth element x_γ. $\mathbf{L}_{A,[\gamma]}$ is the vector of all L_A values except the LLR values of x_γ. In every receiver iteration, the updated a priori $L_{A,[\gamma]}$ obtained from the decoder at $(k-1)$th iteration is used to enhance the reliability of LSD outputs.

In each detector-decoder iteration (or receiver iteration), the output of the LSD is demodulated, deinterleaved, spatially deinterleaved, multiplexed and fed to the decoder. The decoder generates the soft estimates of the transmitted codeword sequences (LLR of Extrinsic information) and feeds it back to the LSD to perform detection for the next iteration.

C. Iterative receiver with PIC-DSC

1) Architecture: The iterative receiver with PIC has a similar structure to that with LSD and it is also shown in Fig. 2. The most prominent difference of the iterative receiver with PIC from that with LSD is that the APPs of the codeword sequences from the decoder are feedback to the detector [7]. The decision of the PIC output in the k-th iteration at time t for transmitter antenna i is written as [10]

$$\mathbf{y}_t^{i,k} = \mathbf{H}_i^H(\mathbf{R}_t - \mathbf{H}\underline{\widehat{\mathbf{X}}}_{t_i}^{k-1}) \tag{6}$$

where $\mathbf{H}_i$, which is the i-th column of channel matrix $\mathbf{H}$, denotes the channel from the i-th transmitter antenna to all receiver antennas and $(\bullet)^H$ denotes Hermitian operation. $\mathbf{R}$ is the received symbol vector. $\underline{\widehat{\mathbf{X}}}_{t_i}^{k-1}$ is an estimation of symbols of the all transmitter antennas excluding the i-th antenna in the $(k-1)$th iteration, which is given by

$$\underline{\widehat{\mathbf{X}}}_{t_i}^{k-1} = (\widehat{\mathbf{x}}_t^{1,k-1}, ..., \widehat{\mathbf{x}}_t^{i-1,k-1}, 0, \widehat{\mathbf{x}}_t^{i+1,k-1}, ..., \widehat{\mathbf{x}}_t^{n_T,k-1})^T \qquad (7)$$

where T denotes the transpose operation.

To combat the bias effect [9] in an iterative receiver with APP feedback, a decision statistic combining (DSC) is employed. In a system with a large number of interferers (antenna size equals to or greater than 4) and for the APP based symbol estimates, the output of DSC is given by

$$\mathbf{y}_{t,c}^{i,k} = \frac{(\sigma_c^{i,k-1})^2}{(\sigma_c^{i,k-1})^2 + (\sigma^{i,k})^2}\mathbf{y}_t^{i,k} + \frac{(\sigma^{i,k})^2}{(\sigma_c^{i,k-1})^2 + (\sigma^{i,k})^2}\mathbf{y}_{t,c}^{i,k-1} \qquad (8)$$

where $\mathbf{y}_{t,c}^{i,k}$ is the combined symbol and $(\sigma_c^{i,k-1})^2$ is the variance of the combined symbol of $(k-1)$th iteration .

D. Complexity

In this paper, we only compare the number of multiplications required for the detection operation as in [14]. Let us consider a system with antenna size $n_T = n_R$ and QPSK modulation. The complexity of the LSD is largely determined by the complexity of generating the list set $\pounds$ as well as generating the soft information $L_E^k(x_\gamma|\mathbf{R})$. Hereby, the estimation of complexity of $\pounds$ in [14] is employed as our reference (Where z in [14] is set to 16).

LSD		PIC-DSC	
generate $L_E^k(x_\gamma\|R)$	$\frac{(2n_T-1)}{n_T}N\Gamma$	PIC	$(4n_T+4)\Gamma$
genrate $\pounds$ by [17]	$(n_T^2+n_TMz)\Gamma$	DSC	4Γ

TABLE I

COMPLEXITY OF THE DETECTORS

From table. I, we find that the overall complexity of LSD is determined by its list size, which is typically related to the number of antennas whereas the complexity of PIC-DSC is linear to the number of antennas.. For a system with sufficiently large list size N, the complexity of PIC-DSC is almost negligible compared to that of LSD.

III. CONVERGENCE BEHAVIOR OF ITERATIVE DETECTION AND DECODING

In this section, we use *variance* for the analysis of the convergence behavior of an IDD scheme. In this paper, we use *Bit variances* [13] calculated as

$$\sigma^2(\alpha) = E\left[\left|\alpha - \tanh(\frac{L(\alpha)}{2})\right|^2\right] \qquad (9)$$

where $L(\alpha)$ is the Log Likelihood ratio (LLR) of α. By measuring the bit variance of the input and output signals, we are able to depict the curve of the *Variance transfer function* (VTR) of that component. The VTR functions of a detector and a decoder can be combined into a *variance exchange graph* (VEG). If the interleaver size is large enough, a very accurate trajectory can be obtained to predict the convergence behavior of that iterative receiver [13].

A. VTR of detectors

Since the VTR functions of the detectors are variables of the fading coefficients, the VTR properties of the detectors are more appropriate to be studied under a fast fading channel rather than a slow fading channel.

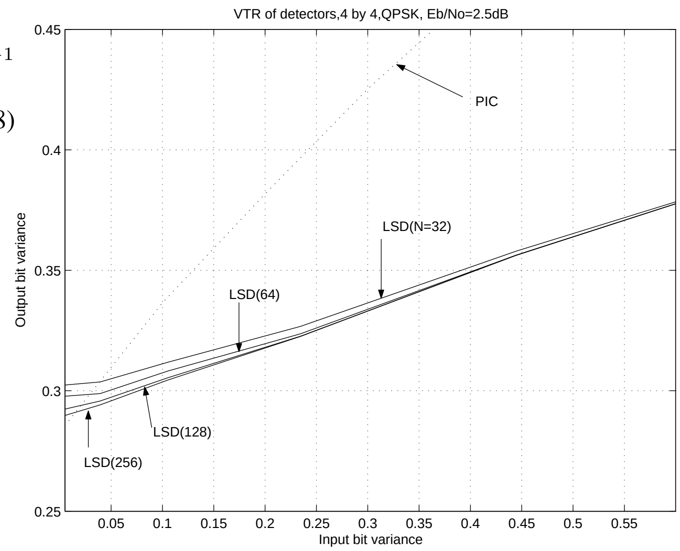

Fig. 3. VTR of detectors for a fast-fading channel,Tx=Rx=4, QPSK, Eb/No=2.5dB

Fig. 3 presents the VTR functions of PIC and LSD with varied list sizes. The frame size is 9216 and the E_b/N_o is 2.5dB. QPSK is used and the full list size of LSD is $2^{n_T M} = 256$. From the VTR functions of the detectors, we observe that at relatively high variance region, LSD with List size $N \geq 32$ has better variance transfer properties than the PIC. At very low input variance region. On the other hand, PIC is able to yield smaller output bit variance than LSD with list size $N \leq 128$. Empirically, the above findings suggest that if the IDD for a frame is able to converge to a very low input bit variance region, the iterative receiver with PIC is likely to perform better than that with LSD($N \leq 128$).

B. VEG of iterative detection and decoding

By combining the VTR functions of the PIC and that of the decoder we obtain the VEG of the iterative receiver for a randomly selected frame. In Fig. 4, the horizontal axis stands for the variance of the signal at the decoder output $\sigma^2_{dec,out}$ and that of the detector input $\sigma^2_{\det,in}$. The vertical axis is for $\sigma^2_{dec,in}$ and $\sigma^2_{\det,out}$. In certain scenarios, we observe that at the high variance region the curve of PIC intercepts with that of the turbo code (5, 7). In this paper, we name the intercepetion between the VTR curve of the detector and that of the decoder at high variance region as "*early-interception*"

(EC). Given that the EC happens, the IDD of that frame is not able to converge to successful decoding.

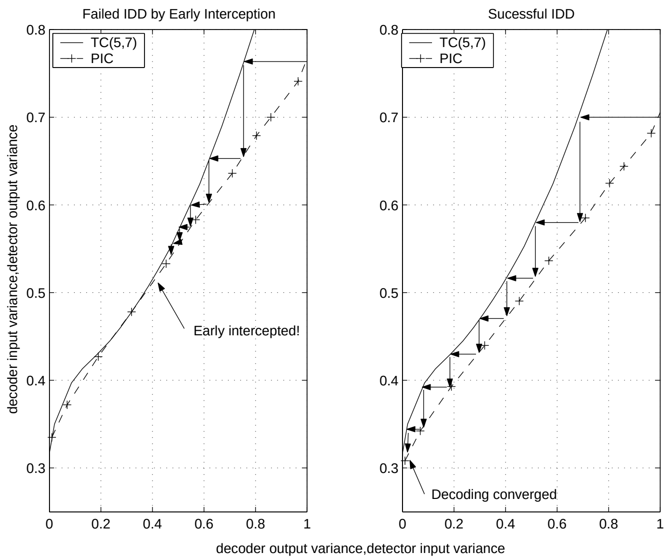

Fig. 4. Illustration of the VEGs for randomly selected frames on a slow fading channel, with and without early interception

The ECR is an essential parameter for the study of the convergence behavior of IDD. Unfortunately, no close-form equations are known to predict the ECR. However, we can obtain ECR through monte carlo simulations on a slow fading channel.

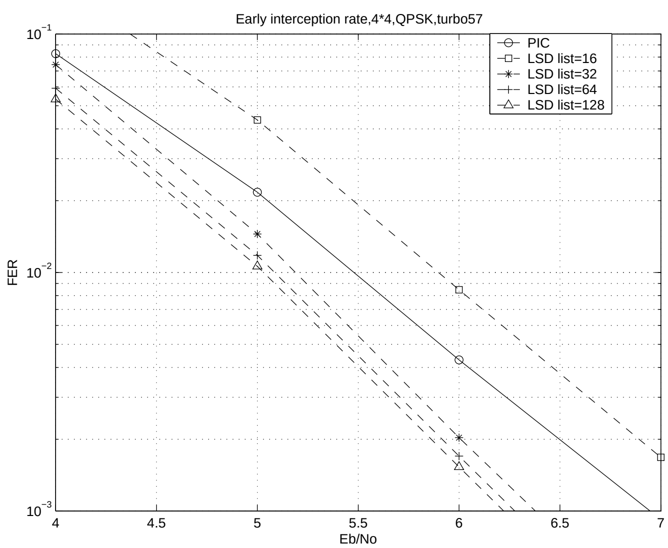

Fig. 5. Early interception Ratio,Tx=4,Rx=4,QPSK,code rate=1/2,receiver iterations=8

Fig. 5 shows the ECR of various IDD schemes. In this part, we consider iterative receivers with PIC and LSD with list size 16, 32, 64 and 128. We find that the PIC has larger ECR than LSD with list size ($N \geq 32$). Moreover, we observe that as the list size of LSD goes up, the improvement on ECR becomes marginal.

IV. Detection switching approach for the Iterative detection and decoding

In the above sections, the VTR attributes of a decoder and detectors are unveiled. In this section, we present a DSW approach which perform switching from the LSD to the PIC. This idea is motivated by the observations in Figure 3, which are concluded as follows:

1. From Fig. 5, we see that the ECR of IDD with LSD ($N = 32$) is very close to that with LSD ($N > 32$). However, its unfavourable VTR property at low variance region, which is shown in Fig.3, results in a considerable performance degradation compared to the receiver with LSD ($N > 32$).

2. From Fig. 3, we observe that the PIC has better VTR property than LSD ($N \leq 128$) at very low detector input variance.

By switching from the LSD ($N = 32$) to PIC as soon as the variance belows a threshold, the unsatifactory VTR property of LSD ($N = 32$) at low variance region is likely to be circumvented. Moreover, the high complexity due to the large list size of LSD($N > 32$) can be avoided.

However, DSW requires the knowledge on FEC decoding outputs. The switching occurs as the decoding statistics converges. To this end, we use the cross entropy (CE)[11][12] of the codeword estimates of two consecutive receiver iterations. In an IDD scheme, the cross entropy can be used to estimate the convergence of the receiver iteration.

Let $p^k(\widehat{u}_t = \pm 1)$ denotes the probability of the codeword extrinsic information at the k-th receiver iteration and let $L^k(\widehat{u}_t)$ be the LLR of that value, where $p^k(\widehat{u}_t = 1) = \frac{\exp[L^k(\widehat{u}_t)]}{1+\exp[L^k(\widehat{u}_t)]}$ and $p^k(\widehat{u}_t = -1) = \frac{1}{1+\exp[L^k(\widehat{u}_t)]}$. The cross entropy of the codeword extrinsic information probatility of two consecutive iterations is represented as [15]

$$T(k) = E_{p^k}\{\log \frac{p^k(\widehat{u}_t)}{p^{k-1}(\widehat{u}_t)}\} \tag{10}$$

$$= -\frac{1}{1+\exp[L^k(\widehat{u}_t)]} \triangle L^k(\widehat{u}_t) + \log \frac{1+\exp[-L^{k-1}(\widehat{u}_t)]}{1+\exp[-L^k(\widehat{u}_t)]}$$

where $\triangle L^k(\widehat{u}_t) = L^k(\widehat{u}_t) - L^{k-1}(\widehat{u}_t)$.

If the IDD converges, we have the following observations:

1) The signs of the LLR values are not likely to change.

2) The magnitudes of the LLR values are very large.

3) $\triangle L^k(\widehat{u}_t)$ has the same sign as $\widehat{u}_t$ and it is of a small value.

Then, the cross entropy can be written as [15]

$$T(k) = \sum_t \frac{\left|\triangle L^k(\widehat{u}_t)\right|^2}{\exp(|\triangle L^{k-1}(\widehat{u}_t)|)} \tag{11}$$

Simply, we start the IDD with LSD and switch to PIC detection once the CE of the codeword extrinsic information feedback is smaller than a threshold [11]. From simulations, we found that threshold value of $T(k) = 0.001T(1)$ turns in good results.. We do not employ the DSC of the PIC in that the bias is not considerable at the stage after the switching.

V. Simulations

A. Frame error performance

In the following simulations, we consider an iterative receiver with turbo code (5, 7). The channel is modeled as a 4×4 slow rayleigh fading channel and the modulation scheme is QPSK. The frame length is 2048. For all the IDD receivers, 8 receiver iterations and 8 turbo decoding iterations are conducted.

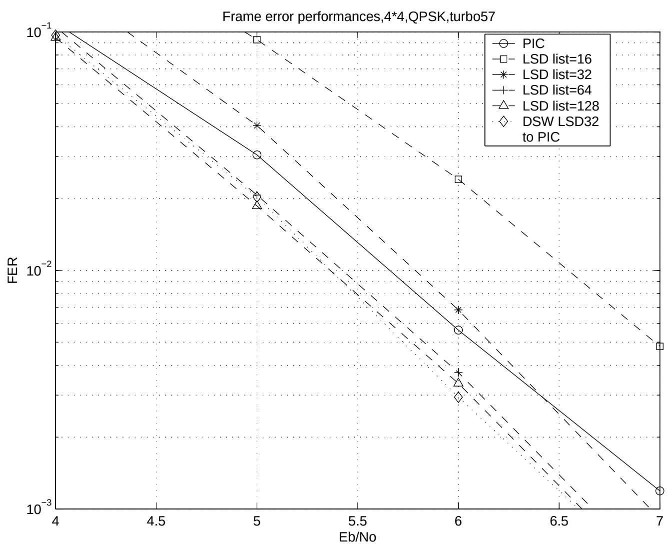

Fig. 6. Frame error rate,Tx=4,Rx=4,QPSK,code rate=1/2 turbo 5,7, frame-size=2048

The frame error ratio of the IDD schemes are shown in Fig.6. We observe that at FER=10^{-3}, the iterative receiver with switching from LSD ($N = 32$) to PIC is 0.4dB better than IDD with LSD ($N = 32$) without switching. Moreover, it is of a slightly better performance to that of the iterative receiver with LSD ($N = 64, 128$). The performance gain over the iterative receiver with PIC-DSC is about 0.6dB at FER=10^{-3}.

B. Complexity

The complexity reduction by introducing the DSW is in light of the reduced list size and the decreased number of receiver iterations with LSD. From simulations we found that the ratio of average number of PIC iterations I_{PIC} and average number of LSD iterations I_{LSD} in the DSW scheme is about 5.75:2.25, where the total number of receiver iterations I is 8. The complexity of the IDD with PIC-DSC, LSD of list size 64 and DSW is shown in Table. II. We see that for the same level frame error performance, the complexity of detection operation of the DSW ($N = 32$) scheme is about 30.4% of the scheme with LSD ($N = 64$). Also, the iterative receiver with DSW($N = 32$) has less complexity than that with LSD($N = 32$) but has better performance.

VI. Conclusions and Future Works

In this paper, a detection switching approach is proposed and is shown to bring about significantly reduced complexity of the detector without performance degradation. In a short future, we will propose a new iteration scheduling by which significantly less decoding operations are required and no performance degradation is incurred.

	Total Multiplication per Frame
PIC-DSC	$(4n_T + 8)\Gamma I = 3.93 \times 10^5$
LSD $N = 32$	$\frac{(2n_T-1)}{n_T} N\Gamma I + (n_T^2 + n_T Mz)\Gamma = 1.21 \times 10^6$
LSD $N = 64$	$\frac{(2n_T-1)}{n_T} N\Gamma I + (n_T^2 + n_T Mz)\Gamma = 2.75 \times 10^6$
DSW(LSD32->PIC)	$(4n_T + 8)\Gamma I_{PIC} + \frac{(2n_T-1)}{n_T} N\Gamma I_{LSD} +$
	$(n_T^2 + n_T Mz)\Gamma = 8.36 \times 10^5$

TABLE II

Total multiplications of the detection operations, QPSK, 4×4 MIMO channel, frame size 2048, 8 receiver iterations

Acknowledgement

The work in this paper is with National ICT Australia and affiliated with the Australian National University. National ICT Australia is funded through the Australian Government's *backing Australia's Ability* initiative and in part through the Australia 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] Tetsushi Abe and Gerhard Bauch,"Comparison of MIMO Receivers: From APP or MMSE? ", in *Turbo-Coding-2006, April 3-7, 2006,* Munich.

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

[5] 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.

[6] 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.

[7] Ka Leong Lo and Slavica Marinkovic, "Performance comparison of Layered Space Time Codes," *0-7803-7400-2/02 IEEE.*

[8] H. El Gamal and A.R. Hammons, "The layered space-time architecture: a new perspective", *IEEE Trans. Inform. Theory,* vol. 47, pp. 2321-2334, Sept. 2001.

[9] Slavica Marinkovic, Branka S. Vucetic and Jamie Evans, " Improved Iterative Parallel Interference Cancellation for Coded CDMA Systems", *ISIT2001, Washington, DC*, June 24-29, 2001.

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

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

[12] Rose Y.Shao, Shu Lin and Marc P.C.Fossorier, "Two simple Stopping Criteria for Turbo Decoding, " *IEEE transactions on communications, Vol 47, No.8,* August 1999.

[13] 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.

[14] 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.

[15] Joachim Hagenauer, Elke Offer and Lutz Papke, "Iterative Decoding of Binary Block and Convolutional Codes", *IEEE transaction on information theory*, vol. 42, No.2, March 1996.