2005

Asterism decoding for Turbo coded MIMO systems

Phillip Conder
University of Wollongong

Tadeusz Wysocki
University of Nebraska-Lincoln, wysocki@uow.edu.au

Asterism decoding for Turbo coded MIMO systems

Phillip Conder, Tadeusz A.Wysocki

Telecommunications and Information Technology Research Institute
University of Wollongong, Wollongong Australia
Email: {pc20,wysocki}@uow.edu.au

Abstract— The area of Multiple Input Multiple Output (MIMO) communications systems has received enormous attention recently as they can provide a roughly linear increase in data rate by using multiple transmit and receive antennas. MIMO combined with an Turbo coding has been shown as a promising way to achieve near capacity for wireless channels.

Where previously described decoders provide linear decoding where the number of receive antennas is at least equal to the number of transmit antennas, are computationally burdensome when $n_t > n_r$. Asterism decoding is a scheme that achieved ML performance for MIMO systems for any number of receive antennas. By applying Asterism decoding to Turbo coded MIMO systems we show that good performance can be achieved for $n_t > n_r$ without the need of additional feedback into the MIMO decoding stage.

I. Introduction

The area of Multiple Input Multiple Output(MIMO) communications systems has received enormous attention recently as they can provide a roughly linear increase [1] in data rate by using multiple transmit and receive antennas. Incorporating a multiple antenna system with error correcting code, such as Turbo coding, is a popular approach to achieving increased data rates.

It was shown in [3], that using such a ML decoder with less receive than transmit antennas could still provide sufficient increase in data rate, hence removing the need for additional receive antennas on a receiving device, such as a mobile terminal.

At the same time as increased data rates, future mobile communications are promised to be small, lightweight mobile devices which present problems to the implementation of multiple antenna systems. Thus reducing the size and cost of mobile terminals by reducing the number receive antennas will be factor in the implementation of MIMO especially for larger number of transmit antennas.

A number of sub-optimal MIMO decoding schemes such as Zero Forcing (ZF) and the Bell Labs Layered Space-Time (BLAST)[2] perform best when the number of receive antennas is greater than the number of transmit antennas, while perform poorly when antenna numbers are equal or less than the number of transmit antennas. While previously described near ML decoders, such as the List Sphere decoder [6], provide linear decoding only where the number of receive antennas is at least equal to the number of transmit antennas are described themselves as "computationally burdensome" when $n_t > n_r$,

while similar Iterative List Sphere decoders [7] do not consider the case of $n_t > n_r$.

Asterism decoding [8] [9] is a scheme that achieved ML performance for MIMO systems for any number of receive antennas, by considering the larger complex constellation created by a multiple transmit antennas and a single receive antenna.

The paper is ordered as follows: Section 2 gives a brief system description of a Turbo coded MIMO system and Section 3 describes decoders such as Maximum Likelihood, Zero Forcing (ZF) and Sphere decoding for $n_t > n_r$. Section 4 describes Asterism decoding for MIMO systems, while Section 5 shows the performance of uncoded and turbo coded MIMO systems with $n_t = 4$ and $n_r = 1, 2, 3$ and 4.

II. Multiple In Multiple Out Systems

The Multiple In Multiple Out approach was first introduced by Lucent's Bell Labs, with their BLAST family of Space Time Code structures [2]. A Vertical Bell Laboratories Layered Space-Time (VBLAST) scheme with Turbo error correction is considered in this paper. A VBLAST (also known as MIMO) system is one where the Turbo encoded bit stream, is interleaved and encoded into complex symbols before de-multiplexed into n_t substreams. Let n_t be the number of transmit and n_r be the number of receive antennas, and $s = (s_1, s_2, ...s_{n_t})^T$ denote the vector of symbols of constellation size C, transmitted in one symbol period. The received vector $R = (R_1, R_2...R_{n_r})^T$ becomes:

$$R = Hs + n \qquad (1)$$

where $n = (n_1, n_2, ...n_{n_r})^T$ is the noise vector of additive white Gaussian noise of variance σ^2 equal to $\frac{1}{2}$ per dimension. The $n_r \times n_t$ channel matrix:

$$H = \begin{pmatrix} h_{1,1} & . & . & h_{1,n_t} \\ . & . & & . \\ h_{n_r,1} & . & . & h_{n_r,n_t} \end{pmatrix} \qquad (2)$$

contains independent identical distribution (i.i.d.) complex fading gains $h_{i,j}$ from the jth transmit antenna to the ith receive antenna. We assume flat fading where the magnitude of the elements of H have a Rayleigh distribution.
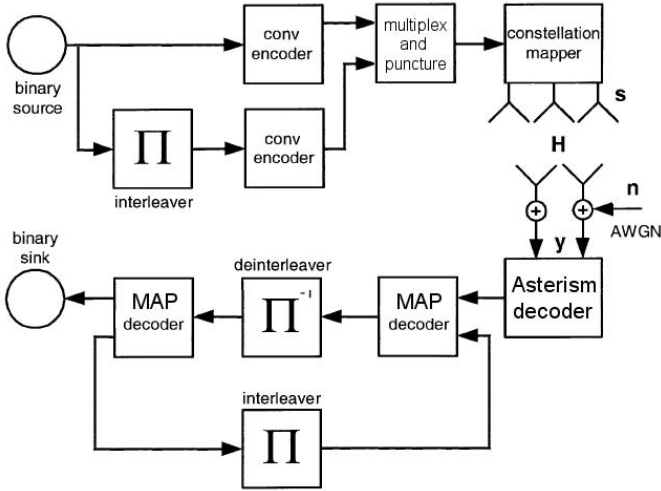
Fig. 1. Simplified block diagram of receiver for $n_t > n_r$ without feedback into the Asterism decoder

A. Turbo encoder

The turbo encoder used in our system is the same structure of [4], with two recursive systematic convolutional (RSC) encoders. One of the encoder is passed the data to be transmitted, while the other passed the same data after it has been permuted by an interleaver. The outputs of the two RSC encoders are multiplexed and punctured (if required). A simplified block diagram of the entire encoding/decoding system described in this paper is shown in Figure 1.

III. PREVIOUSLY MIMO DECODERS FOR $n_t > n_r$

There has been considerable research for decoding MIMO systems, almost all scheme concentrate on systems with the same or more receive antennas than transmit antennas.

Optimal Maximum Likelihood decoding is achieved by minimising

$$\parallel Hs - R \parallel^2 \qquad (3)$$

for all elements of s, which are symbols of constellation of size C. This would produce a search of length C^{n_t}. A system using $n_t = 4$ and $16QAM$ would have to test each of the 65536 possibilities, far beyond being easily implemented on low cost and low power hardware [3].

A. Zero Forcing and VBLAST

The Zero Forcing (or nulling) method of decoding uses a pseudo inverse of H to produce estimates, $\tilde{s}$, of the transmitted symbols. Whereas VBLAST [2] algorithm uses the ZF estimate to successively detect and cancel out transmitted symbols from the receive vector. Each time a detected symbol is cancelled from the received vector the level of diversity increases as the received vector has fewer interferers and better performance is obtained for each detected symbol. Obviously an incorrect symbol selection in the early stages will create errors in the following stages. Therefore the order in which the symbols are detected becomes important to the overall system performance.

Calculation of $\tilde{s}$ becomes difficult for $n_t > nr$ as a square matrix is needed for the pseudo inverse calculation. The channel matrix can be made square by $padding$ the diagonal entries of H with very small numbers:

$$H_{pad} = \begin{pmatrix} h_{1,1} & h_{1,2} & h_{1,3} & h_{1,4} \\ h_{2,1} & h_{2,2} & h_{2,3} & h_{2,4} \\ 0 & 0 & \gamma & 0 \\ 0 & 0 & 0 & \gamma \end{pmatrix} \qquad (4)$$

where γ is a small magnitude constant, we use $\gamma = 10^{-3}$ in our simulations, and by some extra zeros in the received vector:

$$\tilde{s} = H_{pad}^{-1} \begin{pmatrix} R \\ 0 \\ 0 \end{pmatrix} \qquad (5)$$

$\tilde{s}_i$ can be calculated. Error propagation from the early stages of detection is increased due the reduced number of receive antennas, hence a VBLAST decoder performance is far worse than the optimal ML performance.

B. Sphere decoder

The principle of Sphere decoding is to search the closest lattice point to the received vector within a sphere of radius d centered at $\tilde{s}$. The Sphere decoder of [5] decomposes the channel matrix and received vector from complex numbers into its real and imaginary parts to produce a channel matrix of twice the size of the original. The multi-dimensional Sphere decoder now becomes an interval centered around $\tilde{s}$ and a Cholesky factorisation of H is also used to determine the interval size based on the starting radius and any previously decoded symbols.

The Sphere decoder is essentially the same for $n_t > n_r$ but uses the $Padded$ channel matrix H_{pad} to not only determine $\tilde{s}$ but also the Cholesky factorisation. While the $padding$ of H is satisfactory for the calculation of $\tilde{s}$, when used for the Cholesky factorisation, the padded H produces very small numbers for the lowest $n_t - n_r$ levels of the triangular matrix. This means that the Sphere decoder has a complexity of the order of $C^{n_t - n_r}$ i.e. Sphere decoder decomposes to the complexity of ML detection when $n_r = 1$.

IV. ASTERISM DECODING

Asterism decoding, proposed in [8] [9], was created to reduce the computational complexity of Maximum Likelihood decoding and yet retain the performance and flexibility of reducing the number of receive antenna. By considering the larger complex constellation created by a multiple transmit antennas and a single receive antenna.

The large complex constellation generated by (3) for all values of C substituted into s, plotted in Figure 2, can be divided into C smaller groups or Asterisms. Each of these Asterisms can in turn be divided into C smaller Asterisms, and so on for all n_t symbols, as shown in Figure 3.
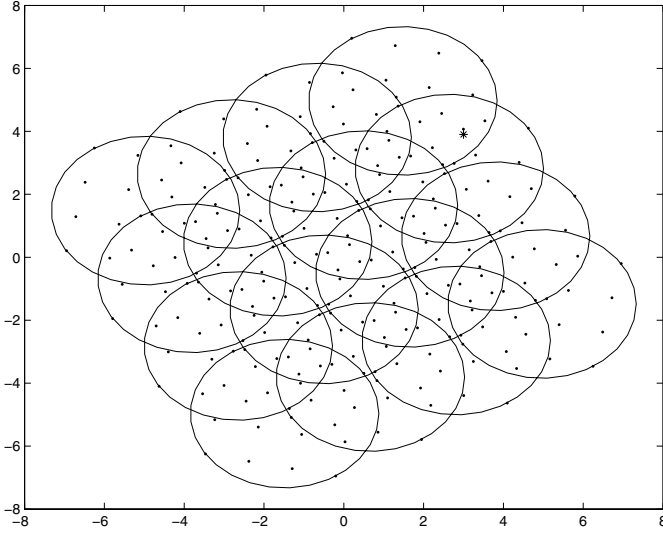
Fig. 2. Complex constellation R grouped into 16 Asterisms.

Fig. 3. Two remaining Asterisms from the complex constellation R that possibly contain the receive vector.

Finding the ML solution without having to test every point by grouping the complex constellation into Asterisms is the main concept behind the Asterism decoding.

For ease of explanation, we make the assumption that the magnitude of H in (3) is decreasing i.e. $|h_1|$ is the largest and $|h_3|$ is the smallest. The radius of the Asterism radius at detection stage k is:

$$Radius_{(k)} = \beta \times \sum_{j=k+1}^{n_t} |h(j)| \qquad (6)$$

where β = largest symbol magnitude, which for 16QAM is $\sqrt{18}$ the magnitude for the symbols $[3+3i, -3+3i, -3-3i, 3-3i]$. These Asterisms at the first detection stage are centered at $h_1 \times s_i$. Every possible combination is covered by these Asterism circles. The size and the amount of overlap of these circles is determined by the number of transmit antennas, the magnitude of the elements of H and the Hamming distance of the constellation.

If the received vector R is inside the one or more circles it is possibly the ML solution. The algorithm then subtracts this possible solution from R and determines whether modified R is in one of the new Asterism circles centered at $h_2 \times s_i$ and of radius $|h_3|$. This recursive process continues until all n_t symbols are found. If there is more than one combination found, the combination with the lowest complex distance measurement is chosen to be the ML solution.

When noise places the received vector outside any of the Asterisms the algorithm will find no symbol combination and fail. This can happen not only at the first stage of decoding, where noise puts R outside the area covered by the largest Asterisms, but also at any later stage of decoding where R may be part of a larger Asterism but not part of a small Asterism of later stages of decoding.

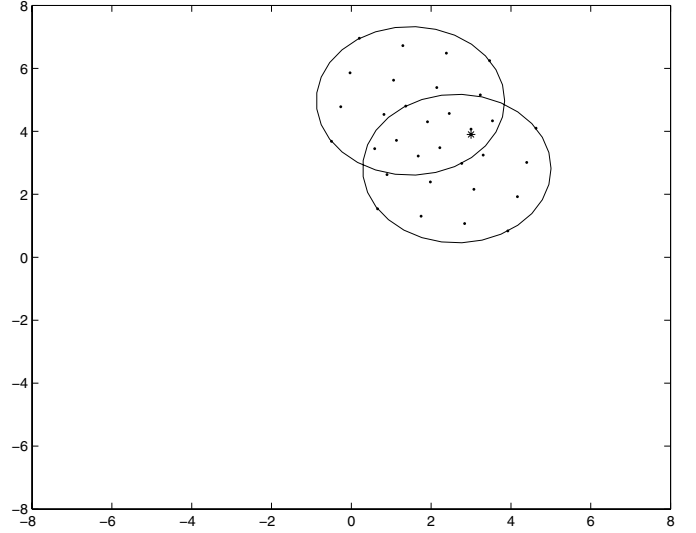To overcome this the decoding the algorithm must not only find which Asterism the received vector is inside, but also allow for the case where it is inside none of the Asterisms. In this case the decoding chooses the Asterism to which it is closest to and continues the process to find the ML solution.

A. Extending to multiple receive antennas

While using Asterism decoding to multiple transmit and a single receiver antenna system produces the ML performance, the performance of this type of system is poor as shown later in Section IV. To overcome this the use of multiple antennas at the receive is now considered. The received vector where $n_t = 3$ and $n_r = 2$ becomes:

$$\begin{bmatrix} R_1 \\ R_2 \end{bmatrix} = \begin{bmatrix} h_{1,1} & h_{1,2} & h_{1,3} \\ h_{1,2} & h_{2,2} & h_{2,3} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_4 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} \qquad (7)$$

To take advantage of the information provided by the additional antenna a number of traditional receive diversity combining schemes, such as equal gain and Maximum Ratio combining, were tested and found to have sub-ML performance. This loss in performance was overcome by applying a modified Maximum Ratio combining at each stage of decoding and was found to have near ML performance. For the first detected symbol the modified Received vectors becomes:

$$R_{m_{1st}} = h_{11}^* R_1 + h_{21}^* R_2 \qquad (8)$$

After the first stage of decoding the modified received vector becomes:

$$R_{m_{2nd}} = h_{12}^* R_1 + h_{22}^* R_2 \qquad (9)$$

this process is continued to detect all n_t symbols are found.

Because the order is which the symbols are detected, the detection order needs to be tested to ensure that the modified Maximum ratio combining does not change the detection ordering and that the symbol with the largest channel magnitude is detected first while the symbol with the smallest channel magnitude is detected last.
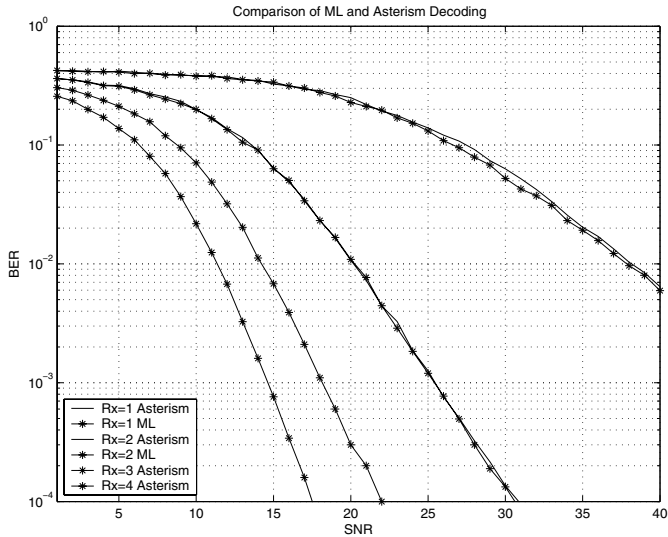
Fig. 4. Performance comparison of Asterism decoding for $n_t = 4$, $n_r = 1, 2, 3$ and 4 using QPSK without error coding.

Fig. 5. Performance comparison of Turbo Asterism decoding for $n_t = 4$, $n_r = 1, 2, 3$ and 4 using QPSK.

V. SIMULATION RESULTS

In this section we provide simulation results, by Monte Carlo methods, to illustrate the performance of Asterism decoding. The channels are assumed quasi static flat-fading, i.e., they are the same over a data burst and changes from burst to burst. The fading coefficients are generated according to a Rayleigh distribution.

It can be seen from Figure 4, that the performance of Asterism decoding is similar to that of ML decoding for $n_r = 1$ and $n_r = 2$, and that the gain between $n_r = 2$ is over $15dB$ over the system with one receive antenna. It is also can been seen that increase in performance by the addition of further receive antennas is less for each additional antenna. For example the system with three receive antennas produces only a gain of approximately $5dB$ and and the addition of a fourth antenna yields only an extra $4dB$.

A. Turbo Decoding

The optimum detector Turbo coded MIMO receiver is well known to be the maximum a posteriori probability (MAP) detector. Unfortunately, a problem with the optimum MAP detector, is that its computational complexity is exponential in the number of bits transmitted simultaneously in each symbol interval.

The turbo decoder used in our simulations consisted of two Log-MAP decoders utilising the Jacobian approximation [10] permuted by a deinterleaver. For our simulation we used a frame of length 9216 bits, two (7,5) convolution coders with puncturing giving a total rate of $1/2$ and all results are for 5 iterations of the Turbo decoder.

Figure 5 shows results of the a $QPSK$ system with $n_t = 4$. The performance of the system quite poor with $n_r = 1$ (approximately 10^{-5} at $33dB$) but with $n_r = 2$ approximately $15dB$ gain is achieved. When $n_r = 3$ the simulation show an
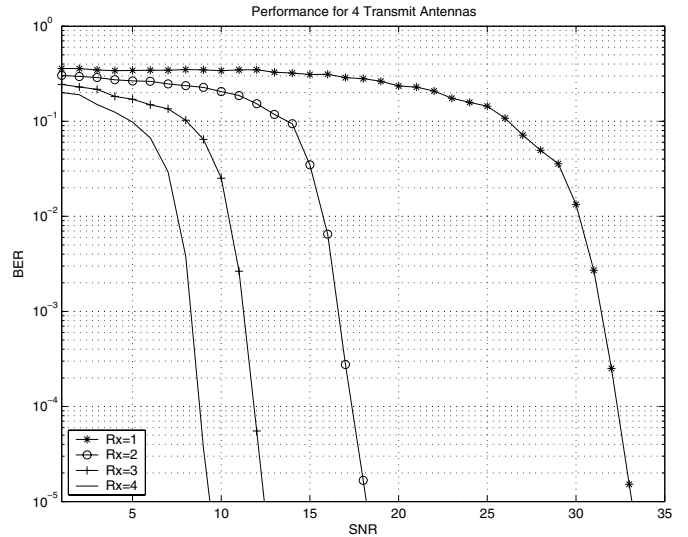
additional $5dB$ gain over the $n_r = 2$ system, while when $n_r = 4$ is only approximately $3dB$ better than $n_r = 3$ results.

VI. CONCLUSION

This paper has extended the previously described method of decoding uncoded MIMO systems called Asterism decoding. The contribution is the decoding of MIMO turbo coding system without the need for additional feedback loops to the MIMO decoder for systems where $n_t > n_r$.

Increasing the performance beyond what has been shown in this paper without significantly increasing the computational complexity is an area of further research.

REFERENCES

[1] E. Telatar, "Capacity of Multi-antenna Gaussian Channels", AT & T Bell Labs, Murray Hill, NJ, Tech. Rep., 1995
[2] G.J. Foschini, "Layered Space-Time Architecture for Wireless Communications in a Fading Environment when using Multiple Antennas", Bell Labs Technical Journal, Autumn 1996.
[3] D. Garrett, L. Davis, G. Woodward, " 19.2 Mbit/s 4 /spl times/ 4 BLAST/MIMO detector with soft ML outputs", Electronics Letters , Volume: 39 Issue: 2 , 23 Jan. 2003.
[4] C. Berrou, A. Glavieux and P. Thitimajshima, "Near Shannon limit Error-Correcting Coding and Decoding: Turbo Codes", ICC 1993.
[5] M. O. Damen, A. Chkeif, and J.-C. Belfiore, "Lattice code decoder for space time codes", IEEE Communications Letters, May 2000.
[6] B. Hochwald and S. ten Brink, "Achieving Near-Capacity on a Multiple-Antenna Channel", IEEE Transactions on Communications, Vol 51, No. 3, March 2003.
[7] Y.L.C. de Jong and T.J. Willink, "Iterative tree search detection for MIMO wireless systems", VTC 2002-Fall, Vancouver, Canada, 2002.
[8] P. Conder and T. A. Wysocki, "Asterism decoding for Layered Space-Time systems using 8PSK", 2004 IEEE Sarnoff Symposium, Princeton, New Jersey, USA, 27-28 April 2004.
[9] P. Conder and T. A. Wysocki, "Extending Asterism decoding to QAM and its complexity in Rician fading MIMO systems", Workshop on IT and Signal Processing 2004 WITSIP '04, Adelaide, Australia, 20-22 December 2004.
[10] L. Hanzo, T. Liew, and B. Yeap, "Turbo Coding, Turbo Equalisation and Space-Time Coding for transmission over fading channels", volume 1., 2002, John Wiley and Sons Ltd.