ABSTRACT

GOEL,VIVEK. Semi-blind turbo detection for Multiple-input Multiple-output wireless systems. (Under the direction of Dr. Huaiyu Dai)

Multiple-input multiple-output (MIMO) techniques constitute the core of future high speed wireless systems. Blind detection in MIMO wireless systems conserves bandwidth and is better suited for use in multiple user scenarios. Since, most of the modern wireless communication systems employ some form of forward error correction, therefore, we can exploit the turbo principle to further improve the performance of blind detection schemes for MIMO.

In this thesis we propose a new semi-blind turbo multiuser detector (MUD) for signal detection in a MIMO wireless system, operating on a single link with Gaussian noise. This turbo MUD (named as T-MUK) performs a sub-optimal joint detection and decoding by iteratively exchanging soft information between the detector stage, that optimizes multiuser kurtosis maximization criterion, and the decoder stage, that runs maximum aposteriori probability (MAP) algorithm. It is shown to achieve good performance at the expense of very few training symbols.

This thesis also introduces a successive interference cancellation based semi-blind MUD for multicell MIMO systems. This MUD uses T-BLAST, which is a near optimal detection structure for single link, to detect desired signals and T-MUK to detect interfering signals and improves the estimate for desired signal by using these in an iterative fashion. Numerical results indicate the advantage of using this MUD structure over T-BLAST (which treats inter-cell interference as additional noise). The semi-blind nature of T-MUK in this MUD allows us to avoid a common problem in multicell systems, that of extracting reliable channel information for interferers using training sequences.

SEMI-BLIND TURBO DETECTION FOR MULTIPLE-INPUT MULTIPLE-OUTPUT WIRELESS SYSTEMS

by

Vivek Goel

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial satisfaction of the
requirements for the Degree of
Master of Science in Electrical Engineering

Department of Electrical And Computer Engineering

Raleigh

2005

Approved By:

Dr. Huaiyu Dai
Chair of Advisory Committee

Dr. Brian L. Hughes

Dr. Winser E. Alexender

Dedication

To my family and dear friends

Biography

Vivek Goel was born in Kotdwar, India on 12th May 1980. He graduated from Indian Institute of Technology Roorkee, India with a Bachelor's degree in Electronics and Communication Engineering in May 2002. He worked at Centre for development of Telematics, Bangalore, India from February 2003 to July 2003 as a Research Engineer. In the fall of 2003, he enrolled in North Carolina State University as a graduate student to pursue Master of Science in Electrical Engineering.

Acknowledgements

I would like to thank my academic advisor, Dr. Huaiyu Dai for introducing me to the exciting area of signal processing for wireless communications and for his constant support and guidance.

Contents

List of Figures

List of Symbols

Symbol	Definition
$\{.\}^T$	Transposition
$\{.\}^*$	Conjugation
$\{.\}^H$	Hermitian transpose (i.e., Conjugate transpose)
$\{.\}^\dagger$	Moore-Penrose Pseudoinverse
$\{.\}^{-1}$	Inverse
$\mathrm{E}\{.\}$	Expectation operator
$\|.\|$	Absolute value (Modulus)
Bold Face Letter	Matrix
Capital Letter	Column or Row vector
$\sigma_{\{.\}}$	Variance
$\mathrm{K}\{.\}$	Kurtosis
sign	Signum function
∇	Gradient operator
$O(.)$	Order of

Chapter 1

Introduction

1.1 Motivation

Recently multiple-input multiple-output (MIMO) communication has emerged as one of the most important concepts in the wireless area. High data rates expected for future broadband wireless access systems, wireless local area networks (WLAN), 3G and 4G cellular systems dictate the use of MIMO techniques. Most of the practical receiver structures developed so far for MIMO communication systems require perfect channel knowledge at the receiver for good performance. This requirement of perfect channel state information (CSI) necessitates use of long training sequences, which results in loss of bandwidth and may not be feasible for fast fading channels. Moreover, in interference-limited cellular environments, it is often advantageous to learn CSI of adjacent-cell signals and detect them to better aid the detection of desired signals. Nonetheless, even if training sequences for interferers are available, much poorer quality is expected for their CSI estimation and detection, due to much lower SNR and asynchronism of the signals from the interferers with respect to the desired signals. These considerations motivate us to study and develop new receiver structures for MIMO communication that can completely eliminate or significantly reduce the length of training sequences. In addition, all modern communication systems employ some form of forward error correction. In such systems joint detection and decoding at the receiver is known to yield best performance, but at a prohibitive computa-

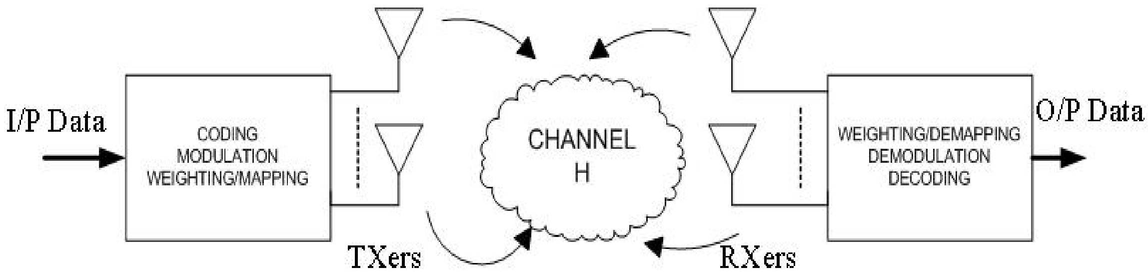

Figure 1.1: Diagram of MIMO wireless system. Coding, modulation, and mapping of the signals onto antennas may be realized jointly or separately.

tional cost. Alternatively, turbo principle can be explored to achieve near-optimal performance at a much lower computational complexity, and is incorporated in our study. In the next few sections we will give an overview of MIMO communication, blind detection and the turbo principle.

1.2 Multiple-Input Multiple-Output Systems

In simplest terms a MIMO system can be defined as a wireless link with the transmitting as well as the receiving end equipped with multiple antennas (see Figure 1.1), along with some signal-processing at the transmit antenna and the receive antenna arrays that results in increase in quality (decrease in bit-error rate (BER)) or the data rate (in bits/sec) of transmission over this link. The most attractive feature of a MIMO system is its ability to take advantage of random fading (due to multipath propagation) (Foschini & Gans mar. 1998), (Telatar june 1995) and when available, multipath delay spread (Raleigh & Cioffi mar. 1998) for increasing the data rates at no extra spectrum cost (while hardware cost and complexity do increase). The increase in data rate offered by a MIMO system is owing to its ability to transmit multiple independent information streams in parallel (known as the spatial multiplexing gain) over a matrix channel created by multiple transmit and receive antennas. Whereas, the increase in quality (or BER minimization) is due to the ability of a MIMO system to utilize the spatial dimension to combat adverse channel conditions (known as the diversity gain).

In order to explain the spatial multiplexing gain in MIMO, consider a wireless system with N_t transmit antennas and N_r receive antennas. At the transmitter N_t data streams are transmitted that mix together in the wireless channel as they use the same frequency spectrum. At the receiver, the objective is to estimate the mixing channel matrix (through training symbols) and to separate individual data streams. Assuming flat-fading channels, that is, each entry of the channel matrix is a scalar coefficient, the separation of data streams is possible only if each receive antenna sees a sufficiently different channel. A highly scattering environment resulting in rich multipath ensures that this condition is satisfied. The key point here is that unlike conventional single-input single-output (SISO) wireless systems where multipath represents an impediment to accurate transmission, MIMO actually exploits multipath to maximize the data rate over a given transmission link. Spatial multiplexing in MIMO is somewhat similar to code-division multiple access (CDMA) transmission in which multiple users/streams share the same time/frequency channel upon transmission and are recovered through their unique codes (signatures). The main difference is that in MIMO, unique spatial signatures of input streams exist naturally due to rich multipath, thus using available spectrum more efficiently.

Traditionally, multiple antennas have been used at the receiver to increase the diversity gain. The idea behind this is to send the signals carrying the same information through different paths, so as to obtain multiple independently faded replicas of the same data, which ensures more reliable reception. More recently (Tarokh, Seshadri & Calderbank mar. 1998), (Alamouti oct. 1998) it has been shown that the diversity gain can be achieved using multiple antennas at the transmitter by joint coding (space-time trellis codes(STTC) or space-time block codes (STBC)) of transmitted data streams. Since, a MIMO system has multiple antennas at both ends, therefore, it can have a maximal diversity gain of N_tN_r. However, the diversity gain of

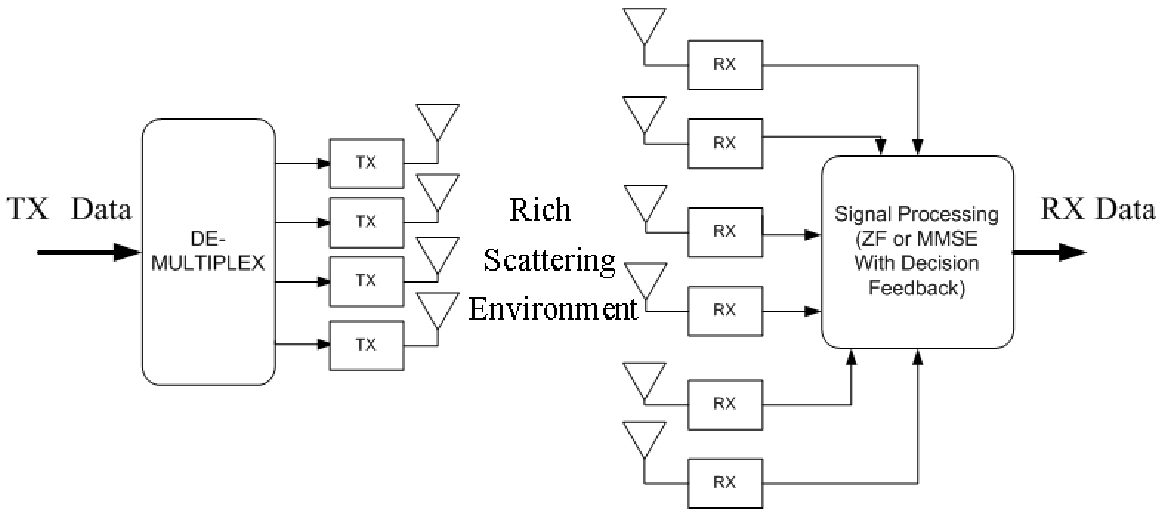

Figure 1.2: V-BLAST System Model

an actual MIMO system depends on the the system architecture and communication environment.

The capacity of a MIMO system with N_t transmit and N_r receive antennas with no transmit channel state information (Foschini & Gans mar. 1998), (Telatar june 1995) is given by

$$C = log_2[det(\mathbf{I}_{N_r} + \frac{\rho}{N_t}\mathbf{H}\mathbf{H}^H)] \quad b/s/Hz, \tag{1.1}$$

where H denotes the transpose-conjugate, $\mathbf{H}$ is the channel matrix, ρ is the SNR at any receive antenna and we assume N_t uncorrelated equal-power sources. Foschini and Telatar both demonstrated that the capacity in (1.1) grows linearly with $n = min(N_t, N_r)$ under certain channel conditions. This information theoretic result serves as an upper bound for actual algortihms/architectures for MIMO developed to achieve a given BER and data rate at a reasonable complexity.

Most of the existing architectures for MIMO focus on either the spatial multiplexing gain or the diversity gain. In (Zheng & Tse may 2003) it is shown that it is possible to achieve a tradeoff between the multiplexing gain and the diversity gain. Bell-labs space-time architectures (BLAST) are the most widely known MIMO architectures that focus on maximizing the spatial multiplexing gain. Diagonal-BLAST

(Foschini 1996) utilizes a diagonally layered coding structure at the transmitter in which code blocks are dispersed across diagonals in space-time. In an independent Rayleigh environment, this processing structure can achieve rates approaching 90% of the Shannon capacity. Vertical-BLAST (Wolniansky, Foschini, Golden & Valenzuela oct. 1998) is the most likely architecture (Figure 1.2) to be utilized in future wireless systems due to its low complexity and ease of implementation. In V-BLAST (without forward error correction), the encoding process for transmitted data is simply a demultiplex operation followed by independent bit-to-symbol mapping of each substream. Whereas, at the receiver conventional adaptive antenna array techniques (e.g. minimum mean square error (MMSE), zero forcing (ZF)) are used, to detect one substream at a time and null remaining substreams, in conjugation with successive cancellation (cancelling out detected components of the transmit vector from the received signal vector) in a manner analogous to the decision feedback equalization. A recent addition to this class of architectures is turbo-BLAST (Section 1.4), which can achieve rates very close to shannon capacity by better exploitation of transmit and receive diversity.

1.3 Blind Detection

1.3.1 Why blind?

Blind techniques have been actively studied over the past 25 years. Many algorithms have been proposed in the context of blind equalization and signal separation. The goal of blind estimation is to determine the channel or the signals based on the prior temporal or spatial knowledge. Although, the use of training sequences is probably the most robust way to estimate the channel, there are several reasons for studying blind algorithms:

- Bandwidth is conserved by eliminating or reducing training set.
- In rapidly time-varying channels, training is not efficient.

- Severe multipath fading during the training period can lead to poor channel estimates.

- Training for interference is often not accessible.

- Training requires synchronization, which may not be feasible in multi-user scenarios.

- In communication intelligence, training is not available.

- In distributed networks, it may not be feasible to send training sequences each time when a new communication link is set up.

Blind algorithms also suffer from some drawbacks as compared to non-blind techniques. In general, blind algorithms tend to be computationally more expensive. Due to the non-linear nature of the blind estimation, many proposed methods converge to a local rather than a global minimum. Also, uniqueness of signal estimates may not be guaranteed if small data lengths are used. All blind algorithms result in an inherent phase and ordering ambiguity in recovered signals. Apart from the first one, all other problems can be resolved by using a short list of training signals. Although the algorithms are no longer blind, they retain many of the advantages associated with blind algorithms. Hence, purely blind and non-blind methods correspond to the two extremes of an array of possible algorithms. In practice, ideas from both approaches can be combined to minimize the training signal requirement of non-blind methods, and yet obtain the robustness of blind methods at low computational costs.

1.3.2 Blind deconvolution/Blind equalization

Deconvolution is a signal-processing operation that ideally unravels the effects of convolution performed by a linear time-invariant system on an input signal. In ordinary deconvolution operation, both the output signal and the system are known, and the requirement is to reconstruct the input signal. In blind deconvolution (BD), only the output signal is known, and the requirement is to reconstruct the input signal and the system itself. This makes BD a more difficult signal-processing task than

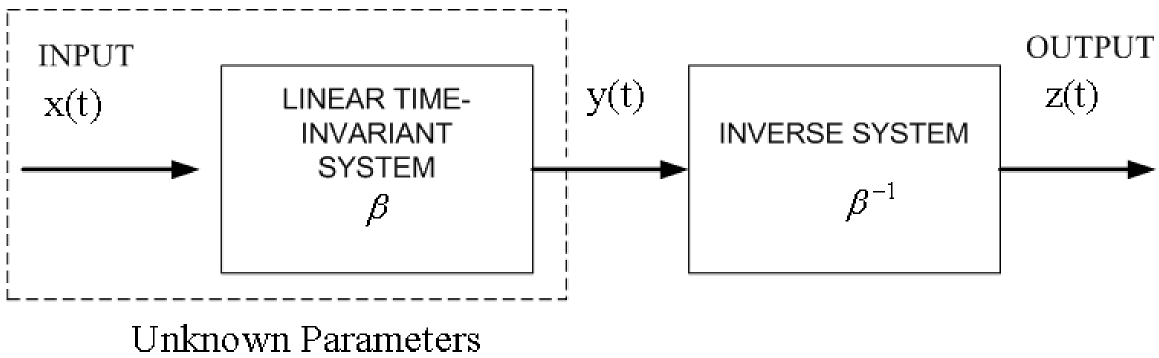

Figure 1.3: Block Diagram for Blind Deconvolution

ordinary deconvolution. Consider an unknown linear time-invariant (LTI) system β with input x(t) assumed to consist of independent and identically distributed (i.i.d.) symbols. The only thing known about the input is its probability distribution. The requirement is to restore x(t), or equivalently to identify the inverse β^{-1} of the system β, as depicted in Figure 1.3.

If the given system β is minimum-phase (i.e., the transfer function of the system has all its poles and zeros inside the unit circle in the z-plane), then not only is the system stable but so is its inverse. Also, magnitude and phase response of a minimum-phase systems are directly related to each other, i.e., if one is known the other can be easily computed. In this case the inverse system β^{-1} that recovers the input x(t) from the unknown system output y(t) is just a whitening filter. In general a whitening filter ensures the temporal decorrelation (independence up to second order moments) of its output, which is good enough to recover the magnitude response of input but the phase information about input signals (present in higher order moments) is lost. However, if the unknown system and its inverse are minimum-phase, the phase response can be recovered from the magnitude response and a whitening filter is sufficient for BD. With this observation, the BD problem is solved. But most practical systems (e.g. a telephone channel, a multipath fading channel, etc.) that require deconvolution are non minimum-phase. In such cases blind deconvolution (BD) is a non-trivial signal-processing task and thus requires use of more complex algorithms. In communication literature deconvolution of the channel output is more commonly known as equalization, therefore in all further discussions the term blind equalization (BE) is used in place of blind deconvolution (BD).

1.3.3 Classification of blind equalization techniques

Blind channel equalization has received considerable attention in communication and signal-processing literature. Higher order statistics (HOS) based BE, second order statistics (SOS) based BE and subspace based BE are some of the widely used (among numerous other) approaches listed in literature for blind identification of the inter-symbol interference (ISI) channel and/or the transmitted signal.

Most of the early algorithms for BE exploited the higher order statistics of the system output for signal recovery. These algorithms use the higher order statistics of the received signal in an implicit or explicit (described in terms of cumulants or polyspectra) sense to recover the transmitted signal. These algorithms provide a basis for the identification (and therefore blind equalization) of a non-minimum phase system by virtue of their ability to preserve the phase information in the observed signal. In addition these algorithms are robust in the presence of noise and converge to a global optimum. The main drawback in using the higher order statistics based methods is that they require a large amount of data which may be unacceptable in time-varying channels. One such blind method using HOS in an implicit sense (Godard nov. 1980), (Triechler & Agee apr. 1983) is the constant modulus blind equalization. This algorithm involves designing an equalizer by stochastic gradient minimization of a cost fuction penalizing the dispersion of the squared equalizer output ($|z_n|^2$) away from a constant ($E|x_n|^4/\sigma_x^2$), which is a function of the source alphabet. This equalizer accomplishes perfect blind equalization under certain conditions on the channel matrix and the source properties. Note that non-linear operation (squaring) on the equalizer output results in an implicit use of HOS of the system output. Another algorithm that uses HOS in explicit sense (Shalvi & Weinstein mar.

1990) is based on multiuser kurtosis maximization criterion. Here the blind equalizer is computed by stochastic gradient maximization of the kurtosis of the equalizer output ($K(z_n) = E(|z_n|^4) - 2E^2(|z_n|^2) - |E(z_n^2)|^2$) under the constraint that product of the channel vector (H) and the equalizer vector (W), $G = H^T W$ is unitary. Apart from these there are a number of other HOS based blind equalization methods (Sato june 1975), (Hatzinakos & Nikias jan. 1989), (Tugnait may 1987) etc.

In (Gardner june 1991), it is shown that SOS of cyclostationary signals contain phase information which can be used for the identification of non-minimum phase systems. This led to the development of algorithms which explore the cyclostationary properties of oversampled communication signals to allow the blind estimation to be accomplished based on the SOS of the channel output. The use of SOS translates into lesser number of observations (faster convergence) as compared to HOS based algorithms. The first SOS based algorithm developed (Tong, Xu & Kailath nov. 1991) employs temporal oversampling on the channel output (y_n), while assuming an i.i.d. system input (x_n). It then enforces a Jordan structure on the input correlation matrix ($R_{x_n}(k) = E(x_n x_{n-k}^H)$) and essentially reconstructs the channel vectors in channel matrix (H) based on the information provided by the change of rank of $R_{y_n}(k) = E(y_n y_{n-k}^H)$. There are a number of other SOS based blind equalization approaches listed in the literature (Tong, Xu, Hassibi & Kailath jan. 1995), (Giannakis nov. 1994). The blind identification using SOS has a unique solution provided certain conditions (Tong et al. nov. 1991), (Xu, Liu, Tong & Kailath april 1994) on channel and input are met. Although these SOS based methods perform well in general, they suffer a performance degradation caused by a model mismatch when a limited number of observations are available.

Another class of algorithms that are more data efficient is subspace based blind identification algorithms. Oversampling of the channel output results in a block Toeplitz channel matrix structure and a Hankel input matrix structure. These algorithms exploit certain inherent subspace structure arising from the combination of the block-Toeplitz channel matrix and the Hankel input matrix in the absence of noise, to provide exact channel information using only a finite number of observations. The channel subspace algorithm (Moulines, Dohamel, Cardoso & Mayrargue feb. 1995),

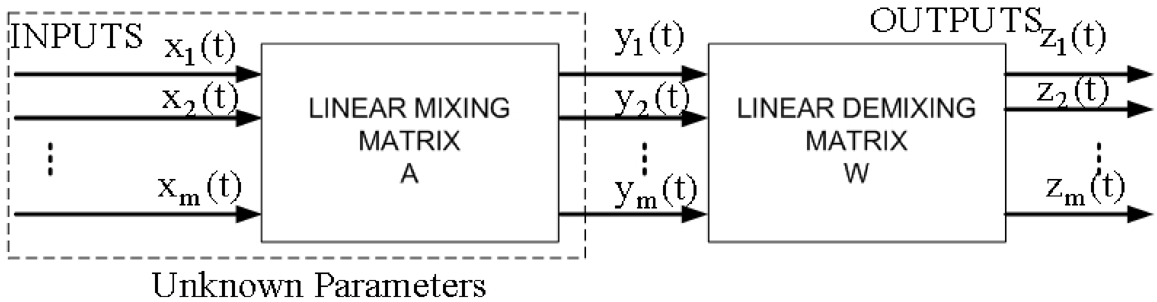

Figure 1.4: Block Diagram for Blind Source Separation

(Slock nov. 1994) depicts the key role of channel structure in blind estimation, it is built on the fact that the column span of the block Toeplitz channel matrix and its corresponding complement are the channel signal subspace and channel signal orthogonal subspace respectively, which can be caclulated from the channel data matrix. Some other algorithms that exploit channel and signal subspace structures are the least square algorithm (Liu, Xu & Tong nov. 1993) and the signal subspace algorithm (Liu & Xu nov. 1995). The performance of this genre of algorithms is fundamentally limited by the nature of the channel. For example, singularity in the channel matrix can result in the divergence of subspace, and result in failure of subspace approaches.

1.3.4 Blind source separation (BSS)

In order to describe the basic blind signal separation problem, consider a set of unknown source signals $x_1(t), x_2(t), \ldots, x_m(t)$ that are mutually independent of each other. These signals are linearly mixed in an unknown environment to produce a $m \times 1$ observation vector (see Figure 1.4)

$$Y(t) = \mathbf{A}X(t), \tag{1.2}$$

where X(t)= $[x_1(t), x_2(t), \ldots, x_m(t)]^T$, Y(t) = $[y_1(t), y_2(t), \ldots, y_m(t)]^T$ and $\mathbf{A}$ is an unknown non-singular mixing matrix of dimensions $m \times m$. Given the observation vector Y(t), the requirement is to recover the original source signals $x_1(t), x_2(t), \ldots x_m(t)$ without any knowledge about the mixing matrix $\mathbf{A}$. The solution to this problem is feasible, except for an arbitrary scaling of each source signal and possible permutation of indices under certain but fairly general conditions. Given that the original source signals are independent and the mixing matrix $\mathbf{A}$ is non-singular, it is possible to find a demixing matrix $\mathbf{W}$ defined as follows:

$$Z = \mathbf{W}Y = \mathbf{WA}X = \mathbf{DP}X, \tag{1.3}$$

where Z is the output signal vector produced by the demixer, $\mathbf{D}$ is the non-singular diagonal matrix, and $\mathbf{P}$ is the permutation matrix. The underlying principle involved in the solution of this problem is called independent component analysis (ICA), which may be viewed as an extension of widely known principal component analysis (PCA). Whereas PCA imposes independence in a statistical sense only to second order while constraining the direction of the vectors to be orthogonal, ICA imposes statistical independence on all the individual components of the output vector, but has no orthogonality constraint. In general it is possible to blindly recover source signals for a $n \times m$ mixing matrix, where the $n \geq m$ case can be handled using ICA, however the $n < m$ case requires additional information about sources along with independence of sources.

1.3.5 Classification of BSS techniques

A number of blind source separation algorithms have been proposed in recent past. In this section, we classify some of these algorithms and highlight their common properties. As a first step, BSS algorithms can be broadly classified into two classes based on properties used for channel or signal estimation, which include temporal signal properties such as higher order statistics, finite alphabet, cyclostationarity, and spatial receiver properties such as calibrated array or special array geometry.

We begin with a discussion on temporal signal properties based algorithms which enforce a known signal property on their estimates. Early fundamental work on HOS based BSS can be traced back to (Donoho 1981), where the problem was treated from an output entropy minimization point of view, and was later expanded in (Comon july 1996) using notion of contrast (a contrast is a cumulant based function of the outputs that is maximized if and only if separation is achieved). Some other methods that rely directly on HOS cumulants were introduced in (Cardoso 1989), (Shamsunder & Giannakis 1994). The HOS based BSS criterion described in (Papadias dec. 2000*b*) (detailed description given in later chapters) builds upon the kurtosis maximization technique for BE and deflation approach (sources are extracted one by one)(Delfosse & Loubaton apr. 1994*a*), to give a BSS algorithm that is globally convergent. Another class of algorithms that use HOS implicitly are constant modulus (CM) type BSS techniques that build upon the CM techniques for BE, notable among these (Vanderveen & Paulraj 1996). Finite alphabet property based BSS algorithms, e.g., (Talwar, Viberg & Paulraj feb. 1994), also use HOS in an implicit sense, which exploit the finite alphabet property of the digital signals to develop a maximum likelihood approach for estimating transmitted signals from the received instantaneous mixture. Some other algorithms rely on assumption that different sources have different cyclostationary features (Agee, Schell & Gardner apr. 1990) or exploit spectral redundancy at the cyclic frequency for source separation (Schell & Gardner 1993).

In the spatial techniques the desired signal is estimated in two-steps. First, the directions of arrivals (DOAs) of the impinging signals are determined using subspace information and prior knowledge of the array geometry. Common high resolution techniques for direction finding include MUSIC (Schmidt 1981) and ESPIRIT (Paulraj, Roy & Kailath 1985). The array responses corresponding to the estimated directions are then used to construct a beamforming weight vector. The weight vector optimally combines the array outputs to extract each desired signal from interference and noise (Otterson, Roy & Kailath 1989). Some widely used optimization criteria for beamforming are minimization of mean squared error (MMSE), maximization of signal to interference plus noise ratio (SINR) and maximum likelihood(ML). The performance of this approach depends strongly on the reliability of prior spatial information, which may not be known precisely in a highly variable propagation environment. In addition, DOA based algorithms require that the number of wavefronts, including multipath reflections and interference, be less than the number of sensors.

1.3.6 Similarities and differences between BSS and BE

Blind equalization and blind source separation are the two most important blind signal-processing problems. These problems are somewhat related (evident from the use of similar concepts for developing algorithms for solving them) with similarities and subtle differences between them. In BSS, sources are corrupted by the superposition of other sources, in BE, a source is corrupted by time-delayed versions of itself. In both cases unsupervised (blind) learning must be used because no error signals are available. The most important dissimilarities regarding the nature of BSS and BE are

- The dimension of the source-separation task is equal to or less than the length of the vector Y(k) that is composed of linear mixture of the unknown source signal vector X(k), whereas the dimension of the deconvolution task depends on the length of the impulse response of the unknown LTI system β.
- Source separation can involve m (where m is the number of unknown sources) different source probability distribution functions (pdfs), whereas deconvolution only involves one source pdf. The difference between pdfs in source separation means that solutions to the problem may require estimation of different source pdfs, or detection of difference between source pdfs.

1.3.7 Blind MIMO detection

In the case of multiple-input multiple-output (MIMO) systems, if the channel is memoryless then the blind detection problem essentially reduces to a blind source separation problem, where channel matrix ($\mathbf{H}_{N_r \times N_t}$) is the unknown mixing matrix. We have already discussed some of the detection algorithms for this case in preceding sections. If the channel has memory (i.e., it introduces a delay spread) then the blind detection problem involves BE along with BSS. The later case (also known as convolutive mixture blind identification) presents a much tougher signal-processing problem. A large number of blind algorithms for MIMO channels with memory are either built upon existing BSS algorithms by incorporating additional features or use some combination of BE and BSS algorithms for extracting source signals out of a convolutive mixture at the receive antenna array. In this thesis, the emphasis is on blind MIMO detection for memoryless channel, but for completeness, we will briefly describe some of the promising detection techniques proposed in recent works for solving the convolutive mixture blind identification problem (FIR-MIMO blind identification).

In (Vanderveen, Talwar & Paulraj may 1995), the class of algorithms for BE, initiated in (Tong et al. nov. 1991), where the signal is recovered by oversampling the channel output are combined with finite alphabet based algorithms for BSS proposed in (Talwar et al. feb. 1994), to derive an algorithm that can identify multiple FIR channels carrying a superposition of digital finite alphabet signals. In the first step of this algorithm ISI caused by the channel is removed by using the BE algorithm and in the next step the instantaneous mixture of the input signals so obtained is separated using the BSS algorithm. In (Papadias feb. 2004), the multiuser kutosis maximization algorithm for BSS (Papadias dec. 2000*b*) is extended to the FIR-MIMO case by restructuring the data model such that the FIR-MIMO source separation problem is essentially reduced to BSS problem with increased dimensions. Analogous to full rank condition on channel matrix in case of BSS, FIR-MIMO case also requires a set of conditions on channel matrix to be fulfilled for blind channel identifiability. Another interesting approach (Liu & H. Luo may 2002) based on autocorrelation matching was proposed recently. This is an SOS based technique that recovers the transmitted signals up to a unitary factor and a finite delay, if and only if the autocorrelation function of the output signal equals that of the transmitted signal, given that the autocorrelation function of transmitted signals and MIMO-FIR channel satisfy autocorrelation matching and FIR invertibility condition respectively. An extension of the classical constant modulus (CM) algorithm for BE is multiuser constant modulus (MU-CMA) algorithm derived in (Papadias & Paulraj june 1997). This algorithm optimizes a cost function which has a weighted sum of CM term and a cross correlation term corresponding to equalized signals. The CM term penalizes the deviations of the equalized signals' magnitudes from a constant modulus, whereas the cross-correlation term penalizes the correlations between them, thus achieving blind separation of a linear dispersive mixing channel.

In summary, we have seen that there is a wide range of possible techniques for blind MIMO detection, which can be well suited for specific wireless communication applications. The choice of particular technique depends on signal, channel and receiver characterstics. Convergence, computational complexity and ability to track variations in a wireless channel are other important issues taken into account when making this choice.

1.4 Turbo Detection

1.4.1 Turbo principle

Error control coding is ubiquitous in wireless and other impaired channels as it assists in approaching the capacity limit for a given channel predicted by the channel coding theorem. Among various existing error control coding schemes, Turbo coding invented by Berrou and Glavieux (Berrou & Glavieux oct. 1996) stands out due to its ability to attain the channel capacity by few hundredths of a decibel. The basis for this novel coding scheme is that parallel or serial concatenated codes separated by an interleaver exhibit near Shannon-limit performance if they are decoded jointly in an optimal fashion. But this type of optimal decoding is practically impossible due to high computational complexity. Turbo coding scheme mitigates this problem of computational complexity by decoding each of the parallel or serially concatenated codes separately and iteratively exchanging information about a set of transmitted symbols between decoders separated by interleavers/de-interleavers. The key idea behind this type of coding scheme, " when several probabilistic machines work together on the estimation of a common set of symbols, all machines have to give the same decision, with the same probability, about each symbol, as a single (global) decoder would," (Berrou aug. 2003) is known as the Turbo principle. In any feedback configuration there exists a possibility of unstability due to positive feedback. In turbo decoding this situation is avoided by using only the extrinsic information (additional information generated about a symbol by given decoder) for feedback. The soft information about a set of symbols exchanged by decoders is in terms of log of ratios of the probability of a symbol having a particular logical value (Pr{d=1}) to the probability of that symbol not having this logical value (1-Pr{d=1}). This soft information is termed as log likelihood ratio (LLR).

In recent years it has been shown that the turbo principle discussed here in the context of error control coding can be successfully applied to several other complex signal processing problems in digital communications. Notable among these are joint channel equalization and decoding (C. Douillard oct. 1995), joint multiuser detection and decoding (Poor sept. 2001), joint channel estimation and decoding (Komninikas & Wesel sept.2001). The common feature of all these is the use of the turbo structure which involves iterative exchange of soft information about symbols between the two blocks separated by interleavers.

1.4.2 Turbo multiuser detection

The type of communication in which several transmitters share a common channel is known as multiaccess communication. Mobile telephones transmitting to a base station, ground stations communicating with a satellite, and local area networks are a few examples of this type of communication system. A common feature of these communication channels is that the receiver obtains a noisy version of the superposition of the signals sent by the active transmitters. A simple solution to avoid this superposition of signals is to ensure an orthogonal multiplexing of signals (various users operate in separate non-interfering channels). But this solution is practically infeasible due to economic considerations and the fact that unintentional superposition of signals is still possible. Thus, in all practical multiacess communication systems users are multiplexed in a non-orthogonal fashion.

Multiuser detection refers to the detection of data from multiple users when observed in a non-orthogonal multiplex. This problem arises naturally, for example, in code-division multiple-access(CDMA) systems using non-orthogonal spreading codes. It also arises in orthogonally multiplexed wireless channels, such as TDMA channels, due to effects such as multipath or non-ideal frequency channelization, and in wire-

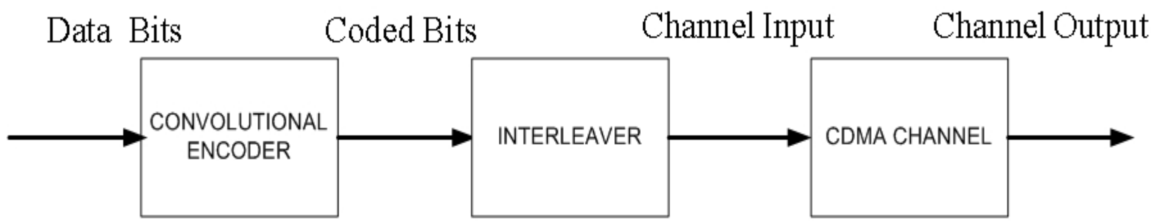

Figure 1.5: Transmitter Configuration for Turbo-MUD

line channels such as digital subscriber lines (DSLs) in which crosstalk is the major impairment. The basic idea of multiuser detection is to exploit the cross-correlations among the signals to be demodulated in order to improve the data detection.

As discussed previously, error coding is an essential component of any modern wireless communication system. A possible configuration is a convolutional encoder mapping data symbols into channel symbols, followed by an interleaver, and finally a CDMA modulator for channel symbols, as shown in Figure 1.5. One can view the configuration of Figure 1.5 as a serially concatenated code, in which CDMA spreading and the wireless channel together form the inner code, and the convolutional code is the outer code. A traditional way of demodulating this concatenation is to first demodulate the CDMA signals, followed by a de-interleaver and a channel decoder. To seek optimality in such situation, one could replace this traditional configuration with an overall demodulator/decoder that uses an optimal (say maximum likelihood) mapping from the received signal to the original data symbols. The complexity of such a system is prohibitively high. This complexity can be mitigated however, by exploiting the turbo principle. In particular we can reduce the complexity of joint decoding and multiuser detection by an iterative exchange of soft information between the multiuser detection and channel decoding, iterating until some kind of convergence is reached (see Figure 1.6). Like Turbo decoding this iterative approach to joint

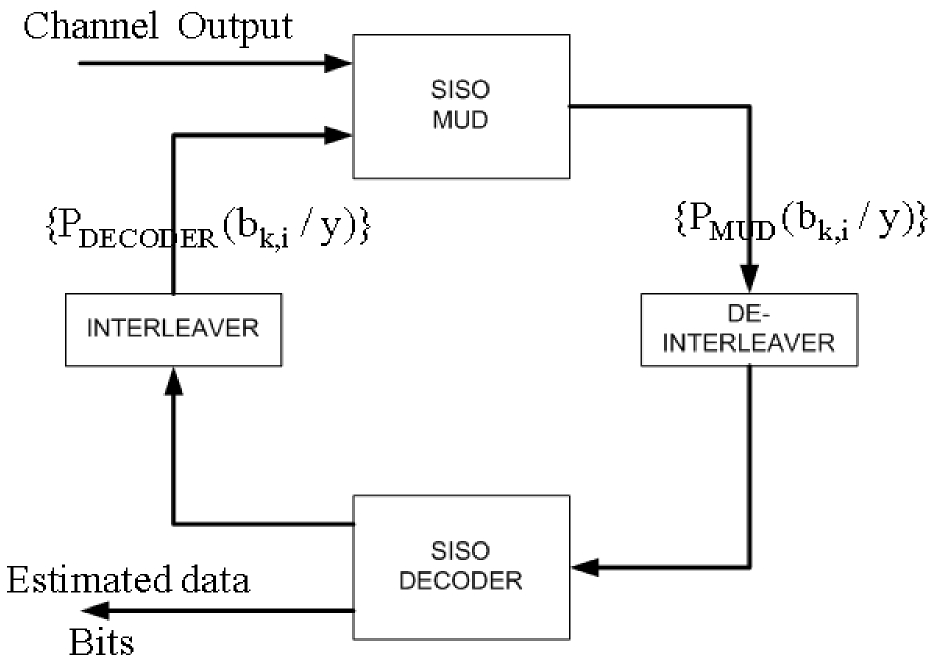

Figure 1.6: General Receiver Structure for Turbo-MUD

multiuser detection and decoding is shown to have very good performance. In (Wang & Poor july 1999) such turbo multiuser detectors with maximum likelihood detection (ML) stage and minimum mean square error with parallel interference cancellation (MMSE-PIC) stage have been proposed and shown to have performance close to the single user bound. Here we have discussed the turbo multiuser detection principle in the context of CDMA cellular systems but the same principle is applicable to other cellular systems and applications as well.

We discussed the similarities between CDMA and MIMO system architectures in previous sections. Due to these similarities, we can view a coded MIMO system as a serially concatenated code as we did in the case of CDMA system. That is, here too we can simplify the problem of joint space-time MUD and decoding greatly by applying the turbo principle. In some recent works (Dai, Molisch & Poor mar. 2004) and (Sellathurai & Haykin oct. 2002), such turbo space-time detection structures built upon Bell labs layered space time architecture (BLAST) have been proposed.

These structures have been shown to reach very close to the capacity limit in the interference-free (single cell) case. The detection structures, such as coded V-BLAST and T-BLAST in (Dai et al. mar. 2004), are composed of a BLAST-like MUD stage with soft metric output and a maximum aposteriori probability (MAP) decoding stage, exchanging soft information iteratively (refer to Chapter 3 for more details about T-BLAST). The detection structure in (Sellathurai & Haykin oct. 2002) is more like D-BLAST with a sub-optimal turbo receiver that performs iterative decoding of the random layered space time codes (applied at the transmitter) and estimation of channel matrix in an iterative and simple fashion.

Following the same line of reasoning the turbo principle is applicable for joint MUD and decoding in coded blind MIMO systems, too. The only implementation of such structure we came across in literature is Sequential Monte Carlo (SMC) based turbo receiver for MIMO systems proposed in (Guo & Wang april 2003). This receiver structure uses SMC methodology for blind detection and a MAP decoder in an iterative manner in order to approach close to the capacity limit. The SMC methodology recently emerged in the fields of engineering and statistics and can be used in solving a wide class of sophisticated statistical inference problems. Under a state-space framework the SMC recursively generate Monte Carlo samples of the state variable or some other latent variables, based on which the posterior distribution of any parameter of interest can be approximated. The SMC methods exhibit global convergence and are robust to choice of initial conditions.

1.5 Thesis Overview and Organization

In this thesis we focus on developing new semi-blind turbo detection structures for single-cell MIMO systems and multi-cell MIMO systems. In the case of single-cell MIMO sytems our goal is to estimate the transmitted symbol substreams in the presence of background noise and intra-cell interference. For multi-cell MIMO systems our goal is to estimate the transmitted symbol substreams (from desired user), which are corrupted by background noise, intra-cell interference and inter-cell interference (due to symbol substreams from users in the adjacent cells). We aim to estimate the transmitted substreams in single-cell MIMO systems and the interfering substreams (from users in the adjacent cells) in multi-cell MIMO systems in a semi-blind fashion (using minimal number of training symbols), so as to reduce the system overhead and to make these detection structures a better choice for use in broader wireless applications. In multi-cell MIMO systems, the estimation and removal of interfering substreams assists in achieving the main objective, that is to get an improved estimate for symbol substreams transmitted by the desired user.

The thesis is organized as follows. In this first chapter, we have motivated our work in the context of multiple-input multiple-output systems, existing BE/BSS/blind MIMO detection techniques and the concept of turbo MUD. In chapter 2, we will discuss the multiuser kurtosis maximization criterion, a promising idea recently proposed for blind MIMO detection, and the stochastic gradient algorithm that can be used to implement this criterion. After this we develop a new semi-blind turbo space-time MUD structure for single-cell MIMO, which we name as turbo - multiuser kurtosis maximization (T-MUK). The detection stage of this turbo MUD structure is a modification of MUK detector such that it can process soft information from the decoder stage with interference cancellation and output soft information to be fed to the decoder. We show that T-MUK structure works even if symbol streams transmitted from different antennas have different power levels. We demonstrate the performance of T-MUK with some numerical results.

In Chapter 3 we propose a new semi-blind group interference cancellation structure for multi-cell MIMO systems. This semi-blind receiver structure uses T-MUK in conjugation with T-BLAST in a group multistage iterative fashion to estimate desired user signals in the presence of co-channel interference. We discuss the possible multi-cell scenarios where this semi-blind structure could be used. Again, we demonstrate the performance and advantage of this structure through computer simulations. Finally, we conclude in Chapter 4 with a summary of the thesis, and directions for future research. The main ideas in this thesis will appear in (Goel & Dai 2005).

Chapter 2

Turbo Multiuser Kurtosis Maximization

In this chapter we review the multiuser kurtosis (MUK) maximization criterion and the corresponding MUK algorithm for blind MIMO detection (Papadias dec. 2000*b*) and describe some of the important issues related with them in detail (Section 2.1). We then proceed to propose a new semi-blind receiver structure for single-user MIMO detection for transmission over frequency flat channels (Section 2.2). This receiver enhances the MUK based receiver by incorporating a decoder stage and performing joint detection and decoding by appealing to the turbo principle, using very small training sequence. The reduction in overhead cost offered by this receiver is its main advantage over conventional receivers for MIMO detection, that typically dictate perfect channel knowledge, which translates to longer training sequences. We then show that this novel receiver structure can also be used for multi-user MIMO detection (Section 2.3) and finally we present some numerical results (Section 2.4) to highlight the usefulness of this receiver structure.

2.1 Multiuser Kurtosis Maximization (MUK) Criterion

2.1.1 MIMO system model

We first consider a single cell MIMO system (no inter-cell interference) and adopt the well known MIMO system model given by

$$Y(k) = \mathbf{H}X(k) + N(k), \tag{2.1}$$

where $Y(k)$ represents the $N_r \times 1$ received signal vector, $X(k) = [x_1(k) \ldots x_{N_t}(k)]^T$ is the $N_t \times 1$ transmitted signal vector, $\mathbf{H}$ represents the $N_r \times N_t$ matrix which captures channel characteristics between the transmit and the receive antenna arrays, $N(k)$ is the additive background noise vector, all at a time instant k. The transmitted signal power is constrained by $E[X(k)^H X(k)] \leq P$, and the additive background noise is circularly symmetric Gaussian with the covariance matrix given by $\mathbf{\Phi}_N = \sigma^2\mathbf{I}$. The entries of the complex matrix $\mathbf{H}$ are independent with uniformly distributed phase and normalized Rayleigh distributed magnitude, modelling a Rayleigh fading channel with sufficient separation between the transmit and the receive antennas. The signal to noise ratio (SNR) is given by $\rho = P/\sigma^2$.

For studying blind methods in MIMO systems, we further assume that N_t streams transmitted from N_t antennas are i.i.d., mutually independent, zero-mean discrete time sequences. The objective of the MUK criterion to be discussed in subsequent sections is to recover these symbol streams from the output of the linear channel ($\mathbf{H}$) that introduces inter-stream interference. In order to achieve this objective, the received signal vector $Y(k)$ is filtered by an $N_r \times N_t$ matrix equalizer $\mathbf{W}$ that produces output vector $Z(k) = [z_1(k) \ldots z_{N_r}(k)]^T$,

$$\begin{aligned} Z(k) &= \mathbf{W}^T Y(k) \\ &= \mathbf{W}^T \mathbf{H} X(k) + N_w(k) \\ &= \mathbf{G}^T X(k) + N_w(k), \end{aligned} \tag{2.2}$$

where $\mathbf{G}$ represents the $N_t \times N_t$ global response matrix (product of the channel matrix and the equalizer matrix), and $N_w(k)$ is the colored noise at the receiver output. In case of perfect signal recovery (in absence of noise) the equalizer ouput vector $Z(k)$ should be an exact replica of the transmitted signal vector $X(k)$, whose components are given by

$$z_j(k) = \sum_{l=1}^{N_t} g_{jl} x_l(k) = G_j^T X(k), \qquad j = 1, ..., N_t, \tag{2.3}$$

where G_j (the j^{th} column of $\mathbf{G}$) is composed of weighting factors for all the channel inputs to the j^{th} equalizer ouput. We can express the same relationship in a more compact vector notation

$$Z(k) = \mathbf{G^T} X(k), \tag{2.4}$$

where

$$\mathbf{G} = [G_1 \dots G_{N_t}] \begin{bmatrix} g_{11} & \cdots & g_{N_t 1} \\ \vdots & \vdots & \vdots \\ g_{1N_t} & \cdots & g_{N_t N_t} \end{bmatrix}. \tag{2.5}$$

The above described instantaneous linear mixture setup is shown in Figure 2.1. We define the variance and the kurtosis of equalizer output (needed later to develop the MUK criterion) as follows

$$E(|z_j|^2) = \sigma_x^2 \sum_{l=1}^{N_t} |g_{jl}|^2, \qquad j = 1, \dots, N_t \tag{2.6}$$

$$K(z_j) = K_x \sum_{l=1}^{N_t} |g_{jl}|^4, \qquad j = 1, \dots, N_t \tag{2.7}$$

where $K_x = E(|x|^4) - 2E^2(|x|^2) - |E(x^2)|^2$ is the unnormalized kurtosis and σ_x^2 is the variance of any input substream $\{x_l(k)\}$.

2.1.2 Necessary and sufficient conditions for BSS

It is well known that statistical blind techniques are inherently incapable of distinguishing between the phase rotated versions of an input signal. Also, statistical blind techniques cannot differentiate between the different input signals if they are

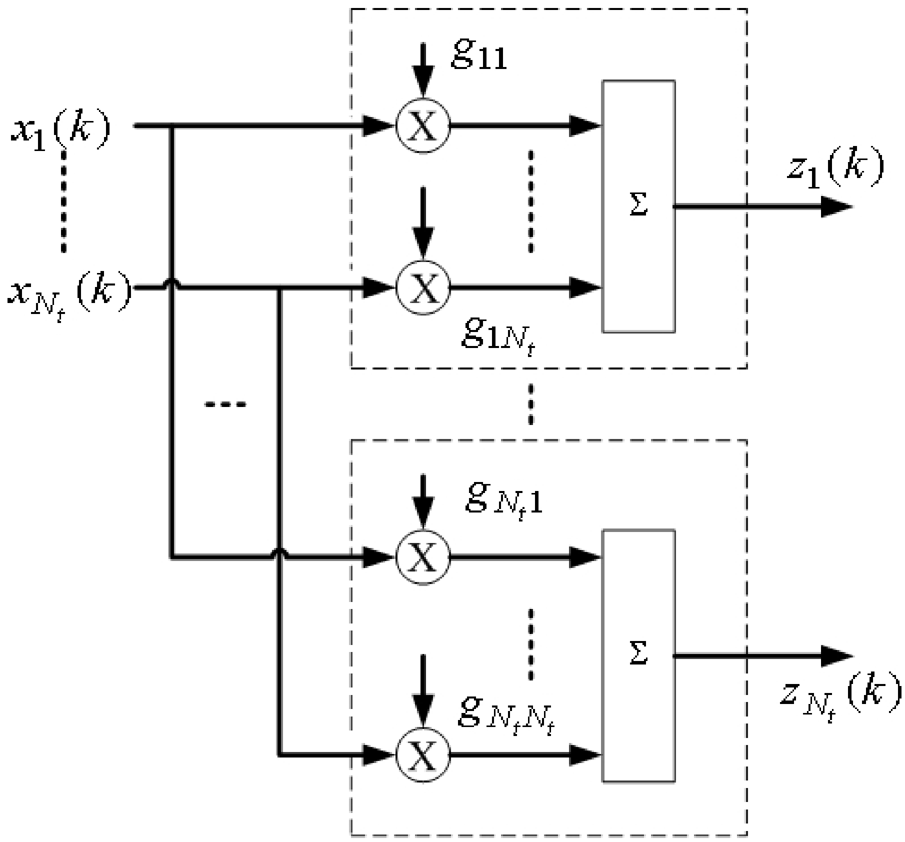

Figure 2.1: Instantaneous Linear Mixture Setup

identically distributed (Cardoso oct. 1998). Thus, we say blind recovery is achieved if, after suitable reordering of the equalizer outputs the following holds

$$z_j(k) = e^{i\phi_j} x_j(k), \tag{2.8}$$

for some $\phi_j \in [0, 2\pi)$ and all $j \in 1, \ldots, N_t$.

If the conditions on the transmitted signals described in the previous section hold, the following set of conditions are necessary and sufficient for the blind recovery of all the transmitted signals at the equalizer outputs:

$$\begin{array}{lll} \text{A1)} & |K(z_j(k)| = |K_x|, & j = 1, \ldots, N_t \\ \text{A2)} & E|z_j(k)|^2 = \sigma_x^2, & j = 1, \ldots, N_t \\ \text{A3)} & E(z_l(k) z_k^*(k)) = 0, & l \neq j. \end{array}$$

Proving that the given set of conditions is necessary is quite straightforward, A1 and A2 are evident from (2.8) and A3 follows from the fact that $E(x_l(k)x_j^*(k)) = 0$ for $l \neq j$. In order to prove that these conditions are sufficient consider the following inequality

$$\sum_{l=1}^{l=N_t} |g_{jl}|^4 \leq (\sum_{l=1}^{l=N_t} |g_{jl}|^2)^2, \qquad j = 1, ..., N_t, \tag{2.9}$$

where the equality holds if and only if there exists a unique non-zero element g_{jl} of unit magnitude for each j. Now from (2.7), A1 and (2.6), A2 we obtain

$$\begin{aligned} \sum_{l=1}^{l=N_t} |g_{jl}|^4 = 1, \\ \sum_{l=1}^{l=N_t} |g_{jl}|^2 = 1. \end{aligned} \tag{2.10}$$

If we put together the inequality in (2.9) and (2.10), we can see that G_j must be of the form

$$G_j = [0 \ldots 0 e^{i\phi_j} 0 \ldots 0]^T, \tag{2.11}$$

where the single non-zero element can be in any arbitrary position. Now combining A3 and (2.4) gives

$$G_l^H G_j = 0, \qquad l \neq j. \tag{2.12}$$

According to (2.11) and (2.12), the N_t different vectors G_j of the form shown in (2.11) will contain their unique non-zero elements at different positions, which corresponds to the recovery of all N_t different inputs. Therefore, after reordering we obtain (2.8).

2.1.3 MUK criterion and MUK algorithm

The necessary and sufficient conditions for BSS discussed in the previous section form a basis for the following optimization criterion (known as the multiuser kurtosis maximization criterion) for blind signal recovery

$$\begin{aligned} \underbrace{max}_{\mathbf{G}} F(\mathbf{G}) &= \sum_{j=1}^{N_t} |K(z_j)| \\ subject\ to:\ \mathbf{G}^H\mathbf{G} &= \mathbf{I}_{N_t}. \end{aligned} \tag{2.13}$$

The constraint comes from the fact that, conditions A2 and A3 give

$$E(ZZ^H) = \sigma_x^2 \mathbf{I}_{N_t}. \tag{2.14}$$

Since, according to A1, A2 and A3, the global optimum of the MUK criterion achieve blind recovery, any adaptive algorithm that tries to optimize the MUK criterion has a chance of reaching the global desired solution (an adaptive algorithm may get trapped at a local optimum instead of reaching the global optimum under certain conditions).

There exist several algorithms in literature to optimize similar MUK criteria, but these algorithms have some drawbacks that make them unsuitable for practical implementations. Most of these algorithms do not guarantee convergence to a global optimum (corresponding to source separation) and may converge to a local optimum depending on initial conditions. The algorithm in (Delfosse & Loubaton april 1994*b*) uses a deflation structure for source separation and is globally convergent but it is not direct on the equalizer parameters (entries of matrix $\mathbf{W}$), and therefore requires a parameter transformation (from some parameter α optimized using stochastic recursion to equalizer matrix $\mathbf{W}$), which introduces bias and extra computational complexity. Another algorithm that exhibits global convergence (Tugnait jan. 1997) is the multistage CM algorithm, but its multistage structure (sources extracted one by one) imposes a large number of required iterations before convergence. In contrast, the adaptive stochastic-gradient with orthogonalization algorithm discussed here (Papadias dec. 2000b) (called the MUK algorithm thus far) is direct on the equalizer parameters, aims at joint recovery of all the sources (for any number of sources) and ensures convergence to the global optimum (even in the presence of Gaussian noise). This convergence to the global optimum is derived assuming channel matrix $\mathbf{H}$ to be full-rank, thus the maximum number of streams that can be recovered is limited by the degrees of freedom of the system. In the following we briefly introduce this algorithm. First we rewrite $F(\mathbf{G})$ as

$$\begin{aligned} F(\mathbf{G}) &= \sum_{j=1}^{N_t} sign(K(z_j))K(z_j) \\ &= sign(K_x)\sum_{j=1}^{N_t}(E|z_j|^4 - 2E^2|z_j|^2 - |E(z_j^2)|^2). \end{aligned} \tag{2.15}$$

Assuming symmetrical inputs ($E(x^2(k) = 0$) so that $E(z_j^2(k) = 0$)

$$F(\mathbf{G}) = sign(K_x)\sum_{j=1}^{N_t}(E|z_j|^4 - 2E^2|z_j|^2). \tag{2.16}$$

By invoking condition A2, $E|z_j(k)|^2$ is a constant, thus the gradient of F($\mathbf{G}$) with respect to $\mathbf{W}$ is given by

$$\nabla(F(\mathbf{G})) = 4sign(K_x)\sum_{j=1}^{N_t}[0_{N_r\times 1}, \ldots, E(|z_j(k)|^2 z_j(k)Y^*(k)), \ldots, 0_{N_r\times 1}], \tag{2.17}$$

where j^{th} term on the right side is a $N_r \times N_t$ matrix with all columns other than j^{th} column being zero vectors. Now, sum together all the terms on the right side into a single $N_r \times N_t$ matrix and update $\mathbf{W}$ in the direction of instantaneous gradient as follows (dropping the expectation operator in (2.17))

$$\mathbf{W_p}(k+1) = \mathbf{W}(k) + \mu\, sign(K_x)Y^*(k)Z_p(k), \tag{2.18}$$

where $Z_p(k)) = [|z_1(k)|^2 z_1(k), \ldots, |z_{N_t}(k)|^2 z_{N_t}(k)]$. For the next iteration the orthogonality constraint for the MUK criterion needs to be satisfied

$$\mathbf{G}^H(k+1)\mathbf{G}(k+1) = \mathbf{I}_{N_t}. \tag{2.19}$$

For this to be feasible we first need to make the channel matrix $\mathbf{H}$ unitary. This can be done by using any second order blind separation method (one such method is discussed in the next section) that typically converge to a unitary instantaneous mixture of inputs. Now, in order to satisfy the constraint given by (2.19), it suffices to ensure that $\mathbf{W}$(k+1) is unitary, i.e.

$$\mathbf{W}^H(k+1)\mathbf{W}(k+1) = \mathbf{I}_{N_t}. \tag{2.20}$$

Since, in general there is no guarantee that $\mathbf{W_p}(k+1)$ will satisfy the constraint given by equation (2.20), we use Gram - Schmidt orthogonalization on columns of $\mathbf{W_p}(k+1)$ and transform it into a unitary matrix $\mathbf{W}$(k+1). Now, keep on updating $\mathbf{W}$, while ensuring that it remains unitary after each update, until it converges to a fixed value.

From this description of the MUK algorithm it can be observed that we require some minimum number of symbol periods (say B) before the equalizer matrix $\mathbf{W}$ converges to a fixed value corresponding to a given channel matrix $\mathbf{H}$. Thus, the MUK algorithm can be successfully applied only if the transmit data block length is greater than B symbols periods. If the channel conditions change, another data block of at least B symbols is required to compute the new equalizer matrix. In (Papadias dec. 2000*b*) it is shown that the MUK algorithm requires about 1200 data symbols for convergence. A simple trick to reduce the number of data symbols needed for convergence is to use a data block of shorter length multiple times in the gradient update equation (2.18).

2.1.4 Channel whitening

In the last sub-section it was shown that in order to apply the MUK criterion directly on the equalizer parameters (coefficients of $\mathbf{W}$), we need to make the channel matrix unitary. Since we know that second order blind separation methods can separate an instantaneous mixture up to a unitary matrix, we can use them to whiten the received signal vector (Y)

$$Y_w = \mathbf{P}Y = \mathbf{PH}X + \mathbf{P}N = \mathbf{U}X + N_p, \tag{2.21}$$

where $\mathbf{P}$ is the pre-whitening matrix computed using second order separation methods and $\mathbf{U}$ is the whitened (unitary) channel matrix. There exists a simple algorithm for pre-whitening, using eigenvalue decomposition of the estimated covariance matrix $\hat{\mathbf{R}}_{YY}$ of the received signals

$$\hat{\mathbf{R}}_{YY} = \mathbf{LDL}^H, \tag{2.22}$$

where $\mathbf{L}$ is a unitary matrix and $\mathbf{D}$ is a real diagonal matrix. Given that ideally $\hat{\mathbf{R}}_{YY} = \sigma_x^2\mathbf{HH}^H$ (assuming $\mathbf{R}_{XX} = \sigma_x^2\mathbf{I}_{N_t}$), the non-zero part of the matrix $\mathbf{LD}^{1/2}$ equals $\mathbf{H}$ up to a unitary ambiguity matrix ($\mathbf{L}_w = \sigma_x\mathbf{HU}$). Matrix $\hat{\mathbf{L}}$ is constructed by taking N_t columns of $\mathbf{LD}^{1/2}$ with largest norms. This is so because these columns completely capture the transmitted signal, while other columns with smaller non-zero norms are present due to noise. This also reduces the computational complexity of the MUK algorithm. Next, we pre-filter the received signal as follows

$$Y_w = \mathbf{P}Y, \tag{2.23}$$

where $\mathbf{P}$ is the pseudo-inverse of matrix $\hat{\mathbf{L}}$. The effective channel between the transmitter and the receiver is now given by a unitary matrix $\mathbf{U} = \mathbf{PH}$.

2.2 Turbo-MUK for Single-User MIMO

In Section 2.1, we discussed a blind MUD structure (multiuser kurtosis maximization algorithm), which scores over numerous other blind MUD structures proposed in literature on account of its low computational complexity and global convergence in the presence of Gaussian noise. Inspired by previous works exploiting the turbo principle in coded CDMA systems (Wang & Poor july 1999) and coded space-time MUD (Dai et al. mar. 2004) we consider extending the same principle to MUK based coded blind MUD structure in order to improve its performance.

We propose a coded blind MUD structure that has an MUK with parallel interference cancellation detection stage and a soft decoding stage. The performance improvement offered by this receiver structure is due to the iterative exchange of soft information between the detection and the decoding stage. Figure 2.2 depicts the system set-up for this coded blind MUD scheme, which we name as turbo-MUK.

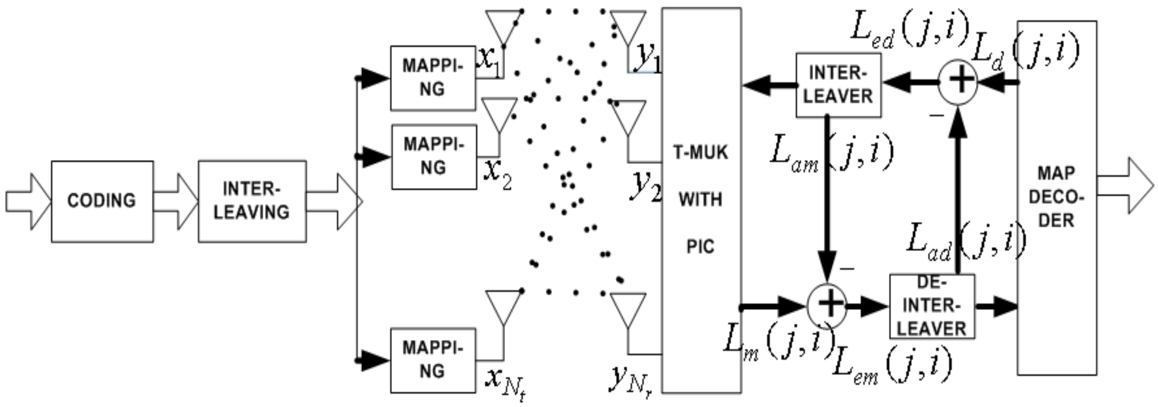

Figure 2.2: Structure of Turbo-MUK for single-user MIMO system

At the transmitter the information bits are first encoded (with convolutional or turbo coding). The coded bit stream is then interleaved and demultiplexed into N_t substreams. Each of these substreams is symbol-mapped and transmitted using a separate antenna from the transmit antenna array. At the receiver MUK criterion is used to decouple the substreams from the instantaneous mixture of transmitted signals received at N_r antennas. Then a soft metric $L_{em}(j, i)$ is computed for each coded bit (i) in a given substream ($j = 1, 2 \ldots N_t$) using apriori information $L_{am}(j, i)$ about these bits received from the decoder stage. The soft-metrics for different streams are multiplexed and de-interleaved and become apriori information L_{ad} for the decoder. The decoder employs maximum aposteriori probability (MAP) algorithm which uses the apriori information to generate extrinsic information L_{ed} about the coded bits. This soft information is interleaved and demultiplexed to give apriori information $L_{am}(j, i)$ used by the detector stage. After several such iterations between the detector and the decoder stages, soft information at the decoder output L_d becomes reliable to take hard decisions about the transmitted information bits.

2.2.1 Detector stage

Detector stage in the turbo-MUK is build upon the MUK detector discussed in the preceding section. The main added features in the turbo-MUK detector stage are its ability to use soft information from the decoder stage to improve the detection process and to generate soft information at its output. Also, it incorporates parallel interference cancellation, i.e., it tries to cancel out the interference from other streams while detecting a particular data stream. In the following discussion about the turbo-MUK detector stage we will stick to the system model used earlier (Section 2.1). The detection process starts with the application of the MUK algorithm on the pre-whitened received data block $\{Y_w(k)\}_{k=1,2,\ldots,B}$, corresponding to the block of information bits transmitted. This gives an equalizer matrix $\mathbf{W}$ that can be applied on the received data to get an estimate for the block of transmitted symbols

$$\begin{aligned} Z(k) &= \mathbf{W}^T Y_w(k) \\ &= \mathbf{G}^T X(k) + N_w(k), \end{aligned} \tag{2.24}$$

where $\mathbf{G}$ is the global response matrix. Its transpose is ideally (in a noise free environment) given by

$$\mathbf{G}^T = \mathbf{\Pi} \begin{bmatrix} e^{j\phi_1} & 0 & \ldots & 0 \\ 0 & e^{j\phi_2} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & \ldots & e^{j\phi_{N_t}} \end{bmatrix}, \tag{2.25}$$

where $\mathbf{\Pi}_{N_t \times N_t}$ is the arbitrary reordering matrix and $\phi_1, \phi_2, \ldots, \phi_{N_t}$ are the arbitrary phase rotations corresponding to N_t transmitted substreams. At high signal to noise ratios it is reasonable to approximate $\mathbf{G}^T$ with its ideal value given in (2.25). Also, we assume that the receiver has knowledge of reordering matrix as well as phase rotation terms corresponding to each substream. With the assumption that the background noise is circularly symmetric Gaussian with covariance $\mathbf{\Phi}_N = \sigma^2\mathbf{I}$, and the fact that combination of the pre-whitening matrix and the equalizer matrix effectively acts as a linear filter, the noise at the output of the equalizer will be Gaussian too with noise covariance matrix given by $\mathbf{\Phi}_{N_w} = \mathbf{W}^T\mathbf{P}\mathbf{P}^H\mathbf{W}^*\sigma^2$. With these assumptions in place we can say that the estimated symbols for the k^{th} transmitted sub-stream are Gaussian distributed and given by

$$z_j = \mu_j x_k + \eta_j, \tag{2.26}$$

where $\mu_j = [\mathbf{G}^T]_{jk}$ and η_j is a Gaussian random variable with zero mean and variance $\upsilon_j^2 = [\mathbf{\Phi}_{N_w}]_{jj}$. z_j is the estimate for the k^{th} transmitted symbol stream because of the ordering ambiguity inherent to blind detection. We can use the fact that estimated symbols have a Gaussian distribution to compute the soft information for the decoder stage in the following manner. Suppose each substream adopts M-ary quadrature amplitude modulation (M-QAM), then for each symbol interval $S_b = log_2 M$ bits are jointly detected for each substream. The detector delivers the aposteriori LLR of the code bit $b_j(i), 1 \leq i \leq S_b, j = 1, \ldots, N_t$

$$\begin{aligned} \wedge[b_j(i)] &= log\frac{P[b_j(i) = +1|z_j]}{P[b_j(i) = -1|z_j]} \\ &= log\frac{P[z_j|b_j(i) = +1]}{P[z_j|b_j(i) = -1]} + log\frac{P[b_j(i) = +1]}{P[b_j(i) = -1]}. \end{aligned} \tag{2.27}$$

The second term denoted by $L_{am}(j,i)$ represents the apriori information about the code bit $b_j(i)$ delivered by the decoder, whereas the first term denoted by $L_{em}(j,i)$ represents the extrinsic information delivered by the detector. Since, z_j is Gaussian distributed, the probability distribution of z_j given x_k is given by

$$P(z_j|x_k = x) = \frac{1}{(2\pi\upsilon_j^2)^{-1/2}} exp(\frac{(z_j - \mu_j x)^2}{2\upsilon_j^2}). \tag{2.28}$$

Denote $X_j^+ = \{x_k = x : b_j(i) = +1\}$ and $X_j^- = \{x_k = x : b_j(i) = -1\}$, the extrinsic information delivered by the detector can be written as

$$\begin{aligned} L_{em}(j,i) &= log\frac{\sum_{x_k \in X_j^+} P[z_j|x_k]P[x_k]/P[b_j(i) = +1]}{\sum_{x_k \in X_j^-} P[z_j|x_k]P[x_k]/P[b_j(i) = -1]} \\ &= log\frac{\sum_{x_k \in X_j^+} exp(\frac{(z_j - \mu_j x)^2}{2\upsilon^2})\Pi_{1 \leq l \leq S_b, l \neq i} P[b_j(l)]}{\sum_{x_k \in X_j^-} exp(\frac{(z_j - \mu_j x)^2}{2\upsilon^2})\Pi_{1 \leq l \leq S_b, l \neq i} P[b_j(l)]}, \end{aligned} \tag{2.29}$$

where $P[b_j(l) = -1] = \frac{1}{1+e^{L_{am}(j,l)}}$ and $P[b_j(l) = +1]\frac{e^{L_{am}(j,l)}}{1+e^{L_{am}(j,l)}}$. In the first iteration apriori information from the decoder stage is set as 0. After suitable reordering, the extrinsic information generated for each substream is multiplexed into a single stream, de-interleaved and then fed to the decoder stage as apriori information L_{ad}. The decoder uses this soft information to generate apriori information for the detector (after interleaving and demultiplexing), which completes the first iteration of the signal estimation process.

The next iteration starts with computing the least square estimate for the channel matrix using soft information from the decoder

$$\hat{\mathbf{H}} = Y\tilde{X}^H(\tilde{X}\tilde{X}^H)^{-1}, \tag{2.30}$$

where $\tilde{X}$ is the soft estimate for the transmitted signal vector. This soft information is also used to substract the interference signals from the received signal, which for some substream j gives

$$\begin{aligned} \tilde{Y}_j &= Y - \hat{\mathbf{H}}\tilde{X}_j \\ &= \mathbf{H}X - \hat{\mathbf{H}}\tilde{X}_j + N, \end{aligned} \tag{2.31}$$

where $\tilde{X}_j = [\tilde{x}_1, \ldots \tilde{x}_{j-1}, \tilde{x}_j = 0, \tilde{x}_{j+1}, \ldots \tilde{x}_{N_t}]$ is the estimated interference vector. Assuming near perfect interference cancellation (possible at high SNRs), now we can perform pre-whitening operation as if there was only one transmitter. The pre-whitening transform is a $1 \times N_t$ row vector P_j obtained by taking the pseudo-inverse of the largest norm column of the matrix $\mathbf{LD}^{1/2}$, where $\mathbf{L}$ and $\mathbf{D}$ are computed by eigenvalue decomposition of the covariance matrix $\mathbf{R}_{\tilde{Y}_j\tilde{Y}_j}$ (see Section 2.1 for definitions of $\mathbf{L}$ and $\mathbf{D}$). Pre-whitening of the received signal obtained after interference cancellation gives an improved estimate for the signals in the j^{th} substream

$$z_j = P_j\tilde{Y}_j. \tag{2.32}$$

The signal estimate is Gaussian distributed in the form given in (2.26) with $\mu_j = P_j\hat{h}_j$ and the variance of η_j given by $\upsilon_j^2 = P_jP_j^H\sigma^2$. This interference cancellation and whitening procedure is repeated for each of the N_t substreams. The extrinsic information about the transmitted signal bits can be computed using (2.29). This extrinsic information is again multiplexed, de-interleaved and fed to decoder stage for next iteration.

2.2.2 Decoder stage

The channel decoding stage of T-MUK uses the BCJR (L. R. Behl mar. 1974) decoding algorithm, which is better known for its ability to minimize probability of symbol error. However, the primary reason for selecting the BCJR algorithm (also known as the MAP algorithm) for the decoder stage here is the fact, that it can generate soft information in the form of log-likelihood ratios (LLRs) for the coded as well as the information bits, provided apriori LLRs for coded bits from the detection stage are available as inputs. In order to describe the functioning and main features of this algorithm we consider a binary $1/n$ convolutional code with the constraint length ρ (decoding for turbo codes is similar). Also, we try to adhere to the notation used in the original paper wherever possible for clarity.

We start by assuming that the input to the encoder block at time t is I_t and the corresponding output is $X_t = (x_t^1, \ldots, x_t^n)$. The state of the encoder at time t is given by

$$S_t = (s_t^1, \ldots s_t^{\rho}) = (I_t, \ldots I_{t-\rho+1}). \tag{2.33}$$

The encoder starts in state $S_0 = 0$. An information sequence of input bits (length L) followed by ρ zero bits causes the encoder to end in state $S_\tau = 0$, where $\tau = L + \rho$. For convolutional codes this encoding process can be represented by a trellis structure. At the decoder, LLRs for coded bits $X_t = (x_t^1, \ldots, x_t^n)$ and information bit I_t corresponding to stage t of the code trellis transiting from state $S_{t-1} = s'$ to $S_t = s$ (where $\{x_j\} \leftrightarrow \{x_t^k\}$ with $j = (t-1)n + k$ for a rate 1/n convolutional code) is given by

$$\wedge(x_t^k) = log\frac{\sum_{C_k^+} \alpha_{t-1}(s')\gamma_t(s', s)\beta_t(s)}{\sum_{C_k^-} \alpha_{t-1}(s')\gamma_t(s', s)\beta_t(s)}, \tag{2.34}$$

where C_k^+ is the set of state pairs (s', s) such that the k^{th} coded bit at stage t is 1 and C_k^- is the corresponding set for -1; and

$$\wedge(I_t) = log\frac{\sum_{I_k^+} \alpha_{t-1}(s')\gamma_t(s', s)\beta_t(s)}{\sum_{I_k^-} \alpha_{t-1}(s')\gamma_t(s', s)\beta_t(s)}, \tag{2.35}$$

where I_k^+ is the set of state pairs (s', s) such that the information bit at stage t is 1 and I_k^- is the corresponding set for -1. In LLR computation equations the term

$\gamma_t(s', s)$ denotes the transition probability for the branch $s' \rightarrow s$ for the stage t of the code trellis and is given by

$$\gamma(s', s) = P(S_t = s | S_{t-1} = s') = \prod_{l=1}^{n} P(x_t^l) = \prod_{l=1}^{n} \frac{1}{2}[1 + x_j tanh(\lambda_1^p(x_j))], \qquad (2.36)$$

where $\{x_j\} \leftrightarrow \{x_t^k\}$ with $j = (t-1)n + k$, and $\lambda_1^p(x_j)$ represents the corresponding apriori LLR delivered from the detector stage that is related to the state transition $s' \rightarrow s$. The other two terms, i.e., $\alpha_t(s)$ and $\beta_t(s)$ can be obtained by recursion as

$$\alpha_t(s) = \sum_{s'} \alpha_{t-1}(s')\gamma_t(s', s), \qquad t = 1, \ldots, \tau, \qquad (2.37)$$

and

$$\beta_t(s) = \sum_{s'} \beta_{t+1}(s')\gamma_{t+1}(s', s), \qquad t = \tau - 1, \ldots, 0, \qquad (2.38)$$

with boundary conditions given by

$$\alpha_0(0) = 1, \;\; and \;\; \alpha_0(s) = 0, \;\; for \; s \neq 0, \qquad (2.39)$$

and

$$\beta_\tau(0) = 1, \;\; and \;\; \beta_\tau(s) = 0, \;\; for \; s \neq 0. \qquad (2.40)$$

The summation of (2.37) and (2.38) are over all states s' for which the transition $s' \leftrightarrow s$ is possible. Since, we are using this decoder in an iterative set-up, to avoid positive feedback (which may result in instability) we are more interested in computing extrinsic information produced by the decoder given by

$$\begin{aligned} \lambda_2(x_t^k) &= log \frac{\sum_{C_k^+} \alpha_{t-1}(s')\beta_t(s) \prod_{l=1, l\neq k}^{n} P(x_t^l)}{\sum_{C_k^-} \alpha_{t-1}(s')\beta_t(s) \prod_{l=1, l\neq k}^{n} P(x_t^l)} \\ &= \wedge_2(x_t^k) - \lambda_1^p(x_t^k), \end{aligned} \qquad (2.41)$$

where $\wedge_2(x_t^k)$ is the LLR for coded bit computed in (2.34) and $\lambda_1^p(x_t^k)$ is the extrinsic information for the coded bit delivered by detector stage.

BCJR algorithm also has some drawbacks. The algorithm requires considerable storage and computation which grows with the constraint length and the block length. Also, large dynamic ranges of α_t and β_t result in numerical problems with the presentation of probabilities. Sub-optimal MAP decoding like Max-Log-MAP algorithm

(P. Robertson june 1995), where the computations are done in the logarithmic domain can be used to solve the later problem. In addition Max-Log-MAP decoding reduces the computational complexity at the cost of a small loss in performance by considering only 2 trellis paths, the ML path and the closest path to the ML path as compared to the MAP algorithm which takes into account all the paths in the trellis. With the substitution of $a_t = log(\alpha_t)$, $b_t = log(\beta_t)$ and $c_t = log(\gamma_t)$, (2.34) can be rewritten as

$$\wedge(x_t^k) = log(\sum_{C_k^+} exp(a_{t-1} + c_t + b_t) - log(\sum_{C_k^-} exp(a_{t-1} + c_t + b_t), \tag{2.42}$$

where

$$c_t(s', s) = \sum_{l=1}^{n} log(P(x_t^l)) = \sum_{l=1}^{n} log(\frac{exp(x_t^l \lambda_1^p(x_t^l))}{1 + exp(x_t^l \lambda_1^p(x_t^l))}). \tag{2.43}$$

The recursion equations to obtain $a_t(s)$ and $b_t(s)$ are given by

$$a_t(s) = log(\sum_{s'} exp(a_{t-1}(s') + c_t(s', s))), \tag{2.44}$$

with

$$a_0(0) = 0, \quad and \quad a_0(s) = -\infty, \qquad for \ s \neq 0, \tag{2.45}$$

and

$$b_t(s) = log(\sum_{s'} exp(b_{t+1}(s') + c_{t+1}(s, s'))), \tag{2.46}$$

with

$$b_\tau(0) = 0, \quad and \quad b_\tau(s) = -\infty, \qquad for \ s \neq 0. \tag{2.47}$$

Using the approximation $log \sum_i e^{x_i} \approx \underbrace{max}_{i} \ x_i$, equations (2.42), (2.44), (2.46) become

$$\wedge(b_t^k) = max_{C_k^+}[(a_{t-1} + c_t + b_t)] - max_{C_k^-}[(a_{t-1} + c_t + b_t)], \tag{2.48}$$

$$a_t(s) = \underbrace{max}_{s'}[a_{t-1}(s') + c_t(s', s)], \tag{2.49}$$

$$b_t(s) = \underbrace{max}_{s'}[b_{t+1}(s') + c_{t+1}(s, s')], \tag{2.50}$$

where (2.49) and (2.50) can be readily recognized as forward and backward Viterbi algorithms.

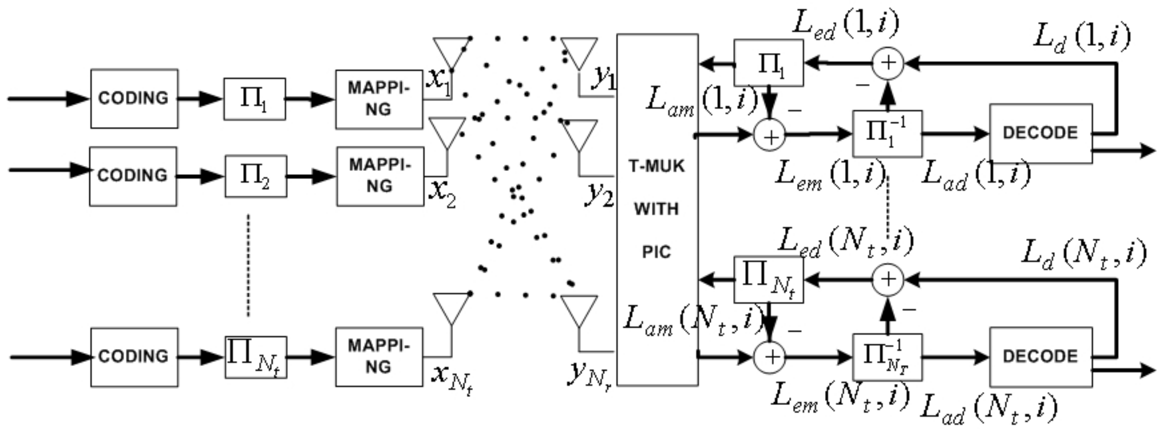

Figure 2.3: Structure of Turbo-MUK for multi-user MIMO system

2.3 Turbo-MUK for Multi-User MIMO

If we consider a number of users (say N_t), with single antenna, transmitting independent data streams, and a receiver with multiple antennas (say N_r), the whole set-up can be seen as an $N_r \times N_t$ distributed MIMO system. If all the assumptions about the transmitted signals and the MIMO channel made in last section hold, then we can use the single-user T-MUK structure for blind multi-user detection, with some minor modifications, for jointly detecting signals transmitted from multiple users (transmitters). Figure 2.3 depicts the T-MUK structure for multiple users. The only difference from the single-user structure is that at the transmitter each substream (corresponding to a particular user) is encoded (convolutional or turbo coding) independently. At the receiver, the MUK detector separates different signal streams and generates soft information as done earlier, but now instead of multiplexing and joint decoding, each signal stream has an independent decoder. Each of these decoders uses the soft information about a given stream from the detector and computes soft information (about the same stream) to be fed back to the detector.

A possible scenario for distributed MIMO system is, different users transmitting at different power levels. In Section 2.1 one of the constraints on the transmitted signal streams, for the MUK criterion to be applicable, is that these streams should be i.i.d. In a multi-user scenario signal streams from different users can be assumed to be independent, also we can assume that all of them have zero mean (assuming all users implement same symmetric modulation scheme, such as 4-QAM), but different signal streams may have different power levels (variances). In the previous discussion about MUK algorithm we assume signal streams with equal power levels, here we show that the channel pre-whitening step of the MUK algorithm enables it to handle the unequal power case as well.

Ideally, estimated covariance matrix for the received signal is given by

$$\begin{aligned}\hat{\mathbf{R}}_{YY} &= \mathbf{H}\mathbf{R}_{XX}\mathbf{H}^H \\ &= \mathbf{H}\begin{bmatrix} \sigma_1^2 & \dots & 0 \\ 0 & \sigma_2^2 & \dots \\ \vdots & \vdots & \vdots \\ 0 & \dots & \sigma_{N_t}^2 \end{bmatrix}\mathbf{H}^H,\end{aligned} \tag{2.51}$$

where $\sigma_1^2, \sigma_2^2, \dots, \sigma_{N_t}^2$ denote the variance (power level) of N_t users. Y, X and $\mathbf{H}$ denote the received vector signal, transmitted vector signal and channel matrix respectively. Running eigenvalue decomposition on $\hat{\mathbf{R}}_{YY}$ gives the decomposition $\hat{\mathbf{R}}_{YY} = \mathbf{L}\mathbf{D}\mathbf{L}^H$ (see Section 2.1), where the non-zero part of the matrix $\mathbf{L}\mathbf{D}^{1/2}$ is now given by

$$\tilde{\mathbf{L}} = \mathbf{H}\mathbf{U}\begin{bmatrix} \sigma_1 & \dots & 0 \\ 0 & \sigma_2 & \dots \\ \vdots & \vdots & \vdots \\ 0 & \dots & \sigma_{N_t} \end{bmatrix}, \tag{2.52}$$

where $\mathbf{U}$ is a unitary matrix and we construct $\tilde{\mathbf{L}}$ as a matrix that contains N_t columns of $\mathbf{L}\mathbf{D}^{1/2}$ with largest norms. We now pre-filter the received signal as

$$\tilde{Y} = \mathbf{P}Y, \tag{2.53}$$

where $\mathbf{P}$ denotes the pseudo-inverse of $\tilde{\mathbf{L}}$ and is given by

$$\mathbf{P} = \begin{bmatrix} 1/\sigma_1 & \dots & 0 \\ 0 & 1/\sigma_2 & \dots \\ \vdots & \vdots & \vdots \\ 0 & \dots & 1/\sigma_{N_t} \end{bmatrix} \mathbf{U}^\dagger \mathbf{H}^\dagger, \tag{2.54}$$

where $\dagger$ denotes pseudo-inverse. From (2.53) and (2.54) it is obvious that pre-whitening removes the power inbalance from the input signal streams. Thus, pre-whitening effectively results in i.i.d. transmitted streams, which makes the MUK algorithm work for different users with unequal powers.

2.4 Numerical Results

In Sections 2.2 and 2.3 we have introduced T-MUK detection for single-user and multi-user MIMO systems. We now highlight the effectiveness of T-MUK through numerical results. In Section 2.2, while describing T-MUK we assumed that the receiver has phase and ordering information before it computes the soft information about the coded bits in the detector stage. However, an actual implementation of T-MUK will require the receiver to extract this information from the received signals. Therefore, we begin this section by listing the phase and ambiguity removal techniques we use in our simulations to obtain numerical results for T-MUK.

It is shown in (Guo & Wang april 2003) that ordering ambiguity can be easily removed by tagging the transmit substreams with different orthogonal spreading codes. We believe that this technique for ordering ambiguity removal can easily be extended to T-MUK also, at the cost of small increase in complexity. However, in our simulations we estimate the arbitrary ordering matrix as $\Pi = |E[Z'X]|$ by using all transmitted data, where Z is the equalizer output and X is the transmitted vector signal. For removing phase ambiguity, standard technique suggested in literature is to use differential encoding, but we found that it is a non-trivial problem to incorporate this technique in a blind receiver using iterative turbo detection. The main problem with the use of differential encoding is that, the soft information (about the coded bits) computed by the decoder needs to be differentially encoded again, before being fed to the detector for next iteration. Now, if the decoder soft estimate about a particular bit is not reliable enough, the inherent structure of the differential encoder (any output bit depends on all previously fed input bits) results in an error propagation to all subsequent bits. Thus, if we use differential encoding we cannot have an iterative information exchange between the detector and the decoder stages. In our simulations we use a very small training sequence (2 symbol periods) to remove the phase ambiguity (thereby making T-MUK a semi-blind MUD instead of completely blind MUD).

We simulate a $N_r \times N_t$ MIMO system, wherein each entry of the channel matrix $\mathbf{H}_{N_r \times N_t}$ is chosen from a complex Gaussian distribution of zero mean and unit variance $h_{ij} \sim N(0,1)$. Each substream employs a 4-QAM modulation and the coding scheme used is rate 1/3, 64 state convolutional code with generators $(G_1, G_2, G_3) = [155, 117, 123]_8$ (code proposed for EDGE). We transmit information in blocks of 384 bits over a quasi-static channel (i.e., channel stays in one state for one block and then changes to another value for next block and so on). We record the bit error probability of single-user T-MUK and multi-user T-MUK (users having unequal power levels) over a large number of data blocks.

The simulation results (BER vs SNR plots) for single-user T-MUK (6×4 MIMO system) are shown in Figure 2.4, from which we can see that first iteration of T-MUK detection gives a gain of more than 8dB as compared to the uncoded MUK detection at a BER of 10^{-3}. Next few iterations, which involve parallel interference cancellation along with the exchange of soft information increase this gain by another $\sim 1.5dB$. We can see that we require just 3 iterations to achieve convergence and get an improved estimate.

Figure 2.5 depict the simulation results for multi-user T-MUK (6×4 MIMO system), where all the users have different power levels and the power of the strongest user is 3db more than that of the weakest user. We can observe that T-MUK gives a gain of greater than 5db over uncoded MUK in the first iteration itself and another $\sim 1dB$ gain in next couple of iterations at a BER of 10^{-3}.

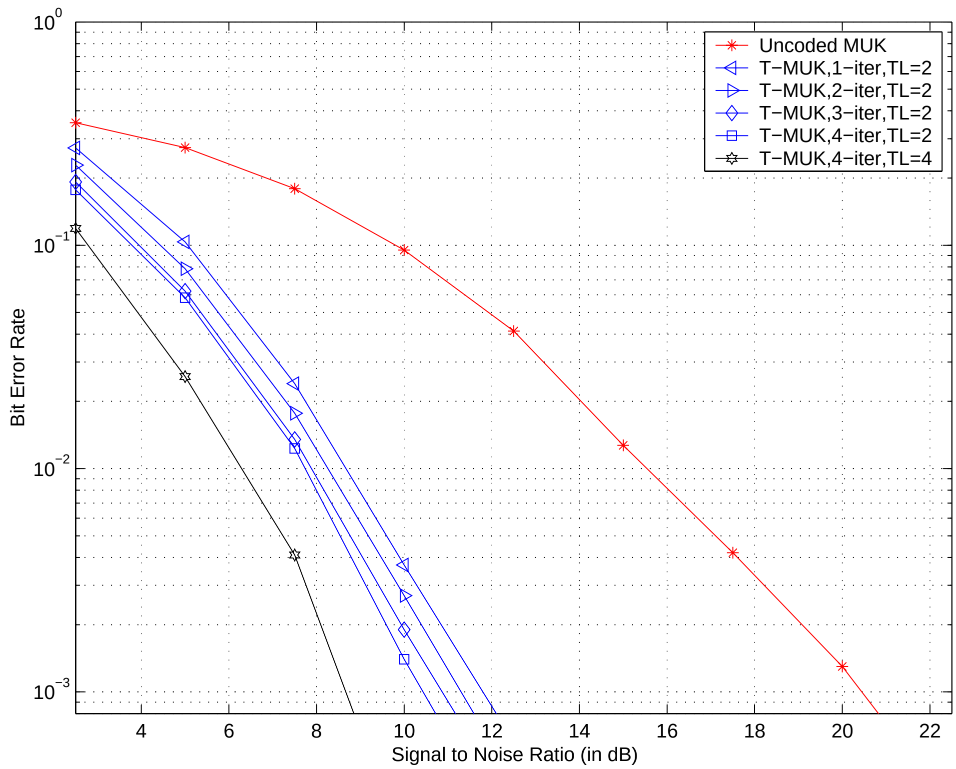

Figure 2.4: Performance of Single-User Turbo-MUK

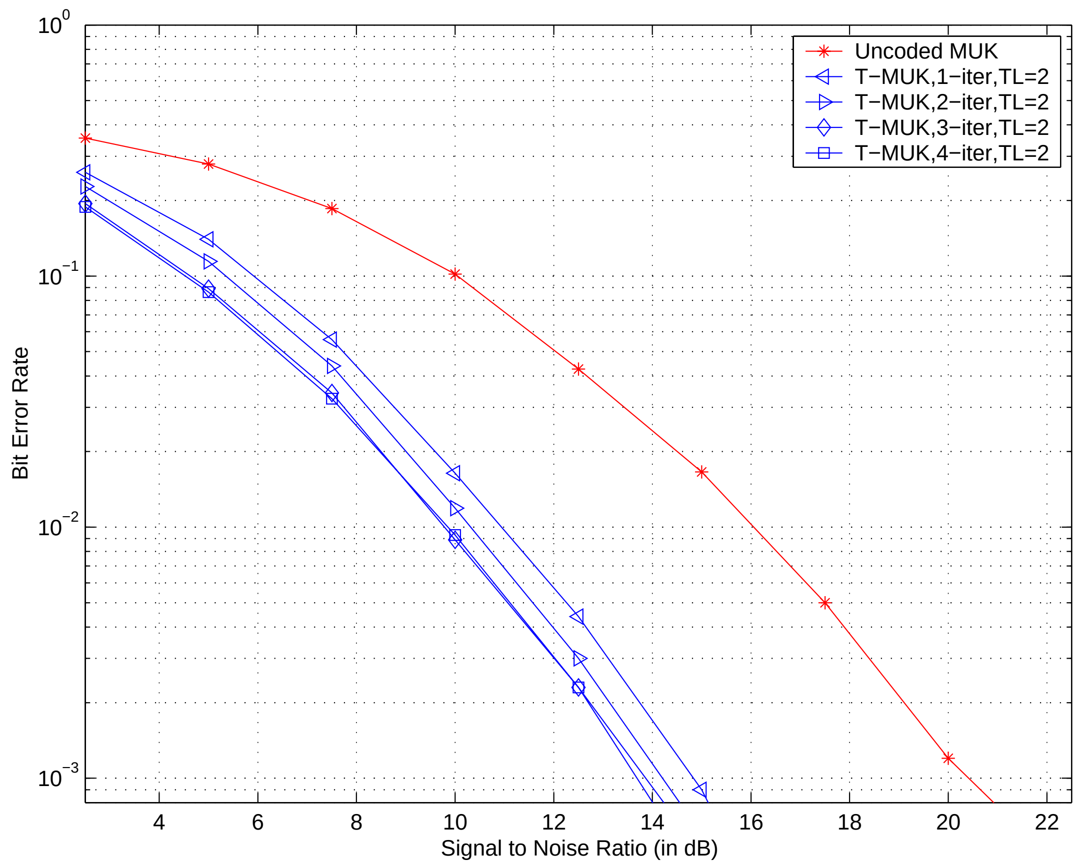

Figure 2.5: Performance of Multi-User Turbo-MUK (Unequal Powers)

2.5 Summary

In (Hassibi & Hochwald april 2003) it is shown that in order to get a meaningful estimate in multiple antenna wireless links, training length should be greater than or equal to N_t symbol periods (where N_t =number of transmit antennas) at high SNRs and it asymptotically approaches half the coherence interval (block length) as SNR is reduced. Therefore, conventional detection structures that rely on training for channel estimation require at least N_t symbol period training. In contrast to this T-MUK structure is shown to perform satisfactorily with a 2 symbol period training length and suits better for use on fast time-varying channels. Also, in the case where the transmitter has a large transmit antenna array the difference between total number of training symbols required by conventional detectors ($O(N_t^2)$) and training symbols required by T-MUK ($O(N_t)$) becomes significant, thereby making T-MUK a better choice in terms of bandwidth usage.

Chapter 3

Semi-Blind interference cancellation

3.1 Introduction

It has been shown that the spectral efficiency of MIMO wireless communication systems grows linearly with the minimum of the number of transmit and receive antennas, when operating on a single link (intra-cell interference only) with Gaussian noise. However, in a cellular environment, co-channel or inter-cell interference (CCI) becomes dominating channel impairment and the capacity of a MIMO system is hardly larger than when using multiple antennas at the receiver only (Cetreux, driessen & Greenstein nov. 2000). In a recent work (Dai et al. mar. 2004) some advanced multiuser detection (MUD) structures have been suggested that are better suited for interference-limited environment as compared to the previously known detection structures (such as coded V-BLAST, T-BLAST etc). Although these novel MUD structures substantially improve the system performance, they dictate perfect channel knowledge about the interferers. This is a rather difficult condition to satisfy because typically SNR for signals from the adjacent cells is worse than that for the desired cell. Rapid variation of channel conditions or asynchronous reception of signals from different cells makes this problem even more complex. This motivates us to investigate an advanced MUD structure that can improve the performance in

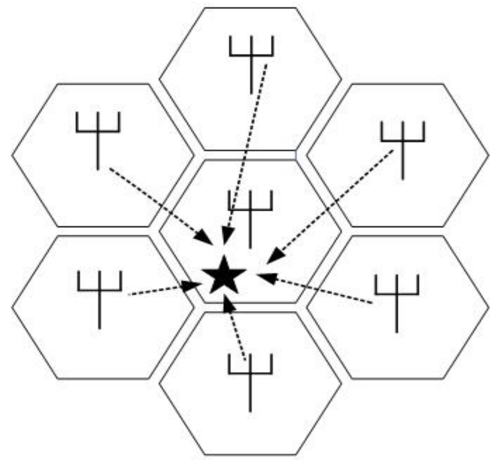

Figure 3.1: Multicell MIMO system (downlink). Each transmitter has N_t antennas and each receiver has N_r antennas

an interference-limited environment without any channel knowledge about the interferers. The receiver structure proposed here employs T-BLAST (discussed later) to detect the desired signal and T-MUK to deal with inter-cell interference and iterates between the two in a group multistage fashion to further improve the performance.

3.2 Multicell MIMO System Model

In order to compare the performance of different MUD structures for a cellular system, we consider a multicell structure where each transmitter has N_t antennas and each receiver has N_r antennas. We take into account interference from the first tier of the center-excited cell configuration and assume a frequency-flat, quasi-static fading environment. The complex path loss between the j^{th} transmit and the i^{th} receive antenna is composed of three components, path loss (that depends on the link length and the path loss exponent), loss due to shadow fading and multipath fading loss. The baseband channel gain between a pair of transmit and receive antennas is modeled as

$$h_{ij} = \sqrt{c\frac{1}{d_{ij}^{\gamma}}\sqrt{s_{ij}}}\left[\sqrt{\frac{R}{R+1}}e^{i\phi_{ij}} + \sqrt{\frac{1}{R+1}}m_{ij}\right], \tag{3.1}$$

where d_{ij} is the length of the link, γ is the path loss exponent, c is the propagation constant, s_{ij} is the log-normal shadow fading variable, R is the Ricean factor, which denotes the ratio of the direct received power to the average scattered power, ϕ_{ij} is the phase shift of the direct path and m_{ij} is scattered component, modeled as a normalized Gaussian variable. In the following discussion, the model is further simplified by assuming absence of large obstacles and the LOS path, so that shadow-fading loss and the effect of Ricean factor on channel gain can be neglected.

In this study, the multicell MIMO system model is a straightforward extension of the single cell MIMO system model discussed in Chapter 2, and is given by

$$Y = \mathbf{H}X + \sum_j \mathbf{H}_{ifj} X_{ifj} + N, \tag{3.2}$$

where X_{ifj} denotes the transmitted signal from the j^{th} interfering user, $\mathbf{H}_{ifj}$ denotes the matrix capturing channel characteristics between the j^{th} interfering user (in an adjacent cell) and the receiver in the cell of interest. The channel matrices $\mathbf{H}$ and $\{\mathbf{H}_{ifj}\}$ are mutually independent, with i.i.d. normalized complex Gaussian elements. The transmitted signals from the desired and interfering users are constrained to have a total transmit power of $E[X^H X] \leq P$ and $E[X_{ifj}^H X_{ifj}] \leq P_{ifj}$, respectively, which includes the path loss factor. The noise is assumed to be white and circularly symmetric Gaussian with covariance $\mathbf{\Phi}_N = \sigma^2 I$. The signal to noise ratio (SNR) is given by $\rho = P/\sigma^2$, and the signal to interference ratio (SIR) is given by $\eta = P/\sum_j P_{ifj}$. The multicell system model discussed here is valid for the downlink as well as the uplink of a cellular system. Figure 3.1 depicts an exemplary cellular downlink scenario.

A special case of this multicell system model is: the transmitter and the receiver in the cell of interest has N_t and N_r antennas respectively, whereas transmitters in the adjoining cells (interfering users) have single antenna. The data streams from different users can be assumed to be mutually independent. These (say N_{ts}) interfering users together with the receiver in the cell of interest form an $N_r \times N_{ts}$ distributed MIMO system, as discussed in Section 2.3. In Figure 3.2 we demonstrate such a system for the cellular downlink scenario.

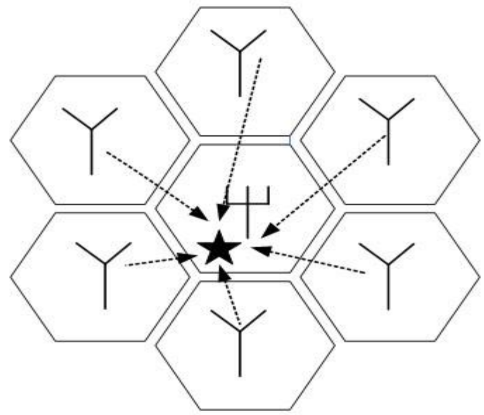

Figure 3.2: Multicell MIMO system model. Each interfering transmitter has a single antenna and each receiver has N_r antennas

3.3 Group Successive Interference Cancellation

In (Varanasi july 1995), the concept of group detection was introduced, where multiuser detection is performed in a sub-optimal fashion by jointly detecting a group or subset of all users. In simple terms group detection can be interpreted as a cascade of a decorrelator (for curbing the users outside the group) followed by an optimal detector for the users in the group of interest. Another interesting idea (Wyner jan. 1974) originally developed in the context of scalar-output Gaussian multiaccess channels, is that of successive interference cancellation (SIC). It involves decoding the users sequentially, that is, the first user is decoded by regarding interference from other users as noise. The decoded and re-encoded symbols of the first user are then subtracted from the received data and the second user is decoded by regarding the interference from remaining users as noise and so on. This strategy achieves the total capacity of Gaussian multiaccess channel.

In a multicell MIMO system, the focus of signal-processing at a given receiver is to detect the desired signal substream in presence of other substreams from the transmitter in the same cell (intra-cell interference) and substreams from transmitters in the adjacent cells (inter-cell interference). In (Dai et al. mar. 2004) a multiuser detection (MUD) scheme for multicell MIMO (Group IC MUD) is investigated that combines group detection with successive interference cancellation. In this scheme data streams from each cell are considered as a separate group. Information from one group is detected at a time (at the same time try to curb signals from all other groups) and fed to other groups for successive interference cancellation. The success of this scheme relies on the correct detection of interference. If the estimate for interference is poor, interference cancellation scheme will worsen the performance instead of improving it.

The group IC MUD scheme as well as other MUD for multicell MIMO systems, such as MMSE MUD and ML MUD, require perfect channel state information (CSI) for the transmitter in the given cell as well as interfering transmitters (channel matrices $\mathbf{H}$ and $\mathbf{H}_{ifj}$ respectively) for good performance. As pointed out earlier, inherent structure of cellular set-up makes this a difficult task and it requires long training sequences from interfering transmitters to get reliable CSI. Even if we assume that channel changes slowly and it is possible to send a long enough training sequence, getting good estimate for $\mathbf{H}_{ifj}$ is rather difficult in a cellular environment due to lack of synchronism between transmitters in different cells.

The simplest method to estimate channel, given training sequences from all transmitters is to use least square estimation

$$\hat{\mathbf{H}}_{if} = Y_{tr} X_{if,tr}^{H} (X_{if,tr} X_{if,tr}^{H})^{-1}, \tag{3.3}$$

where the subscript 'tr' denotes training symbols. But if we perform this estimation for an individual transmitter (3.3), signals from other transmitters add to noise level resulting in a poor channel estimate. The other feasible approach is to use joint least square estimation (Ranta, Hottinen & Honkasalo 1995) and estimate all the channel matrices simultaneously

$$\hat{\mathbf{H}}_{T} = Y_{tr} X_{T,tr}^{H} (X_{T,tr} X_{T,tr}^{H})^{-1}, \tag{3.4}$$

where $\mathbf{H}_T = [\mathbf{H} \;\; \mathbf{H}_{if}]$ and $X_{T,tr} = [X_{tr}^T \;\; X_{if,tr}^T]^T$. This is where synchronism between transmitters becomes crucial, if training sequences from different transmitters are time aligned (case X) then we can get very reliable channel estimates (see Figure

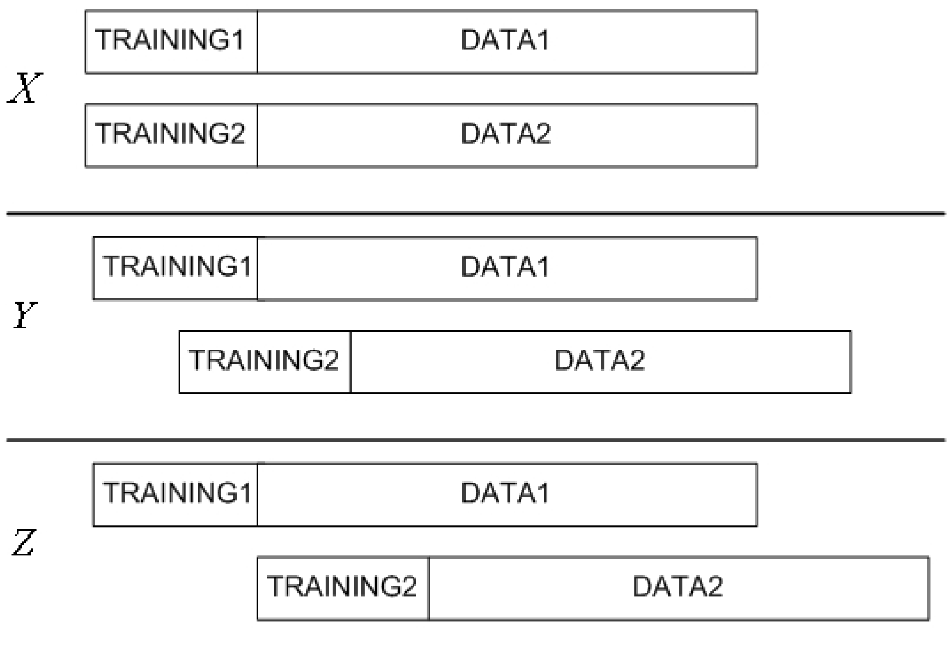

Figure 3.3: Effect of asynchronous training sequences

3.3). But if these sequences are partly overlapping or do not overlap at all (case Y and case Z) then the quality of channel estimates deteriorate depending on the extent of overlap between various sequences (Nordstrom, Holst & Lindoff 2002). In contrast to this case the semi-blind successive interference cancellation structure proposed in Section 3.5 can work well even when training sequences (very small in length) do not overlap, because these sequences are used for removing phase ambiguity and not for joint channel estimation and can occupy any position in a given data frame.

3.4 T-BLAST

Turbo-Bell labs space-time layered architecture (T-BLAST) is an example of application of turbo multiuser detection to the well-known BLAST architecture. In this scheme, at the transmitter the information bits are coded and interleaved as a whole, then demultiplexed into N_t substreams and symbol-mapped individually (Figure 3.4).

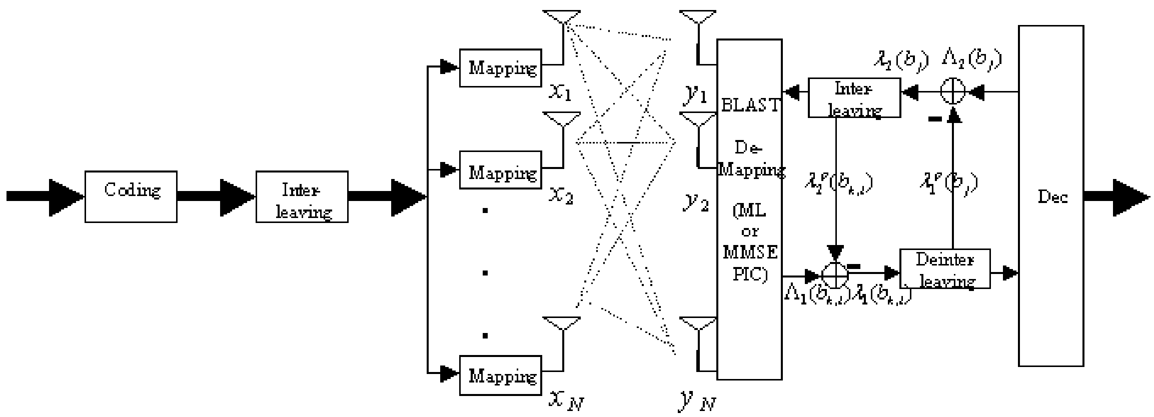

Figure 3.4: Structure of T-BLAST

At the receiver, we process the entire data stream iteratively between a soft detection (ML or MMSE multistage parallel interference cancellation (PIC)) and a soft decoding stage (MAP). Since, T-BLAST brings in full receive diversity due to joint coding, it approaches very close to capacity (shown in (Dai et al. mar. 2004)) for a single link MIMO channel (no interference). T-BLAST is an important block in the semi-blind IC-MUD discussed next and also serves as a lower-bound (baseline case) on the performance of the MUD designed for multi-cell MIMO systems. T-BLAST fails to achieve capacity in the interference-limited (multicell) scenario.

3.5 Semi-Blind Successive Interference Cancellation

Based on the Turbo-MUK algorithm discussed in Chapter 2, a new receiver structure is proposed for MIMO systems in a cellular environment. In contrast to traditional single-cell MIMO receivers, it actively combats the CCI in the interference-

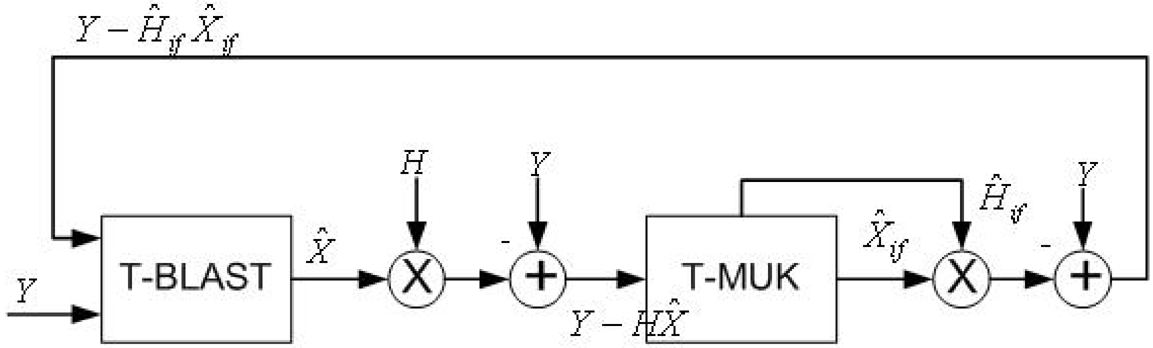

Figure 3.5: Semi-blind group interference cancellation detector

limited environment. Also different from recent work (Dai et al. mar. 2004) that assumes perfect CSI for interferers, it is a semi-blind approach which may assume certain advantages in practice as mentioned earlier.

Our proposed structure assumes a group interference cancellation nature, as shown in Figure 3.5. The detection process is initiated by using turbo-BLAST on the received signal vector Y to get an estimate of the transmitted vector signal $\hat{X}$, which treats any intercell interference as additional noise. Then with the reconstructed desired user's component subtracted from the received signal (assuming perfect channel information $\mathbf{H}$ for the desired cell is available), the resultant vector signal is given by

$$Y_{if} = Y - \mathbf{H}\hat{X}. \tag{3.5}$$

We then employ turbo-MUK on this vector signal Y_{if} to estimate channel $\hat{\mathbf{H}}_{if}$ and transmitted vector signal $\hat{X}_{if}$ for some strongest interfering user(s). With these estimates we attempt to eliminate their detrimental impact from the received vector signal, and feedback a "cleaner" copy to the turbo-BLAST block as

$$Y_S = Y - \hat{\mathbf{H}}_{if}\hat{X}_{if}. \tag{3.6}$$

We expect turbo-BLAST to output a better estimate for the desired transmitted vector signal $\hat{X}$ this time due to reduction in interference power. This iterative procedure is repeated until certain convergence is reached. The same set-up can be used for the multicell MIMO where the interferers have single transmit antenna whereas the transmitter and the receiver in the cell of interest are equipped with multiple antennas (Section 3.2). In this case T-MUK is used to jointly detect interfering signals from different users.

Some comments are readily in order. Due to lack of degrees of freedom, turbo-MUK can be employed only to deal with some of the strongest interferers. This is realizable with the pre-whitening operation as we mentioned in Chapter 2. This approach actually conforms to our current understanding on the interference channel. As is known, detection of the interfering users is not always optimal except in the strong-interference case, nor is treating them as pure ambient noise optimal, except when they are very weak. As the other side of the same coin, this group IC approach works well at high SNR regimes with sufficient power imbalances among users, and the performance is expected to deteriorate otherwise.

3.6 Numerical Results

In this section we justify the usefulness of semi-blind MUD we described in previous sections with some numerical results obtained through computer simulations. The modulation scheme employed is 4-QAM and the coding scheme used is a rate 1/3, 64 state convolutional code with generators $(G_1, G_2, G_3)[155, 117, 123]$. Ordering and phase ambiguity removal techniques used are similar to that discussed earlier in the context of turbo-MUK. We transmit data blocks of 384 bits and record the block error probability for different multicell systems. The baseline for comparison in all systems is the performance of the original T-BLAST. All simulations involve 2 iterations (for soft information exchange) between the decoder and detector stage for the T-BLAST block, 3 iterations between the detector and decoder stage for the T-MUK block and 3 iterations between the T-BLAST and the T-MUK blocks (for interference cancellation).

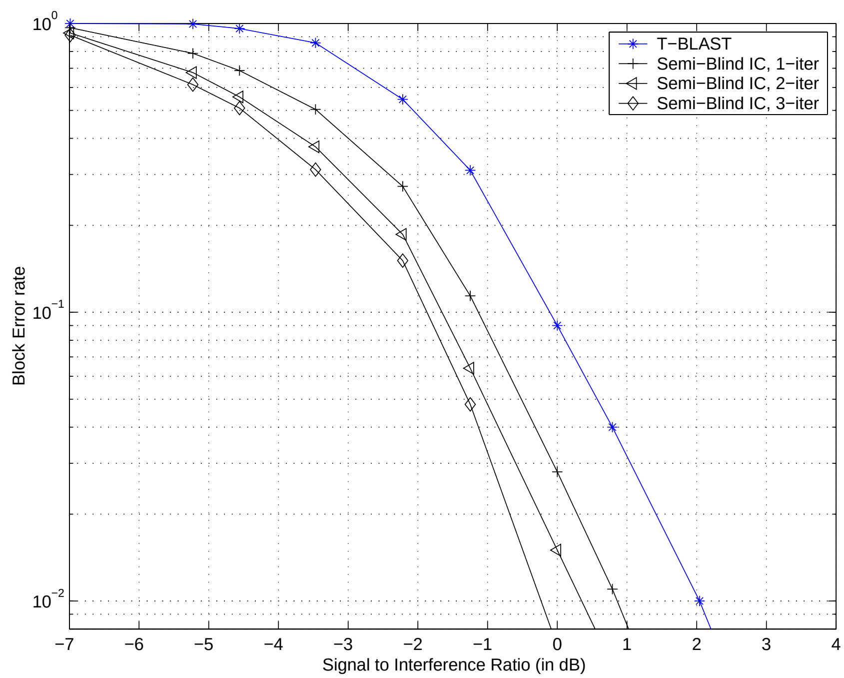

Figure 3.6: Results for a multicell system having interferers with multiple antennas, where the strongest interferer is 6db stronger than other interferers.

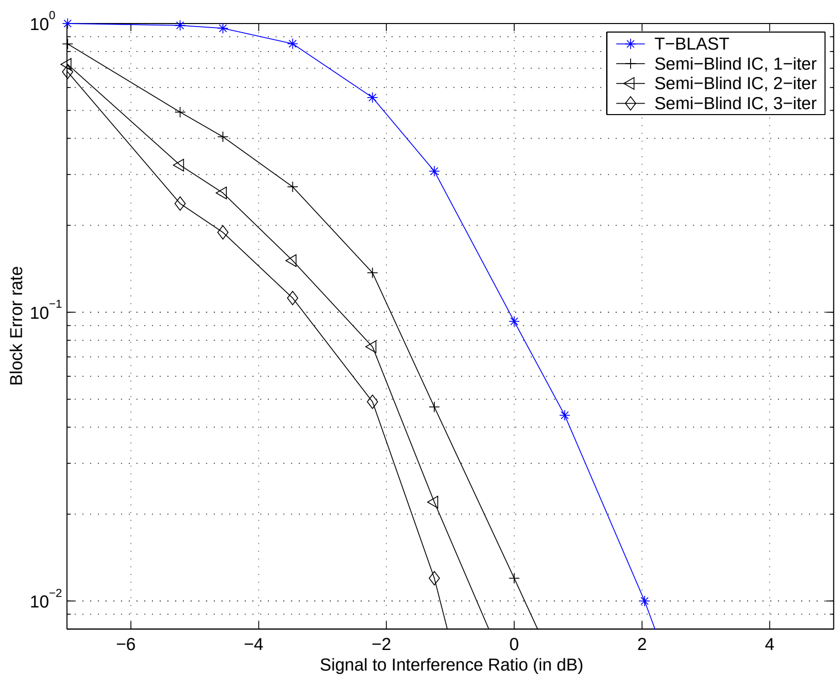

Figure 3.7: Results for a multicell system having interferers with multiple antennas, where the strongest interferer is 9db stronger than other interferers.

We first consider a system where each cell employs a 6×4 MIMO system, with interference to noise ratio fixed at 30db. We assume that one of the adjacent cell interferer is much stronger than the other ones. The simulation result in Figure 3.6 depicts the case where the power of the dominant interferer is 6db stronger than the sum of the rest, from which we can observe that the semi-blind MUD offers a performance gain of $\sim 2db$ over the lower-bound. In Figure 3.7 we show the simulation result for the same system with 9db power difference between the interferers, and the performance gain offered by the semi-blind MUD increases to $\sim 3db$ in this case. The decrease in performance gain with decrease in power difference between the interferers is due to the limited degrees of freedom available at the receiver. Since, the maximum number of different data streams a MUK detector can separate out is limited by the number of receive antennas, in the above described system turbo-MUK can separate out the data streams corresponding to only one of the interferers. The turbo-MUK detector selects the data streams corresponding the strongest interferer without any knowledge about the channel or the relative power levels of interferers, whereas the data streams from the other interferers add to background noise. Thus, the higher is the power level of weaker interferers, the higher is the effective noise floor and greater is the performance degradation suffered by the semi-blind MUD.

We then consider another interesting multicell MIMO system, where the cell of interest employs a 4×4 MIMO system, whereas each interferer in adjacent cells is equipped with single antenna, with interference to noise ratio fixed at 30db. We assume interfering users with differing power levels such that the strongest user is 3db stronger than the weakest user. The simulation results for this situation are shown in Figure 3.8. The performance gain offered by semi-blind MUD is $\sim 3db$, which not only highlights its effectiveness in combating inter-cell interference but also shows its ability to handle interferers with unequal power levels. Now, we modify this situation to show the limitation imposed on performance of the semi-blind MUD by the available of degrees of freedom at the receiver. Instead of ignoring the 2 interferers in the farthest adjacent cells, we assume that these interferers are 6db weaker than the strongest interferer and re-simulate this scenario. The simulation results (Figure 3.9) indicate that the performance gain is reduced to $\sim 1.5dB$. The reason behind

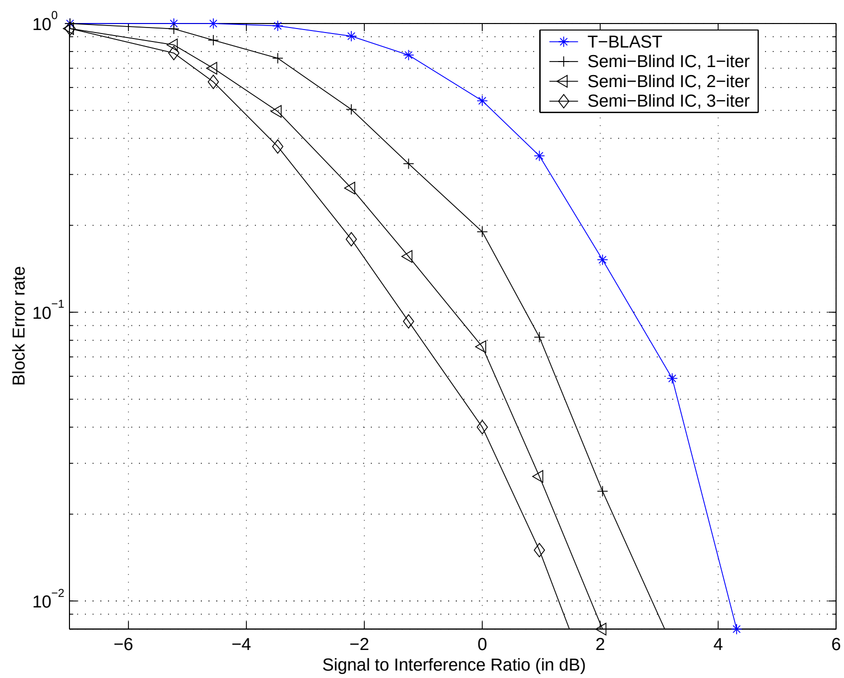

Figure 3.8: Results for a multicell system having 4 adjacent cell interferers with single antenna, where the strongest interferer is 3db stronger than the weakest interferer.

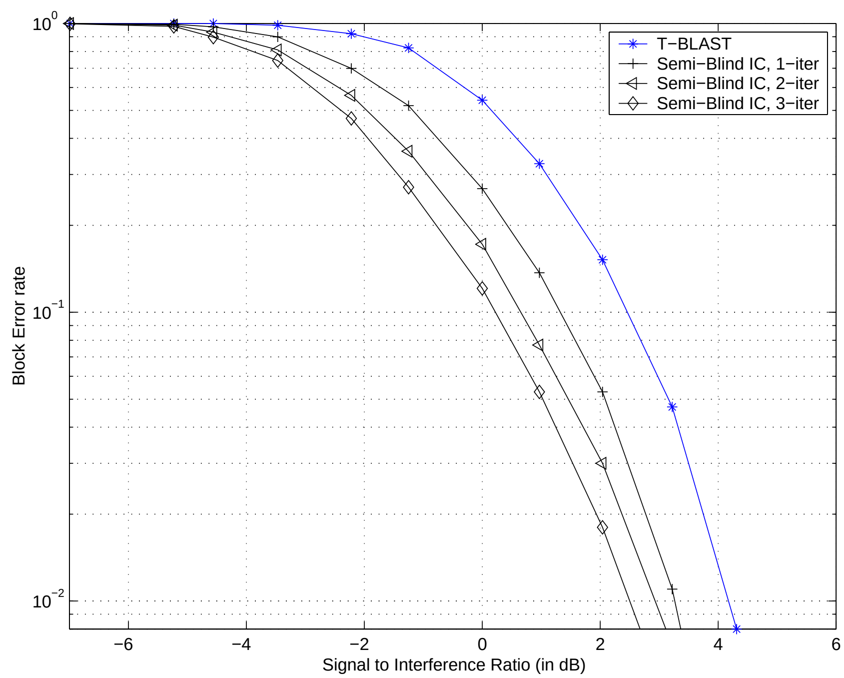

Figure 3.9: Results for a multicell system having 6 adjacent cell interferers with single antenna

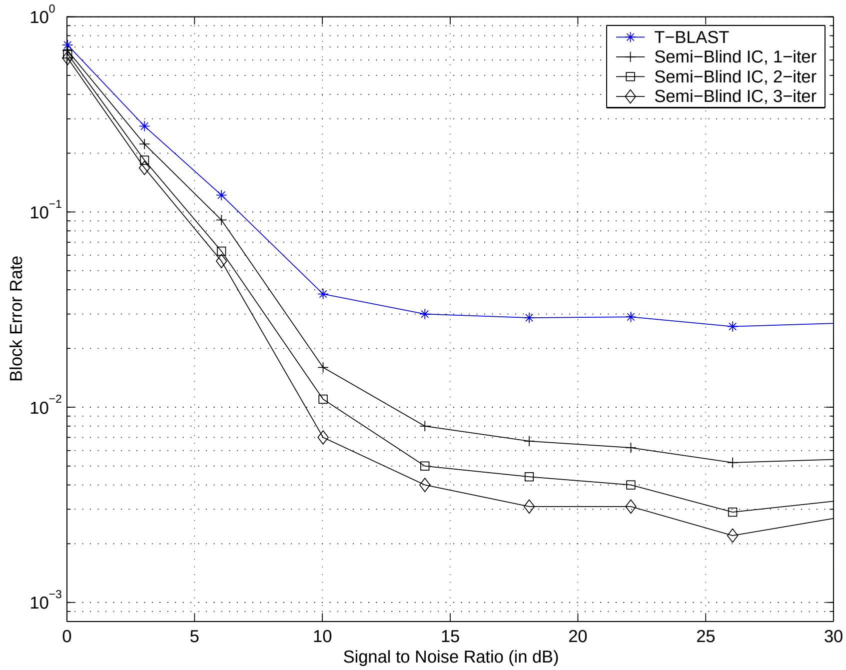

Figure 3.10: Results for an interference-limited multicell system having interferers with multiple antennas, where the strongest interferer is 6db stronger than other interferers. SIR = 1db

this behavior is that the turbo-MUK detector selects the 4 strongest users (limited by 4 receive antennas at the receiver), whereas the 2 weaker users add to the background noise level.

Since, the main focus of the semi-blind MUD is to combat inter-cell interference, we study the various systems discussed so far in an interference-limited environment also, which helps us get a better idea about the effectiveness of this MUD. In Figure 3.10 and Figure 3.11 we have shown the simulation results for 6×4 MIMO systems with the power difference between interferers set at 6db and 9db respectively. The SIRs are fixed at 1db and 0db respectively and the noise floor is -21db in both cases. Similarly, Figure 3.12 and Figure 3.13 show the results for 4×4 MIMO systems with single antenna interferers. The SIRs in these cases are fixed at 2db and 3dB respectively. The feature common to all these results is that the error floor attained by semi-blind MUD in high SNR region is much lower as compared to error floor

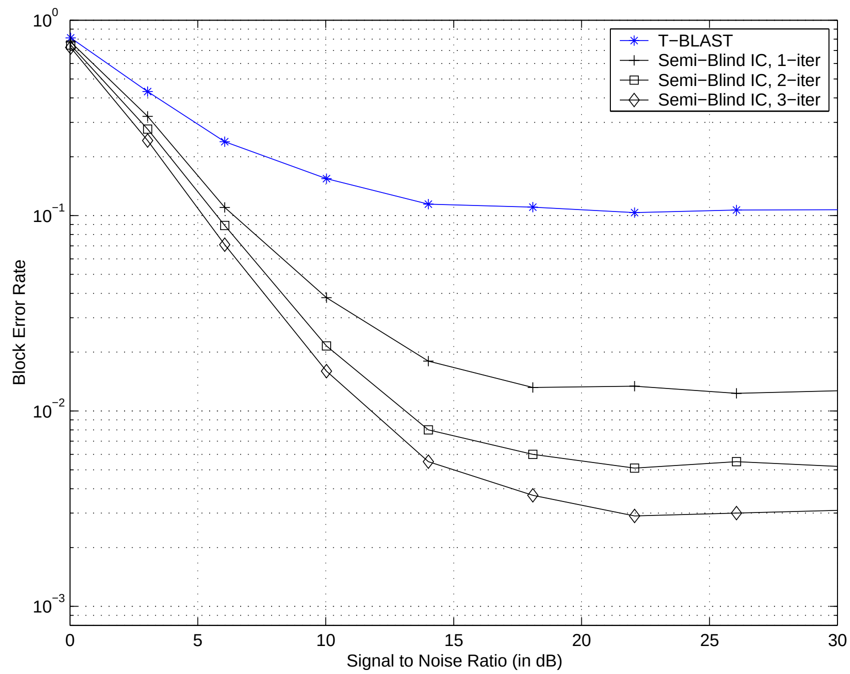

Figure 3.11: Results for an interference-limited multicell system having interferers with multiple antennas, where the strongest interferer is 9db stronger than other interferers. SIR = 0db

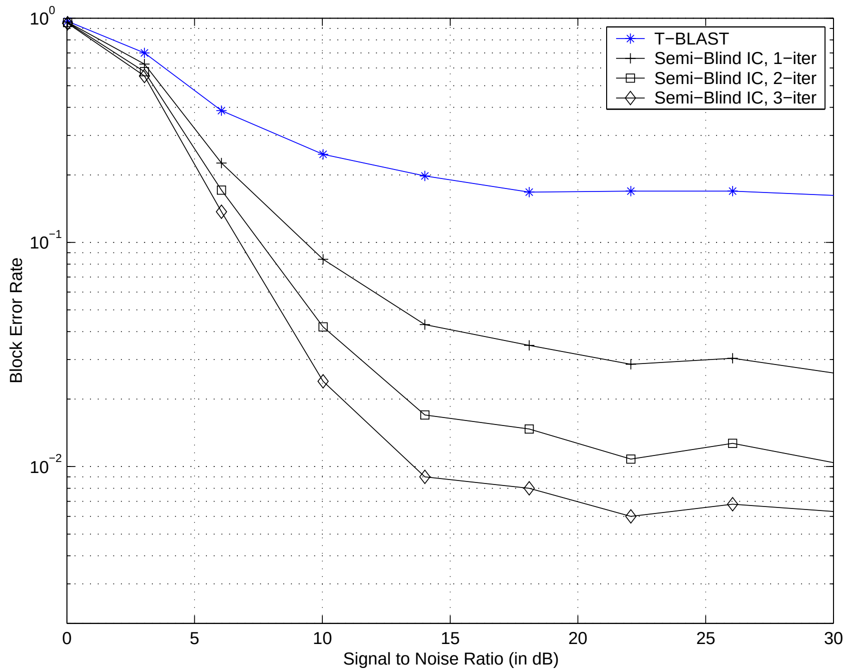

Figure 3.12: Results for an interference-limited multicell system having 4 adjacent cell interferers with single antenna, where the strongest interferer is 3db stronger than the weakest interferer. SIR =2db

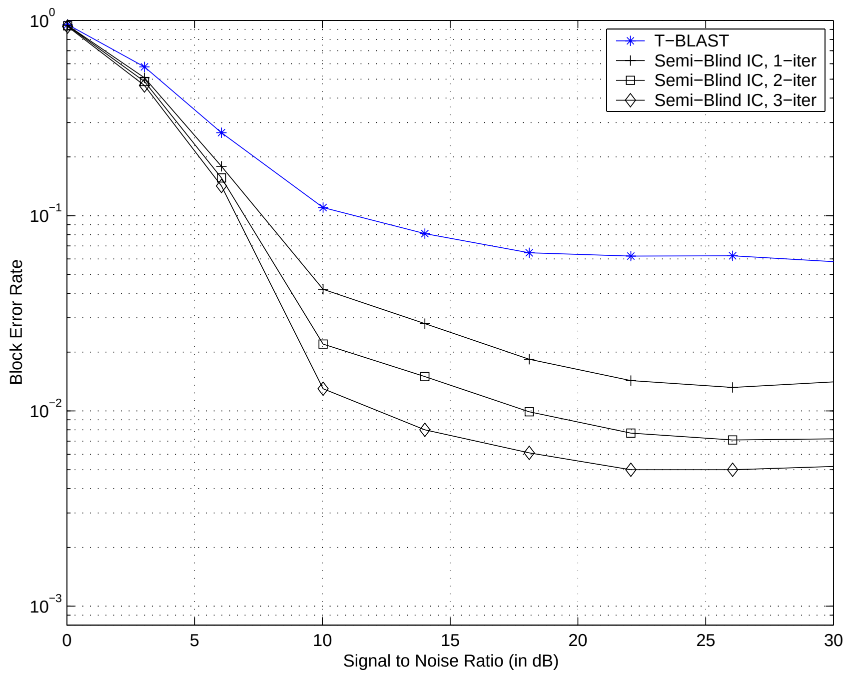

Figure 3.13: Results for an interference-limited multicell system having 6 adjacent cell interferers with single antenna. SIR=3db

attained by the original T-BLAST MUD.

3.7 Summary

In this chapter we have proposed a semi-blind MUD that can cancel out the detrimental effect of the strongest adjacent cell interferer(s), thus improving the estimate for desired user's signals. This semi-blind MUD gives superior performance as compared to previously known MUDs, such as T-BLAST, that treat inter-cell interference as additional Gaussian noise, which is clearly a sub-optimal approach in the presence of strong inter-cell interference. The main advantage of the semi-blind MUD proposed here over other MUDs in (Dai et al. mar. 2004), is its ability to circumvent the channel estimation using training sequences required by these MUDs. This not only makes semi-blind MUD more preferable for use on fast changing channels but also eases the requirement of synchronized transmission from users in different cells. The performance of the semi-blind MUD is limited by the degree of freedom available at the receive antenna array.

Chapter 4

Conclusions and Future work

4.1 Conclusions

For MIMO systems, the receiver structures using training sequences to estimate the channel state information and the transmitted signals are probably the best choice under normal conditions. However, this training based approach for channel estimation is not feasible for bandwidth-limited or rapidly time varying channels and difficult to implement in multicell systems. In this work we develop blind/semi-blind receivers for such situations. Among the algorithms used in blind receivers, the one based on multiuser kurtosis (MUK) maximization particularly interests us due to its globally convergent behavior and ease of implementation. We propose a semi-blind receiver structure (turbo-MUK) for coded MIMO systems, which greatly improves the performance of the MUK receiver by incorporating parallel interference cancellation and soft information exchange between the detector and the decoder stages, with the aid of very few training symbols. Numerical results demonstrate that the newly proposed receiver works well for the case where all the transmitted streams come from one user (single-user MIMO) as well as the case where transmitted streams are from different users with unequal powers (multi-user MIMO).

The existing space-time MUD for single-cell MIMO such as coded V-BLAST and T-BLAST suffer performance degradation in multi-cell environments because they treat inter-cell interference as additional noise. We propose a semi-blind MUD that cancels out the strongest adjacent cell interferer(s) from the received signal, thus giving significant performance improvement over T-BLAST. Numerical results show that this MUD is very effective in an interference-limited environment. These results prove that in presence of strong interference, detecting interfering users is better than treating them as ambient noise. The semi-blind nature of MUD not only conserves bandwidth but also makes it better suited for cellular environment because of its ability to avoid channel estimation using training sequences. The applicability and performance of the semi-blind MUD in various wireless scenarios depend on the degrees of freedom available at the receiver, where larger degrees of freedom translate into an ability to detect more interferers and better performance. From the complexity point of view, the semi-blind MUD proposed here is applicable to uplink processing at base stations currently and will become more relevant for downlink processing with increase in computing power at mobile stations in future.

4.2 Future Work

There are several possible directions in which the ideas presented in this thesis may be extended.

- The possibility of using non-symmetric symbol space constellations or rotationally invariant convolutional codes, instead of resorting to training sequences, to remove the phase ambiguity can be explored. This can make the new MUD structures proposed in this thesis completely blind, hence more suitable for use in cellular environment or use over rapidly varying channels. Since, the performance of the semi-blind MUD depends on accuracy of phase rotation estimate (or length of training sequence), a MUD using these phase ambiguity removal techniques may offer a better performance as compared to semi-blind MUD.

- The performance of the semi-blind MUD in a multicell MIMO system can be studied for generic cellular environments, where shadow fading loss and LOS components are further considered. This will give us a better idea about the utility of the proposed semi-blind/blind MUD.

- The MUK algorithm used in the detection stage of turbo-MUK can be replaced by constant modulus type variants of the MUK algorithm (Papadias sept 2000*a*), which have a faster convergence speed, thereby making turbo-MUK even more suitable for use on fast changing channels.

- In this thesis we have focused on the application of the semi-blind MUD in cellular environments. We believe that this detection structure can be useful in other multicell MIMO systems such as wireless LANs.

Bibliography

Agee, B. G., Schell, S. V. & Gardner, W. A. (apr. 1990). Spectral self coherence restoral: a new approach to blind adaptive signal extraction using antenna arrays. *Proc. IEEE* **78**, 753–767.

Alamouti, S. (oct. 1998). A simple transmitter diversity scheme for wireless communication. *IEEE Journ. on Select. Areas in Commun.* **16**, 1451–1458.

Berrou, C. (aug. 2003). The ten-year-old turbo codes are entering into service. *IEEE Communications Magazine* **41**(8), 110–116.

Berrou, C. & Glavieux, A. (oct. 1996). Near optimum error correcting codes and decoding:turbo-codes. *IEEE Transactions on Communiations* **44**(10), 1261–1271.

C. Douillard, e. a. (oct. 1995). Iterative correction of inter-symbol interference: turbo equalization. *European Trans. Telecomm.* **6**(5), 507 – 511.

Cardoso, J. F. (1989). Source separation using higher order moments. *in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.* **4**, 2109 – 2112.

Cardoso, J. F. (oct. 1998). Blind signal separation: statistical principles. *Proceedings of IEEE* **86**(10), 2009–2025.

Cetreux, S., driessen, P. F. & Greenstein, L. J. (nov. 2000). Simulation results for an interference limited multiple-input multiple-output cellular system. *IEEE Comm. Letters* **4**(11), 334–336.

Comon, P. (july 1996). Contrasts for multichannel blind deconvolution. *IEEE Signal Processing Letters* **3**, 287–314.

Dai, H., Molisch, A. F. & Poor, H. V. (mar. 2004). Downlink capacity of interference limited mimo systems with joint detection. *IEEE Trans. Wireless Comm.* **3**(2), 442–453.

Delfosse, N. & Loubaton, P. (apr. 1994*a*). Adaptive separation of independent sources: a deflation approach. *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.* **4**, 41–44.

Delfosse, N. & Loubaton, P. (april 1994*b*). Adaptive separation of independent sources: A deflation approach. *in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.* **4**, 41–44.

Donoho, D. (1981). On minimum entropy deconvolution. *in Applied Time Series Analysis II, New York:Academic* pp. 565–609.

Foschini, G. J. (1996). Layered space-time architecture for wireless communications in a fading environment when using multiple antennas. *Bell Laboratories Technical Journal* **1**(2), 41–59.

Foschini, G. J. & Gans, M. (mar. 1998). On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Pers. Commun.* **6**, 311–335.

Gardner, W. A. (june 1991). A new method of channel identification. *IEEE Trans. Commun.* **39**(6), 813 – 817.

Gesbert, D., Shafi, M., Shiu, D. S., Smith, P. & Naguib, A. (april 2003). From theory to practice: an overview of mimo space-time coded wireless systems. *IEEE Journal on Selected Areas in Communication* **21**(3), 281–295.

Giannakis, G. B. (nov. 1994). Linear cyclic correlation approach for blind identification of fir channels. *Proc. 28th Asilomar Conference on Signals, Systems and Computers* pp. 420 – 423.

Godard, D. N. (nov. 1980). Self-recovering equalization and carrier tracking in two-dimensional data communication systems. *IEEE Transactions on Communications* **28**, 1867 –1875.

Goel, V. & Dai, H. (2005). Turbo-blast with semi-blind interference cancellation in multicell mimo systems. *Proc. 2005 Fall IEEE Conference on Vehicular Technology(VTC).*

Guo, D. & Wang, X. (april 2003). Blind detection in mimo systems via sequential monte carlo. *IEEE Journal on Selected Areas in Commun.* **21**(3), 464–473.

Hassibi, B. & Hochwald, B. (april 2003). How much training is needed in multiple antenna wireless links?. *IEEE Trans. Inform. Theory* **49**, 951–963.

Hatzinakos, D. & Nikias, C. (jan. 1989). Estimation of multipath channel response in frequency selective channels. *IEEE Jour. Selected Areas in Commun.* **S7**(1), 12 – 19.

Haykin, S. (2000). *Unsupervised adaptive filtering Vol.I Vol. II.* John Wiley and Sons. New York, N.Y.

Komninikas, C. & Wesel, R. D. (sept.2001). Joint iterative channel estimation and decoding in flat correlated rayleigh fading. *IEEE Journal on Selected Areas in Communication* **19**(9), 1706–1716.

L. R. Behl, e. a. (mar. 1974). Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Trans. Inform. Theory* **20**, 284–287.

Liu, H. & Xu, G. (nov. 1995). Closed-form blind symbol estimation in digital communiation. *IEEE TRans. Signal Proc.* **43**(11), 2714 – 2723.

Liu, H., Xu, G. & Tong, L. (nov. 1993). A deterministic approach to blind equalization. *Proc. 27th Asilomar Conf. on Acoust., Speech and Signal Processing* pp. 751 – 755.

Liu, R. & H. Luo, e. a. (may 2002). Autocorrelation - a new differentiating domain for multiple access wireless communications. *IEEE ISCAS 2002* **3**, 643–646.

Moulines, E., Dohamel, P., Cardoso, J. & Mayrargue, S. (feb. 1995). Subspace methods for identification of multi-channel fir filters. *IEEE TRans. Signal Proc.* **43**(2), 516 – 525.

Nordstrom, F., Holst, J. & Lindoff, B. (2002). Interferer cancellation in tdma systems. *5th Nordic signal processing symposium, NORSIG-2002*.

Otterson, B., Roy, R. & Kailath, T. (1989). Signal waveform estimation in sensor array processing. *In Proc. 23rd Asilomar Conf. on signals, Systems and Computers* pp. 787–791.

P. Robertson, e. a. (june 1995). A comparison of optimal and sub-optimal map decoding algorithms operating in the log-domain. *Proc. IEEE International Conf. on Communcations* pp. 1009–1013.

Papadias, C. (feb. 2004). Unsupervised receiver processing techniques for linear space-time equalization of wideband multiple input/ multiple output channels. *IEEE Trans. on Signal Processing* **52**(2), 472 – 482.

Papadias, C. B. (dec. 2000*b*). Globally convergent blind source separation based on multiuser kurtosis maximization criterion. *IEEE TRans. Signal Processing* **48**(12), 3509 – 3518.

Papadias, C. B. (sept 2000*a*). Globally convergent algorithms for blind source separation. *Proc X Eur. Signal Process. Conf., Finland*.

Papadias, C. B. & Paulraj, A. (june 1997). A constant modulus algorithm for multiuser separation in presence of delay spread using antenna arrays. *IEEE signal processing Lett.* **4**(6), 178–181.

Paulraj, A., Roy, R. & Kailath, T. (1985). Estimation of signal parameters via rotational invariance techniques :espirit. *In Proc. 19th Asilomar conf. on Circuits, systems and Comp.*

Poor, H. V. (sept. 2001). Turbo multiuser detection: a primer. *Journal of Communcations and Networks* **3**(3), 196–201.

Raleigh, G. G. & Cioffi, J. M. (mar. 1998). Spatio-temporal coding for wireless communication. *Wireless Pers. Commun.* **46**(3), 357–366.

Ranta, P. A., Hottinen, A. & Honkasalo, Z. C. (1995). Co-channel interference cancelling receiver for tdma mobile systems. *Proc. of IEEE Int. Conf. on Commun. (ICC)* pp. 17–21.

Sato, Y. (june 1975). Two dimensional applications of zero-forcing equalization method. *IEEE Trans. Commun.* **23**, 684 – 687.

Schell, S. V. & Gardner, W. A. (1993). Maximum likelihood and common factor analysis-based blind adaptive spatial filtering for cyclostationary signals. *in Proc. ICASSP* **4**, 229–295.

Schmidt, R. O. (1981). A signal subspace approach to multiple emitter location and spectral estimation. *Ph.D dissertation, Stanford Univ., Stanford, CA*.

Sellathurai, M. & Haykin, S. (oct. 2002). Turbo-blast for wireless communications: theory and experiments. *IEEE Trans. on Signal Processing* **50**(10), 2538–2546.

Shalvi, O. & Weinstein, E. (mar. 1990). New criteria for blind deconvolution of nonminimum phase systems. *IEEE trans. Inform. Theory* **36**, 312 – 321.

Shamsunder, S. & Giannakis, G. (1994). Multichannel blind signal separation and reconstruction. *6th IEEE Digital Signal Processing Workshop*.

Slock, D. T. M. (nov. 1994). Blind fractionally-spaced equalization, perfect reconstruction filter banks and multi-channel linear prediction. *Proc. IEEE Internat. Conf. on Acoust., Speech and Signal Procesing* pp. IV585 – IV588.

Talwar, S., Viberg, M. & Paulraj, A. (feb. 1994). Blind estimation of multiple cochannel digital signals using an antenna array. *IEEE Signal Process. Letters* **1**(2), 29 –31.

Tarokh, V., Seshadri, N. & Calderbank, A. R. (mar. 1998). Space-time codes for high data rate wireless commmunications: performance criterion and code construction. *IEEE Trans. on Inform. Theory* **44**, 744–765.

Telatar, E. (june 1995). Capacity of multiantenna gaussian channels. *AT and T Bell Labs, Tech. Memo.*

Tong, L., Xu, G. & Kailath, T. (nov. 1991). A new approach to blind identification and equalization of multipath channel. *Proc. 25th Asilomar Conference on Signals, Systems and Computers.*

Tong, L., Xu, G., Hassibi, B. & Kailath, T. (jan. 1995). Blind identification and equalization based on second-order statistics, a frequency domain approach. *IEEE Trans. Infor. Theory* **41**(1), 329 – 333.

Triechler, J. R. & Agee, B. G. (apr. 1983). A new approach to multipath correction of constant modulus signals. *IEEE Trans. Acoust., Speech, Signal Processing* **31**, 349 – 372.

Tugnait, J. (jan. 1997). Blind spatio-temporal equalization and spectral response estimation for mimo channel using a godard cost function. *IEEE Trans. Signal Processing* **45**, 268–271.

Tugnait, J. K. (may 1987). Identification of linear stochastic systems via second and fourth order cumulant matching. *IEEE Trans. Inform. Theory* **33**, 393 – 407.

Vanderveen, A. & Paulraj, A. (1996). An analytical constant modulus algorithm. *IEEE Trans. Signal Processing* **44**, 1136 – 1155.

Vanderveen, A. J., Talwar, S. & Paulraj, A. (may 1995). Blind estimation of fir channels carrying multiple finite alphabet signals. *In Proc. IEEE ICASSP* **2**, 1213–1216.

Varanasi, M. (july 1995). Group detection for synchronous gaussian code-division multiple-access channels. *IEEE Trans. Inform. Theory* **41**, 1083–1096.

Wang, X. & Poor, H. V. (july 1999). Iterative (turbo) soft interference cancellation and decoding for coded cdma. *IEEE Transactions on Communcations* **47**(7), 1046–1061.

Wolniansky, P. W., Foschini, G. J., Golden, G. D. & Valenzuela, R. A. (oct. 1998). V-blast: an architecture for realizing very high data rates over the rich-scattering wireless channel. *URSI International Symposium on Signals, Systems and Electronics* pp. 295–300.

Wyner, A. D. (jan. 1974). Recent results in the shannon theory. *IEEE Trans. Inform. Theory* **IT-20**, 2–10.

Xu, G., Liu, H., Tong, L. & Kailath, T. (april 1994). A least squares approach to blind channel identification. *Proc. International Conf. Acoust., Speech and Signal Processing*.

Zheng, L. & Tse, D. N. C. (may 2003). Diversity and multiplexing: A fundamental tradeoff in multiple- antenna channels. *IEEE Trans. on Information Theory* **49**(5), 1073 – 1096.