Coding-Assisted Blind MIMO Separation and Decoding

Citation for published version:
Zhao, X & Davies, M 2010, 'Coding-Assisted Blind MIMO Separation and Decoding', *IEEE Transactions on Vehicular Technology*, vol. 59, no. 9, pp. 4408-4417.

Coding Assisted Blind MIMO Separation and Decoding

Xu Zhao, Mike Davies, *Member, IEEE*

Abstract—**Despite the widespread use of forward-error correcting coding (FEC), most multiple input multiple output (MIMO) blind channel estimation techniques ignore its presence, and instead make the simplifying assumption that the transmitted symbols are uncoded. However, FEC induces code structure in the transmitted sequence that can be exploited to improve blind MIMO channel estimates. In this paper, we exploit the iterative channel estimation based on a posteriori information for blind MIMO separation and decoding. Experiments show the improvements achievable by exploiting the existence of coding structures and that it can approach the performance of a BCJR equalizer with perfect channel information in a reasonable SNR range. Also, through splitting the FEC codeword over multiple blocks, the impact in performance of a bad-conditioned channel matrix can be kept at a reasonable level.**

Index Terms—**Blind Separation, blind channel estimation, independent component analysis, turbo equalization, expectation maximization (EM) channel estimation.**

I. Introduction

In wireless MIMO systems, all practical receivers are designed based on the requirements of acquiring channel state information (CSI) and the channel needs to be estimated in advance before decoding operations. However, obtaining an accurate estimate can be problematic in some environments. For example, if the channel response varies rapidly with time, if the channel is very singular or the signal to noise ratio (SNR) is low. Moreover, with the ever-growing demands of increasing data rate and the requirements of saving the limited bandwidth, several blind or semi-blind systems have been studied in the last decade. The typical subspace method described in [1][2] utilizes the orthogonality between the channel matrix and the noise subspace in order to compensate for extra degrees of freedom provided by the noise subspace. The main drawback of subspace-based MIMO channel estimation is that it needs the number of received antennas to be larger than the number of transmit antennas, otherwise, it requires precoding preprocessing. Other schemes [3] using singular value decomposition (SVD) employ a simple block pre-coding structure. The advantage is that the CSI can be recovered without ambiguity when applying a proper modulation. Nevertheless, this advantage is obtained at the cost of decreasing the spatial diversity.

X. Zhao and M. Davies are with the Institute for Digital Communications Joint Research Institute for Signal and Image Processing School of Engineering and Electronics, The University of Edinburgh. Kings Buildings, Mayfield Road, Edinburgh EH9 3JL, UK (Tel.: +44(0)131 6505659, Fax.: +44(0)131 6506554, e-mail: x.zhao@ed.ac.uk, mike.davies@ed.ac.uk).

Manuscript received September, 2009; revised April, 2010, accepted July, 2010. The associate editor coordinating the review of this paper and approving it for publication was Dr. Kambiz Zangi.

The most popular equalizers use Godard's method [4] or the constant modulus algorithm (CMA) [5]. While in the MIMO setting these do not require more receive antennas than transmit ones, such methods essentially estimate a linear equalization operator and encounter difficulties if the channel matrix is not well conditioned. In this case the maximum likelihood (ML) channel estimate receiver is generally much more robust.

On the other hand, FEC coding, which restricts the transmitted sequence to a limited coding space so as to increase the minimum distance, can correct the potentially wrong decoding due to noise contamination. Using the FEC code, the decoder at the receiver can feedback a posteriori information to the equalizer. The practical challenge, nonetheless, is the tremendous complexity demanded by this joint optimal ML decoding. To solve this problem, the iterative soft decoder has been studied and has been found to approach the optimal ML decoding performance at a practical and reasonable computational burden [6]. Furthermore, with powerful digital processors in the last decade, contributions of FEC to decoding with affordable complexity were explored in [7][8]. Looking from a broader angle, we can take blind equalization as part of the decoding process. Thus, we can try to find a blind equalization and channel correcting scheme that together approximate the Shannon bound. Such methods combine the blind iterative channel estimation and turbo equalization. As illustrated in left block diagram of Figure 1 in the next page, the equalizer uses the channel estimates to compute soft information of the transmitted symbols. The channel estimator then applies these soft symbols to improve the channel estimates, which in turn yields better symbol estimates, and so forth. In contrast, the FEC aware channel estimator based on soft symbol, *a priori* information, feeds this information into the decoder in order to get more reliable soft bit information. Next, this posterior information is fed back to the channel estimator, and so on, as illustrated in right block diagram of Figure 1. Such a scheme utilizes FEC information in blind equalization. Nevertheless, there has been little work relating FEC to MIMO channel estimation. Although the independent, identically distributed (i.i.d.) assumption usually made in MIMO blind separation [9][10] no longer holds (due to the FEC coding), it has been shown that FEC does not impair the performance of some blind equalization techniques [11].

Some previous research has explored the FEC property on blind channel estimation. For example, [8] combined blind channel identification and turbo equalization. They used the Expectation Maximization (EM) algorithm to update channel state information. The covariance matrix is computed as a sample average in which the likelihood of the received sym-

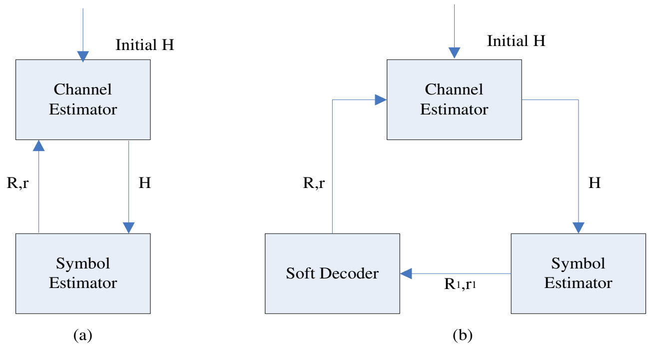

Fig. 1: Joint channel estimation and symbol detection: Uncoded diagram vs. FEC diagram

bols weights the data. In contrast, in this work, pairwise joint probabilities are used to measure expectations in the E-step.

In another work [12] exploited the turbo equalization of unknown ISI channels using a trellis to represent the channel. This develops a hidden Markov model (HMM) [13] for the noisy channel output, and the Baum Welch algorithm, a specific instance of the EM algorithm, is applied to estimate the HMM parameters including the observations before adding noise. This approach need not estimate an explicit channel directly but can be calculated using the estimated symbols. In this paper, the soft-output BCJR equalizer depends on a channel estimate that is obtained from the EM algorithm iteratively.

In [14] [15] a blind iterative channel estimator is used that is also based on an EM algorithm. They applied a turbo equalizer loop with a decision feedback equalizer. Such schemes enjoy a low complexity. In the development of the EM channel estimator, they used a sample average to replace the ensemble average. In our work, the marginal and joint probabilities of each symbol are used to evaluate ensemble averages and this computation generates more accurate information.

In [16] [17], Gunther and co-workers presented a generalized BCJR and LDPC algorithm to compute joint posteriori probabilities of symbols given noisy observations to suppress intersymbol interference (ISI) at the output of channels. These pair-wise joint posterior probabilities are applied in the EM channel estimator. Both schemes are suitable for a single channel with a small number of channel taps since, in such a case, the surface of likelihood function is simple and smooth enough to allow the EM algorithm to converge to a desirable point, otherwise, the generalized BCJR may not converge to the correct state. In [17], for a single channel, 30 EM iterations were used to evaluate the system performance. This number of operations introduces a considerable computational burden.

Per-Survivor Processing (PSP) [18] is a seminal work in joint channel estimation and symbol detection. The PSP method embeds the data-aided channel estimation into the Viterbi algorithm. Each state has a separate channel estimate which is based on the survivor path leading to that state. PSP performs an ML estimation of the channel parameters. Then it estimated the candidate channel by applying a LMS algorithm or a Kalman filter to each survivor path. This kind of method was subsequently developed as an approximate minimum variance channel estimator. A similar but simplified approach was proposed by [19], where the author maximizes the data log likelihood by weighting a quadratic function associated with each survivor path by the path probability. The authors then used an EM algorithm to update the channel estimate iteratively. To avoid the initialization problem, they enumerated many initializations for the EM and select the most probable one in terms of the likelihood. Without an efficient proposal for initialization of the EM, this scheme is not feasible for multi-dimensional dimensional systems with large constellations.

The methods described above were employed in single input single output channels and small constellations and these techniques do not simply extend to the multi-dimension (MIMO) and large constellation QAM systems. In these cases, the increased dimensionality of the MIMO channel can make the convergence problematic.

To tackle the MIMO, large constellation scenario, an efficient hybrid system for blind equalization was proposed in [20], in which the sphere decoder (SD) algorithm is integrated into the EM algorithm for the large multi-dimensional channel estimate. The initialization of the EM is provided by a fast and simple nonlinear independent component analysis (ICA) method which is specifically designed for QAM modulation. Such an efficient combination makes this feasible for a real time communication system. The numerical simulations demonstrate the effectiveness of this combined technique. However, it still suffered a loss of performance when the channel was close to singular or the noise level was high.

Other work on semi-blind channel estimation appears in [21], where the authors used a pilot sequence and a Wiener filter to initialize and update the channel respectively. This Minimum Mean Square Error (MMSE) based iterative channel estimator uses soft information from the output of the decoder to improve the mean square error of the channel estimates. However, taking the mean values of the data symbols calculated by the posteriori probabilities, is not an accurate way to improve the channel estimates.

A similar idea using the a posteriori probabilities computed by a soft iterative decoder to improve the ML channel estimator computed via the EM channel estimation is presented in [22][23]. In [22] the authors use of List Sphere Detection for reduced computational complexity. Both methods require training symbols to gain a sufficiently good initial channel estimate to allow the subsequent application of the EM algorithm to correctly converge. Both methods also used the EM algorithm with a strong code, such as turbo codes and LDPC codes in the feedback loop. This further adds to the computational burden. In contrast to this, the scheme we propose here works without using a strong code.Consequently, much computation can be avoided.

In this paper, a blind MIMO receiver that combines soft channel estimation and a soft MIMO equalizer and decoder is proposed. In this hybrid design, we improve the receiver's performance through efficiently incorporating the soft bit information from the decoder into the EM channel estimator. This system uses an efficient independent component analysis method suitable for QAM modulation to gain good initial

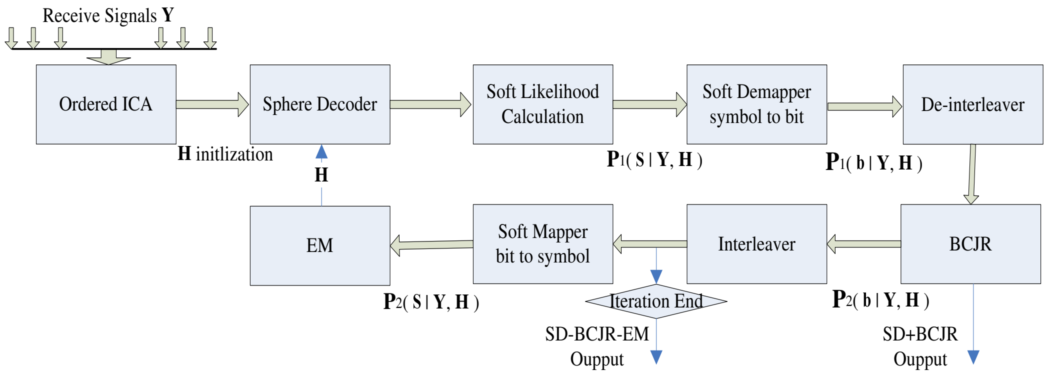

Fig. 2: Receiver architecture of the proposed coding assisted system.

estimates for the EM algorithm, a selective sphere decoder process that computes the likelihood values (soft information) and a simple error correcting operation. This scheme has low complexity and improved convergence, being more likely to converge to the desirable stationary point. Moreover, by sending the bit interleaved coded modulation (BICM) bits on differently fading channels, we can further make use of the temporal diversity. Combining this with the spatial diversity due to the statistical independence of transmitted sequences, blind estimate and adjustment of the channel matrix can be performed simultaneously. Such splitting the FEC codeword over multiple blocks is shown to help avoid singular channels. Empirically this provides us with outstanding performance of MIMO blind equalization and decoding at a reasonable computational cost. This idea can also be easily extended to an MIMO-OFDM system, especially for the fast fading channel acquisition and tracking.

We consider the $M \times M$ MIMO narrow band system model,

$$Y = HS + N, \tag{1}$$

where $Y \in \mathrm{C}^{M\times T}$ is the matrix containing observed signals from the receiver antennas, and $S \in \mathrm{C}^{M\times T}$ is the complex discrete source signal matrix. $N \in \mathrm{C}^{M\times T}$ is the noise matrix with covariance, Σ, which is assumed to be uncorrelated with the source signals and T is the sampled points of observations. $H \in \mathrm{C}^{M\times M}$, the Rayleigh channel, is the unknown linear square channel matrix whose elements are assumed to be drawn independently from a complex Gaussian distribution and we assume that it is invertible. Note that, H is instantaneous but we do not guarantee it is orthogonal. This square channel matrix can be expanded into the non-square overdetermined MIMO systems where the number of received antennas is greater than the number of transmitted antennas via principal components analysis (PCA) [24] or singular value decomposition (SVD) [25] techniques. For the transmission of a frame of K_b bits, the transmitter encodes the K_u information bits using a convolutional code of rate r, where $K_u = K_b \times r$. The coded bits are interleaved and mapped into QAM symbols, forming a sequence of $K_s = K_b/\log_2 P$ symbols, where P is the number of possible symbols in the QAM constellation. Then the QAM sequence of symbols is split into M substreams corresponding to one Rayleigh fading channel, and is transmitted in parallel from each one of the M antennas. The problem above arises not only in MIMO systems, but also in multiuser DS/CDMA systems [26]. It further reduces to SIMO blind equalization when there is only one source signal or when fractionally spaced equalization is employed in single antenna communication systems.

The paper is organized as follows. Section II illustrates the architecture of the proposed blind MIMO separation and decoding scheme and each part is introduced in detail. Section III discusses the multiple blocks FEC method for the singular channel matrix scenario. Simulations are shown in Section IV and conclusions are given in Section V.

II. The Proposed Coding Assisted Blind MIMO Separation and Decoding

In this work, a blind MIMO channel equalization algorithm is designed in which the BCJR and EM algorithms are iterated. Figure 2 shows the receiver structure using iterative equalization, whereby a soft equalizer interacts with a soft-input-soft-output error control decoder. Given initial estimates H^{ini} from the efficient nonlinear ICA method [20], the SD-BCJR algorithm computes the signal a posteriori probabilities $p(\mathbf{s}_k|Y, H^i)$ by utilizing the code structure and then feeds these to the EM algorithm. The EM algorithm uses these a posteriori probabilities to evaluate the conditional expectations in (7) and (8) as we will introduce next. Thus we update the new channel state information by (9) and pass this back to the SD-BCJR algorithm again. As the iterations proceed, estimates become more accurate and the a posteriori symbol probabilities become more precise.

A. Soft MIMO Equalizer and The BCJR Decoder

The optimal ML receiver has exponential complexity with the signal modulation size and number of transmit antennas, thus limiting its real time application. The sphere decoder, on the other hand, is capable of achieving near ML performance [27] and can be designed to provide the soft (likelihood) output information [28]. Thus, we propose a blind soft equalizer and decoder architecture combining the SD decoder and a simple error correcting operator with low complexity. The low-complexity may enable iterative equalization for fast wireless

Rayleigh channels. An important requirement for the proposed blind MIMO equalization is the calculation of soft information both for the channel estimation and the soft MIMO detector and decoder.

Channel coding has been extensively researched in the literature. Here we introduce the techniques used in this paper. The well known Bahl−Cocke−Jelinek−Raviv (BCJR) algorithm [29] is used to compute a posteriori probabilities (APP) of inputs to a finite state machine for the received signals. It provides the exact a posteriori probabilities for convolutional codes and has also been applied to other error correcting codes such as turbo code [30]. While a full complexity BCJR soft decoder is used in this work, an efficient sliding window type scheme [31] can be applied in practical applications, which leads to suboptimal performance with a much lower complexity.

For a BPSK modulation scheme, the log-likelihood ratio (LLR) takes the form:

$$L(b_k) = \ln \frac{p(b_k = 1|Y, \tilde{H})}{p(b_k = 0|Y, \tilde{H})}, \quad (2)$$

where $\tilde{H}$ is the channel estimate. This LLR value shows the reliability of the information bit. Given a convolutional code at the transmitter, the BCJR algorithm calculates the APP exactly if we know the true channel state information.

B. Soft Mapper and Demapper

The BCJR algorithm is designed for a convolutional bit sequence. In a large constellation QAM modulation, soft symbol information needs to be transferred to bit information for the following BCJR operations. In this section, a description of the QAM soft mapper and demapper employed in our proposed decoding scheme is given. Thanks to the bit interleaver in both the transmitters and receivers, the marginal posteriori probabilities of the coded symbols can be expressed as the product of the bit a posteriori probabilities.

The soft demapper and mapper take the following three steps to compute the output symbol-wise APP to be fed back to the EM channel estimator.

1) Demapping with a priori probabilities.
   Define a coded symbol s, with m bits, as $s = \{b_0, \ldots, b_m\}$. The demapper extracts a soft value of each coded bit for subsequent decoding. The following gives a description of this demapper. For a number of m coded bits, the L-value of bit b_j is given as

$$L(b_j) = L_a(b_j) \ln \frac{\sum_{s_k \in S_1^m} p(\mathbf{y}|s_k) e^{L(s_k)}}{\sum_{s_k \in S_0^m} p(\mathbf{y}|s_k) e^{L(s_k)}}, \quad (3)$$

   where S_1^m and S_0^m define the subsets of S in which the bit b_m takes the values 1 and 0, respectively, s_t is the mapping symbol with the bit b_m taking the values 1 and 0, and $L(s_k)$ is the likelihood of symbol s_k at time index k. The probability of each mapping symbol s_t is computed from the equalizer output. The soft demapper given by equation (3) calculates the marginal probabilities and ignores bit dependencies within the codeword.
2) The BCJR algorithm (or other FEC algorithms). This calculates bit-wise a posteriori probabilities based on the marginal probabilities obtained in step 1.
3) Mapping bit APPs to symbol APPs. For each symbol s_k, the channel joint symbol-wise posterior $p(s)$ can be approximated by the product of the input marginal bit-wise posterior $p(b_j)$. It is given as:

$$p(s_k) = \prod_{j=1}^{m} p(b_j), \quad (4)$$

   where

$$p(b_j) = \frac{e^{L(b_j)}}{1 + e^{L(b_j)}}, \quad (5)$$

   denotes the input bit-wise priors offered by the channel decoder in step 2. Generally, LLR clipping techniques [32][33] can be applied to reduce complexity. Here, we fix the LLR clipping level $L_{clip} = 3$ as used in [34].

Note that, a Gray mapping was employed in this work and the optimized symbol mappings for BICM with interleaved decoding were researched in high order constellations, the interested reader is refered to [35]. The soft mapper used in step 3 may lose some information since the joint symbol probability is set equal to the product of marginal bit probability. This assumption is only exact if the bit stream probabilities are strictly independent.

C. EM Channel Estimation with a Posteriori Probability

Most prior work in blind iterative channel identification can be tied to the EM algorithm [36]. It is a general methodology for maximum likelihood or maximum a posteriori estimation. The first use of EM with soft symbol estimates was proposed in [37]. An adaptive version of EM was applied in the identification problem in [38] and some modified EM algorithms were proposed in [39][40]. The EM algorithm updates are analytically simple and numerically stable for distributions that belong to the exponential family. Here we explore EM channel estimation that exploits a posteriori information.

Considering the system model, equation (1), the EM algorithm estimates the channel H based on received signals $Y = \{\mathbf{y}_k\}_1^T$. It maximizes the log likelihood, $\log P(Y|H)$ with an initial channel H^{ini}, by iteratively calculating,

$$H^{i+1} = \underset{H}{\operatorname{argmin}} \, \mathrm{E}\{-\log P(Y|H^i, S^i)|P(S^i|Y, H^i)\}, \quad (6)$$

where H^i is the ith estimate of the channel and S^i is the ith estimate of the symbols. As we know, the EM iteration in (6) only guarantees convergence to a local maximum of $P(Y|H)$ [41].

The update of the equation (6) can be written in a closed-form solution [42] as follows,

$$r^i = \sum_{k=1}^{T} \mathbf{y}_k \mathrm{E}\{\mathbf{s}_k^i|Y, H^k\} \quad (7)$$

$$R^i = \sum_{k=1}^{T} \mathrm{E}\{\mathbf{s}_k^i (\mathbf{s}_k^i)^*|Y, H^i\} \quad (8)$$

$$H^{i+1} = (R^i)^{-1} r^i \tag{9}$$

Equations (7) and (8) depend on first-order statistics and the second-order statistics of the symbols respectively. Note that the computation of (7) and (8) also require the a posteriori probabilities $P(\mathbf{s}_k|Y, H)$ and $P(\mathbf{s}_k\mathbf{s}_k^*|Y, H)$, which are approximated in equation (4). We emphasize that the EM algorithm can make full use of the soft a priori information of the coded bits from the BCJR decoder and these posterior probabilities allow us to exploit the coding structure and thereby provide more accurate channel estimates.

An important problem in the performance of the EM algorithm is the appropriate selection of the initial estimate. In the case of low order constellation modulations and small number of received antennas, the EM algorithm may converge to the desirable point after several re-initializations of the iterative procedure [43]. However, if the likelihood surface is complicated, which happens in high order modulation and with a large number of received antennas, the EM is liable to converge to a local minimum rather than global minimum. Such convergent behaviour has been studied by many researchers, see for example [44] and the references therein.

In [45] we have shown that the EM algorithm locally has a Newton-like convergence in digital communication systems. This makes the EM algorithm suitable for real time applications in wireless communications. However, there is still the important but unsolved problem of whether the EM algorithm can converge to the correct solution, i.e., the consistent solution of the true channel parameters. Our iterative joint channel estimation and symbol decoder with coding assist can partly solve such difficulty, as we will see.

D. Iterative Procedure

In MIMO channels, this soft decoding strategy for blind equalization consists of four stages:

1) Blindly estimate the channel state information from the statistics of the received signals. Here, we use an efficient nonlinear ICA approach to get the initial channel state information estimate as in [20].

2) Estimate the soft bits, i.e., the LLR of each transmitted bit, using the list version of the sphere decoder or its variants and the current channel state information estimate.

3) Make the soft bit information more reliable through a simple BCJR soft decoder.

4) Update the channel state information by the EM algorithm with the soft bit information input and iteratively feed it back to re-estimate the soft bits in step 2 again for further improvement.

Note that, this hybrid architecture does not calculate the full symbol probabilities in order to reduce the system complexity. It uses a number of approximations during the iterative procedure, *i.e.* in the list sphere decoder, the soft mapping and de-mapping.

III. Codewords Over Multiple Blocks

In blind MIMO separation, poor channel estimates principally occur when the SNR is low or when the channel is singular [20]. The latter is related to the channel matrix condition number, γ, which is the ratio of the largest singular value over the smallest singular value. It is a measure of how ill-conditioned the matrix is at receiver. When the channel is very singular, a precise blind estimate may be problematic. However, with a well-conditioned channel matrix, blind separation can usually provide good estimates. Here we utilize this reliable information to rescue the information in singular channel. Typically, as we will see, singular channel matrices occur with low probability. Hence the information received from neighboring good channels can be used to correct the bad information from singular channels.

One potential problem is that a long codeword will effect the system complexity and the decoding delay. Thus the next question is how many blocks are needed to form a codeword. Roughly speaking, this is a function of many parameters, such as: the SNR, the performance of blind separations and the frequency of occurrence and the condition number of the singular channels.

The number of poorly conditioned channel matrices observed will depend on the statistics of the propagation medium. Let us assume that we have an $M \times M$ MIMO system with a Rayleigh block-fading channel. That is we assume that for each block the channel elements $H_{i,j}$, are independently and identically distributed as $H_{i,j} \sim \mathcal{CN}(0, 1)$. In this case the probability density of the normalized condition number, $\tilde{\gamma} = \gamma/M$, can be written in closed form [46] as:

$$p_{\tilde{\gamma}}(x) = \frac{8}{x^3}\exp(-\frac{4}{x^2}). \tag{10}$$

With this equation, we can calculate the probability that the condition number will be greater than a specific value, γ_1 as:

$$P\{x > \gamma_1\} = \int_{\gamma_1}^{\infty} \frac{8}{x^3}\exp(-\frac{4}{x^2})dx \tag{11}$$

which by the change of variable $y = 4/x^2$ can be simplified to:

$$\begin{aligned} P\{x > \gamma_1\} &= \int_{4/\gamma_1^2}^{0} -\exp(-y)dy \\ &= 1 - \exp(-\frac{4}{\gamma_1^2}) \end{aligned} \tag{12}$$

This can help us to find a reasonable block length within a codeword. For example, in a 4×4 MIMO system, if the blind separation algorithm can not provide satisfactory performance when the channel condition number $\gamma_1 > 20$ under practical SNRs, we can calculate its probability by

$$P\{x > 20\} = 1 - \exp(-\frac{4}{(20/4)^2}) \approx 0.15 \tag{13}$$

If we assume that the channel matrices for each block are independent, then, the expected frequency of a singular channel is 0.15 and a code length of 7 blocks would typically encounter a single singular channel.

This illustrates that a long codeword spanning different fading factors can exploit the temporal diversity to increase the error correction capability. Such an advance is applied to the blind channel estimation iteratively and, as we will see in the next section, improves the final performance both in the channel estimate and the BER significantly.

IV. Simulations

We consider MIMO systems with 4×4 or 8×8 transmitters and receivers and QAM16 modulation is used throughout. The channel H is therefore an 4×4 or 8×8 complex instantaneous matrix, which is constant for each block interval (256 symbols) and follows a Rayleigh fading distribution. N follows the complex additive white Gaussian distribution. The results have been obtained for transmitting blocks of $K_b = 4096$ bits in a 4×4 system and $K_b = 812$ bits in a 8×8 system. For the error correcting system, a rate of $r = 1/2$ parallel concatenated convolutional code of memory 3 with two nonsystematic convolutional (NSC) code has been used. The generator polynomials are $G_1(D) = 1 + D + D3 + D4$, $G_2(D) = 1 + D3 + D4$ and the interleaver is set to pseudo random.

A SD [47] is employed in the detector to provide the soft symbol information. The SD computes the symbol likelihood based on 16 constellation points in each dimension. This calculation could be simplified by a list SD [32] or list-fixed-complexity SD [48] but with a potential performance degradation.

The SD output is then conveyed to the BCJR and EM algorithms respectively. As the EM algorithm exhibits very fast convergence only two iterations of the EM channel estimation updates are employed in these simulations (with the exception of those in the final subsection).

To explore the error correcting code correctly, the permutation problem of ICA must be overcome. Here a channel re-ordering technique is used. Given initial estimates H_{ICA}, we estimate the channel permutation matrix, P, by making the $H_{ICA}^{-1}H = DP \approx I$ as close to the identity as possible, where D is a diagonal matrix. The channel estimates $H^{ini} = H_{ICA}P$ are then ordered with P. This operation is called ordered ICA in Figure 2. In practice, a small pilot can be inserted into the data block to indicate the correct permutation or in CDMA systems, distinct unique codes (spread code) can be applied to the transmitted data steam in each antenna so that the correct permutation can be identified at the receivers.

A. Channel estimation performance

In comparison with other effective methods of blind MIMO separation, such as the *Split Threshold nonlinear function* and the SD-EM approach [20], our scheme shows promising performance for this type of problem. The former used an efficient score function which is specified for QAM signals to obtain a good separability. The latter proposed an efficient hybrid blind MIMO equalization and decoding scheme using soft information in the EM channel update but ignoring any additional FEC information. Figure 3 and 4 illustrate the separability improvements with the aid of a channel code in the 4×4 and 8×8 MIMO systems. The performance is evaluated in terms of the Inter-Component Interference (ICI) which measures the distance between the estimated channel and the true value. It is defined as:

$$ICI(P) = \frac{1}{n}\sum_i\sum_j\left[\left(\frac{|P_{ij}|}{max|P_{ij}|}\right)^2 - 1\right], \tag{14}$$

where $P = \tilde{H}^{-1}H_{real}$ and H_{real} and $\tilde{H}$ are the true channel and estimated channel respectively.

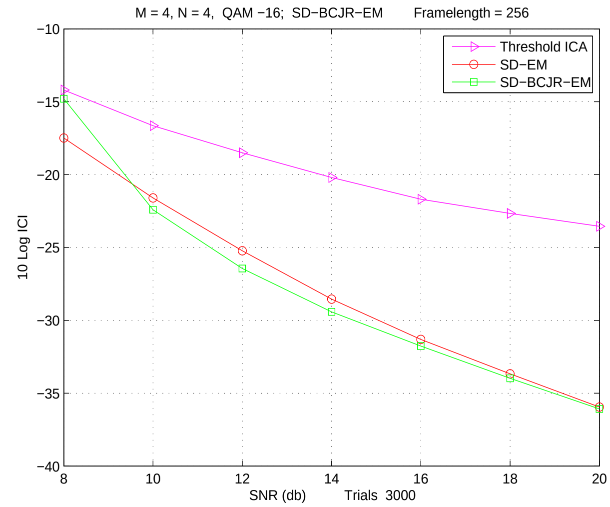

Fig. 3: Channel separability of the split threshold nonlinear ICA, the SD-EM and the coding assisted SD-EM algorithm with a rate $r = 1/2$ convolutional code over different SNR in 4×4 MIMO systems.

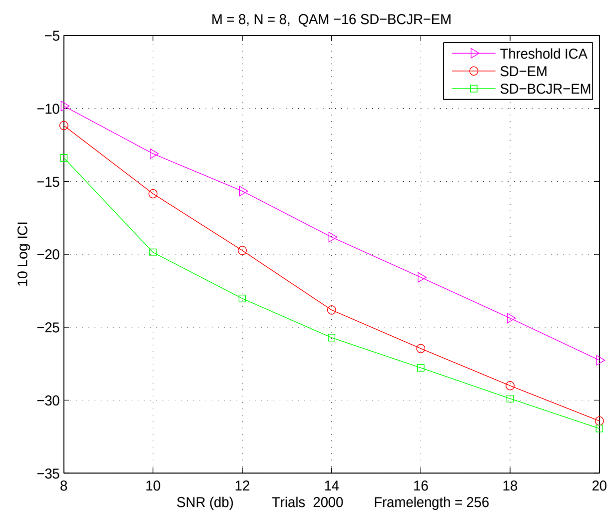

Fig. 4: Channel separability of the split threshold nonlinear ICA, the SD-EM and the coding assisted SD-EM algorithm with a rate $r = 1/2$ convolutional code over different SNR in 8×8 MIMO systems.

B. BER performance

In digital communications, the ultimate goal is to obtain the optimum BER performance. BER performance results are shown in Figures 5 and 6 for the 4×4 and 8×8 MIMO systems respectively. Comparisons are made with: a zero forcing

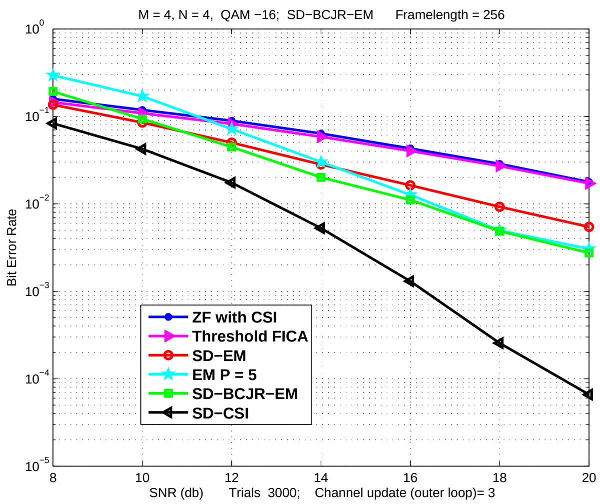

Fig. 5: BER performance of the ZF scheme, the split threshold nonlinear ICA, the SD-EM, the pilot assisted EM method, the coding assisted SD-EM algorithm and known CSI SD with a rate $r = 1/2$ convolutional code in 4×4 MIMO systems.

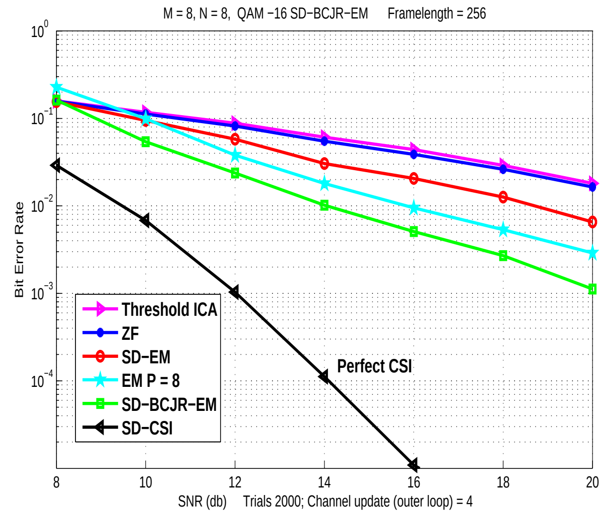

Fig. 6: BER performance of the ZF scheme, the split threshold nonlinear ICA, the SD-EM, the pilot assisted EM method, the coding assisted SD-EM algorithm and known CSI SD with a rate $r = 1/2$ convolutional code in 8×8 MIMO systems.

(ZF) scheme assuming known CSI, the two blind methods mentioned above and training sequence schemes using Least Squares initial estimation [23] with 5 pilot symbols and 8 pilot symbols, respectively.

In order to illustrate the benefits from the channel improvement rather than the FEC, the BER is measured *before* the error correcting operations. The figures show that the coding assisted iterative structure improves the system performance significantly compared with the other blind schemes and it is generally better than using the pilot sequences to provide the initial channel estimates. The gains also appear to be more substantial in the larger MIMO system.

Although, the BER performance is improved, there is still a clear performance gap between the SD-BCJR-EM algorithm and the optimal (with known CSI) SD solution in both figures. We believe such a phenomenon is introduced by a small proportion of very singular channel estimates in which the amplitudes of one or two channels are so weak that the blind estimation can not identify them accurately in noisy environments. We will see below that this problem can be mitigated by allowing the codewords to span over multiple blocks.

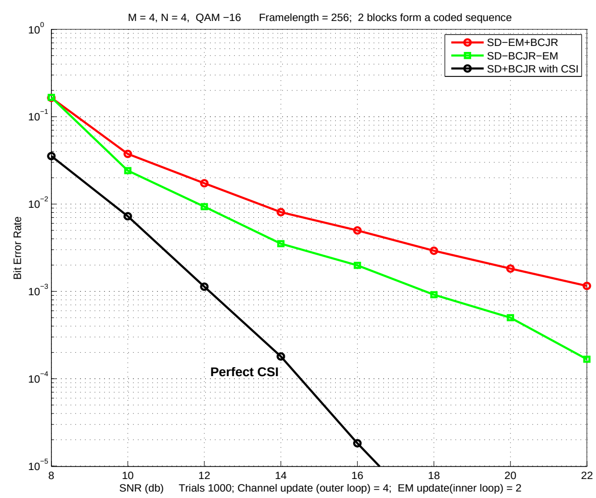

Fig. 7: BER improvements by utilizing time diversity and channel coding. Triangle line is the performance of the SD-EM algorithm following the BCJR algorithm and the square line is the performance of the coding assisted SD-EM algorithm. 2 blocks form a codeword

C. Performance for codewords spanning multiple blocks

For a setup similar to the system above but which provides an increased time diversity, we applied a codeword to 2, 4 and 8 channel realizations in which the channel realizations are independent. The BER performance is measured after the BCJR algorithm. Figures 7, 8 and 10 show that the performance is also improved by the time diversity. This enables, through the interleaving operation, poorly initialized (singular) channel estimates to be re-estimated using higher quality information from the well conditioned channel blocks. In the case of 2 blocks forming a codeword, the gap between our scheme and the optimum solution with perfect CSI is still large. However, the performance improvement is most striking when the code extends over 4 or 8 blocks. When the code spans 4 blocks the BER is within a small factor from the performance with perfect CSI, whereas when the code spans 8 blocks their performances are virtually indistinguishable above 12dB SNR.

D. Convergence over a singular channel

Our final simulations indicates the iterative gain of the coded soft channel estimation measured over $20,000$ Monte

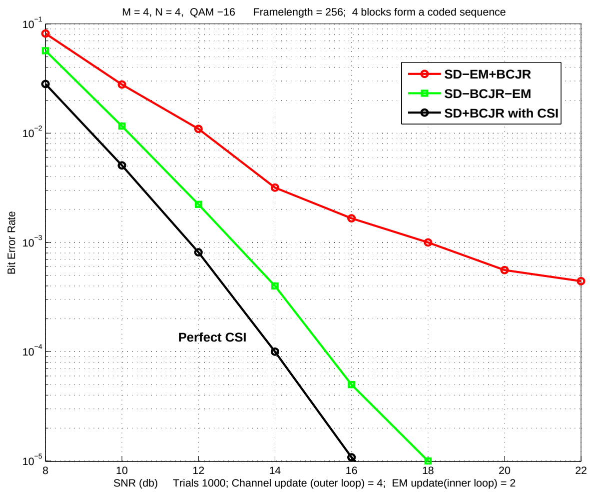

Fig. 8: BER improvements by utilizing time diversity and channel coding. Triangle line is the performance of the SD-EM algorithm following the BCJR algorithm and the square line is the performance of the coding assisted SD-EM algorithm. 4 blocks form a codeword

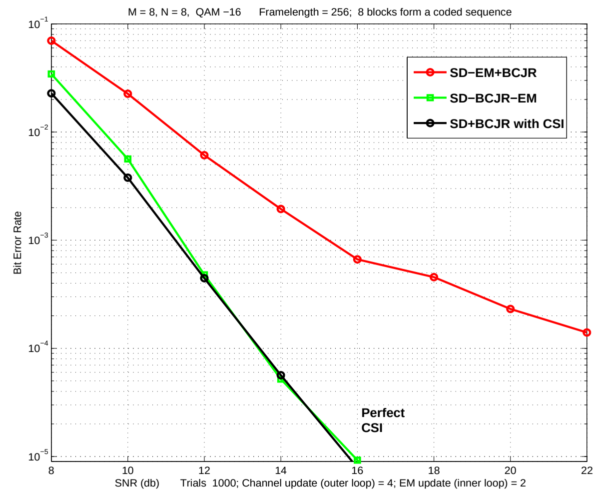

Fig. 10: BER improvements by utilizing time diversity and channel coding. Triangle line is the performance of the SD-EM algorithm following the BCJR algorithm and the square line is the performance of the coding assisted SD-EM algorithm. 8 blocks form a codeword

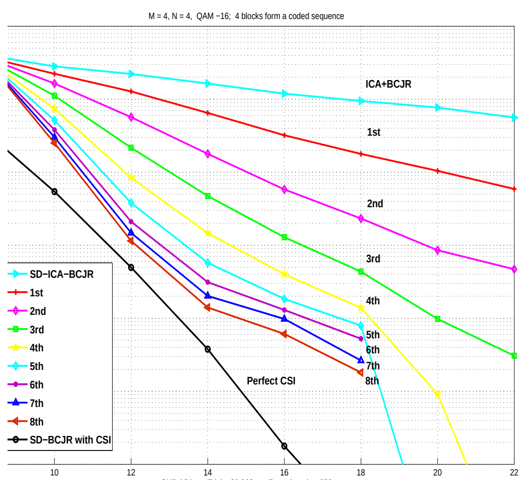

Fig. 9: Iterative BER improvements of the coding assisted SD-EM algorithm in a 4×4 system with 16-QAM modulation, 8 iterations are used.

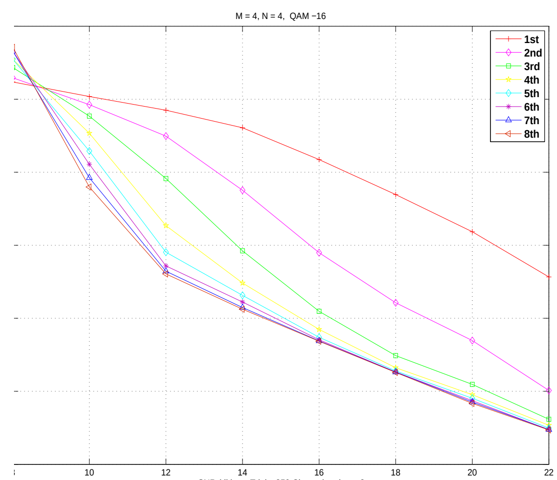

Fig. 11: Iterative channel improvements of the coding assisted SD-EM algorithm in a 4×4 system with 16-QAM modulation.

Carlo runs. Similar to the setup above with 4 fixed channel realizations. One channel matrix is constrained to be very singular with a condition number over 25 so that it is difficult to obtain a good initial estimate of the channel matrix. The other channels have good condition numbers. In this scenario it is reasonable to expect that more iterations of the EM algorithm might be necessary in order to obtain a good channel estimates for the singular channel matrix. Here we see that the convergence of the algorithm is still quite reasonable.

The BER performance is evaluated for different number of iterations of the EM channel estimation. Eight iterations are studied. Clearly, from Figure 9, we can see that the performance progresses towards the optimal curve with CSI known at the receiver and the most significant improvements in BER occur in the first 5 iterations. While there is still a gap between our scheme's performance and that with perfect CSI we speculate that the gap will further decrease if the block length is increased (though with the inevitable increase in decoding latency).

The quality of the singular channel matrix estimate is illustrated in Figure 11 where the ICI is calculated for each iteration over a range of SNR. Here, again, it can be seen that the EM algorithm has effectively converged (this time in terms of the channel estimate) in about 5 iterations.

The convergence rate of the EM algorithm appears to be slower in the presence of singular channels. Thus to gain full advantage of the coding assisted channel estimation we could monitor the singularity of the estimated channels and use one or two iterations when the channel condition appears good while increasing the number of iterations when a singular channel is observed. A trade-off can therefore be made between the performance and the complexity of the estimator.

V. Conclusion

A coding assisted MIMO blind separation and decoding scheme is proposed. Three techniques of separation, diversity and channel coding are used to improve fading link performance (BER). By utilizing a posteriori information, our scheme provides substantial gain over the uncoded system. The existence of coding structures partly solves the problems of EM getting trapped in a local minimum when the channel is close to singular or the SNR is low. This happens frequently when the number of receiver antennas, the size and the dimension of the data are large. The new scheme appears to avoid local minima and converges to the global minimum or at least a good approximation of it. Moreover, this system extends FEC to the multiple blocks in order to form a large codeword and then exploits the temporal diversity. This extension also improved system performance.

References

[1] S. Zhou, B. Muquet, and G.B. Giannakis, "Subspace-based (semi-) blind channel estimation for block precoded space-time OFDM," *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1215–1228, May 2002.

[2] R. Zhang, "Blind OFDM channel estimation through linear precoding: a subspace approach," *In Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 631–633, Nov, 2002.

[3] X. Liu and M.E. Bialkowski, "SVD-based blind channel estimation for a MIMO OFDM system employing a simple block pre-coding scheme," in *International Conference on "Computer as a Tool" EUROCON 2007*, 9–12 Sep. 2007, pp. 926–929.

[4] D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1867–1875, Nov. 1980.

[5] J. Treichler and B. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 31, no. 2, pp. 459–472, Apr. 1983.

[6] M. Tuchler, R. Koetter, and A. C. Singer, "Turbo equalization: principles and new results," *IEEE Transactions on Communications*, vol. 50, no. 5, pp. 754–767, May 2002.

[7] J. Garcia-Frias and J.D. Villasenor, "Blind turbo decoding and equalization," in *Proc. IEEE 49th Vehicular Technology Conference*, 16–20 May 1999, vol. 3, pp. 1881–1885.

[8] P. Ha and B. Honary, "Improved blind turbo detector," in *Proc. IEEE 51st VTC 2000-Spring Tokyo Vehicular Technology*, 15–18 May 2000, vol. 2, pp. 1196–1199.

[9] C-Y Chi, C-C Feng, C-H Chen, C-Y Chen, *Blind Equalization and System Identification*, Springer, 2006.

[10] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, New York: John Wiley & Sons, 2001.

[11] J. Mannerkoski and V. Koivunen, "Autocorrelation properties of channel encoded sequences-applicability to blind equalization," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3501–3507, Dec. 2000.

[12] J. Garcia-Frias and J.D. Villasenor, "Combined turbo detection and decoding for unknown isi channels," *IEEE Transactions on Communications*, vol. 51, no. 1, pp. 79–85, 2003.

[13] L.R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *IEEE Proceedings*, vol. 77, no. 2, pp. 257–286, 1989.

[14] R.R. Lopes and J.R. Barry, "Exploiting error-control coding in blind channel estimation," in *Proc. IEEE Global Telecommunications Conference GLOBECOM '01*, 2001, vol. 2, pp. 1317–1321 vol.2.

[15] R.R. Lopes and J.R. Barry, "Blind iterative channel identification and equalization," in *Proc. IEEE International Conference on Communications ICC 2001*, 2001, vol. 7, pp. 2256–2260 vol.7.

[16] J. Gunther, D. Keller, and T. Moon, "A generalized BCJR algorithm and its use in iterative blind channel identification," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 661–664, Oct. 2007.

[17] J.H. Gunther, M. Ankapura, and T.K. Moon, "A generalized LDPC decoder for blind turbo equalization," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3847–3856, 2005.

[18] R. Raheli, A. Polydoros, and Ching-Kae Tzou, "Per-survivor processing: a general approach to MLSE in uncertain environments," *IEEE Transactions on Communications*, vol. 43, no. 234, pp. 354–364, 1995.

[19] H. Nguyen, B.C. Levy, "Blind and semi-blind equalization of CPM signals with the EMV algorithm," *IEEE Transactions on Signal Processing*, vol. 51, no. 10, pp. 2650–2664, Oct. 2003.

[20] X. Zhao, M. Davies, "A feasible blind equalization scheme in large constellation MIMO systems," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP 2008*, March 31 2008–April 4 2008, pp. 1845–1848.

[21] F. Sanzi, Sven Jelting, and J. Speidel, "A comparative study of iterative channel estimators for mobile OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 2, no. 5, pp. 849–859, 2003.

[22] F.J. Vázquez-Araújo, M. González-López, L. Castedo , and J. Garcia-Frias, "Serially-concatenated LDGM codes for MIMO channels," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 2860–2871, 2007.

[23] M. González-López, J. Míguez, and L. Castedo, "Maximum likelihood turbo iterative channel estimation for space-time coded systems and its application to GSM in subway tunnels," *EURASIP Journal on Applied Signal Processing*, vol. 2004, pp. 727–739, May 2004.

[24] I.T Jolliffe, *Principal Component Analysis*, Springer Series in Statistics, 2002.

[25] Gene H. Golub, Charles F. Van Loan, *Matrix Computations (3rd ed.)*, Johns Hopkins, 1996.

[26] H. Liu, *Signal Processing Applications in CDMA Communications*, Norwood, MA: Artech House, 2000.

[27] M. El-Khamy, H. Vikalo, and B. Hassibi, "Bounds on the performance of sphere decoding of linear block codes," in *Proc. IEEE Information Theory Workshop*, 29 Aug.–1 Sept. 2005, p. 5pp.

[28] H. Vikalo and B. Hassibi, "On joint detection and decoding of linear block codes on Gaussian vector channels," *IEEE Transactions on Signal Processing*, vol. 54, no. 9, pp. 3330–3342, Sept. 2006.

[29] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (corresp.)," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, Mar 1974.

[30] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo-codes. 1," in *IEEE International Conference on Communications ICC 93. Geneva. Technical Program, Conference Record*, 1993, vol. 2, pp. 1064–1070 vol.2.

[31] K.S. Sooyoung, H-M. Kim, and J. Gwak, "Reduced complexity sliding window BCJR decoding algorithms for turbo codesldpc," in *Lecture Notes in Computer Science Proceedings of the 7th IMA International Conference on Cryptography and Coding*, 1999.

[32] B.M. Hochwald and S. Ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 389–399, 2003.

[33] Y.L.C. de Jong, T. J. Willink, "Iterative tree search detection for MIMO wireless systems," *IEEE Transactions on Communications*, vol. 53, no. 6, pp. 930–935, June 2005.

[34] D. L. Milliner, E. Zimmermann, J. R. Barry, and G. Fettweis, "Channel state information based LLR clipping in list MIMO detection," in *Proc. IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications PIMRC 2008*, 15–18 Sept. 2008, pp. 1–5.

[35] F. Schreckenbach, N. Gortz, J. Hagenauer, and G. Bauch, "Optimization of symbol mappings for bit-interleaved coded modulation with iterative decoding," *IEEE Communications Letters*, vol. 7, no. 12, pp. 593–595, 2003.

[36] A. Dempster, N. Laird, and D. Rubink, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society*, vol. Series B, 39(1), pp. 1–38, 1977.

[37] G. K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels," *IEEE Transactions on Communications*, vol. 42, no. 7, pp. 2406–2413, July 1994.

[38] L.B. White, S. Perreau, and P. Duhamel, "Reduced computation blind equalization for FIR channel input Markov models," in *Proc. IEEE International Conference on Communications ICC 95 Seattle, Gateway to Globalization*, 18–22 June 1995, vol. 2, pp. 993–997.

[39] S. Min and C.L. Nikias, "An ML/MMSE estimation approach to blind equalization," in *IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP*, April 1994, vol. 4, pp. 569–572.

[40] H.A. Cirpan and M.K. Tsatsanis, "Stochastic maximum likelihood methods for semi-blind channel equalization," in *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems & Computers*, 2–5 Nov. 1997, vol. 2, pp. 1629–1632.

[41] T.K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 47–60, Nov. 1996.

[42] O. Bermond and J.F. Cardoso, "Approximate likelihood for noisy mixtures," in *First International Workshop on Independent Component Analysis*, 1999, pp. 325–330.

[43] G.A. Al-Rawi, T.Y. Al-Naffouri, A. Bahai, and J. Cioffi, "Exploiting error-control coding and cyclic-prefix in channel estimation for coded OFDM systems," *IEEE Conmunications Letters*, vol. 7, no. 8, pp. 388–390, 2003.

[44] C. F. Jeff Wu, "On the convergence properties of the EM algorithm," *The Annals of Statistics*, vol. 11(1), pp. 95103, 1983.

[45] X. Zhao, M. Davies, "ICA with the EM algorithm in the low noise case," in *ICA Research Network International Workshop 2006*, Sept. 2006.

[46] A. Edelman, *Eigenvalues and Condition Numbers of Random Matrices*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge (MA), 1989.

[47] K. Su, I. Berenguer, I.J. Wassell, and X. Wang, "Efficient maximum-likelihood decoding of spherical lattice space-time codes," in *IEEE International Conference on Communications ICC '06*, 2006, vol. 7, pp. 3008–3013.

[48] L. G. Barbero and J. S. Thompson, "Extending a fixed-complexity sphere decoder to obtain likelihood information for turbo-MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 57s, pp. 2804 – 2814, 2008.

PLACE
PHOTO
HERE

Xu Zhao received the B.Eng. degree in automation engineering from Huazhong University of Science and Technology, Wuhan, China, in 1998, the M.Sc. degree from Queen Mary, University of London, U.K., in 2005. He is currently working toward the Ph.D. degree in the Institute for Digital Communications, at the University of Edinburgh. His research interests are in wireless communications, blind signal separations, independent component analysis and MIMO systems.

PLACE
PHOTO
HERE

Mike Davies (M00) received the B.A. (Hons.) degree in engineering from Cambridge University, Cambridge, U.K., in 1989 and the Ph.D. degree in nonlinear dynamics and signal processing from University College London, London (UCL), U.K., in 1993. Mike Davies was awarded a Royal Society Research Fellowship in 1993 and was an Associate Editor for IEEE Transactions in Speech, Language and Audio Processing, 2003-2007.

He currently holds the Jeffrey Collins SFC funded chair in Signal and Image Processing at the University of Edinburgh. His research interests include: statistical signal processing, array signal processing, sparse approximation, compressed sensing and their applications.