LDPC-Coded MIMO Systems With Unknown Block Fading Channels: Soft MIMO Detector Design, Channel Estimation, and Code Optimization

Jun Zheng, *Student Member, IEEE,* and Bhaskar D. Rao, *Fellow, IEEE*

Department of Electrical and Computer Engineering

University of California, San Diego

La Jolla, CA 92093-0407

E-mail: juzheng@ucsd.edu, brao@ece.ucsd.edu

Abstract

In this paper, we consider the design of a practical LDPC-coded MIMO system composed of M transmit and N receive antennas operating in a flat fading environment where channel state information is assumed to be unavailable both to the transmitter and the receiver. A soft iterative receiver structure is developed which consists of three main blocks, a soft MIMO detector and two LDPC component soft decoders. We first propose at the component level several soft-input soft-output MIMO detectors whose performances are much better than the conventional MMSE-based detectors. In particular, one optimal soft MIMO detector and two simplified sub-optimal detectors are developed that do not require an explicit channel estimate and offer an effective tradeoff between complexity and performance. In addition, a modified EM-based MIMO detector is developed which completely removes positive feedback between input and output extrinsic information and provide much better performance compared to the direct EM-based detector that has strong correlations especially in fast fading channels. At the structural level, the LDPC-coded MIMO receiver is constructed in an unconventional manner where the soft MIMO detector and LDPC variable node decoder form one super soft-decoding unit, and the LDPC check node decoder forms the other component of the iterative decoding scheme. By exploiting the proposed receiver structure, tractable extrinsic information transfer functions of the component soft decoders are obtained, which further lead to a simple and efficient LDPC code degree profile optimization algorithm with proven global optimality and guaranteed convergence from any initialization. Finally, numerical and simulation results are provided to confirm the advantages of the proposed design approach for the coded MIMO system.

This research was supported by UC Discovery grant No.02-10109 sponsored by Ericsson.

I. Introduction

Communication systems using multiple antennas at both the transmitter and the receiver have recently received increased attention due to their ability to provide great capacity increases in a wireless fading environment [1] [2]. However, MIMO capacity analysis and system design is often based on the assumption that the fading channel coefficient between each transmit and receive antenna pair is perfectly known at the receiver. This is not a realistic assumption for most practical communication systems especially in fast fading channels.

For communication systems with unknown channel state information (CSI) at both ends, conventional receivers usually have a two-phase structure, data-aided channel estimation using the preset training symbols followed by coherent data detection by treating the estimated channel as the actual channel coefficients. Due to the importance of channel estimator, which directly determines estimation quality and hence the overall system performance, various MIMO channel estimation algorithms have been studied [3]– [5]. However, conventional channel estimators form estimates based only on the training symbols, thereby failing to make use of the channel information contained in the received data symbols. Consequently, the two-phase model limits the performance and can not approach the MIMO channel capacity (or the maximum achievable information rate), especially in a fast fading environment (with small channel coherence time). Possible solutions to the above problem include use of blind source signal separation algorithms [6]– [8], MIMO differential modulation [9]– [11], and unitary space-time modulation (USTM) [12]– [18]. However, none of these schemes can approach the non-coherent MIMO capacity limit due to their sub-optimal code structure, and in the later case, USTM, only asymptotic (or the diversity) optimality is achieved in high SNR regimes and the approach suffers from exponential decoding complexity.

In order to achieve better spectral efficiency than the conventional data-aided estimation algorithm that uses large number of training symbols for accurate channel estimation, the so-called code-aided joint channel estimation and data detection algorithms have recently received much attention. By treating the unknown channel as unobserved (or missing) data, ML sequence estimation of the coded data frames using the EM algorithm was proposed by Georghiades [19] and Kaleh [20] over single input single output fading channels and extended to MIMO channels by Cozzo [21]. Alternatively, several recent publications [22]- [24] have developed EM-based algorithms that can iteratively improve the channel estimate based on the soft extrinsic information from the outer soft decoder, and the schemes work well in an iterative receiver structure.

In this paper, we focus on the design of practical LDPC-coded MIMO systems employing a soft iterative receiver structure consisting of three soft decoding component blocks, a soft MIMO detector and two soft LDPC component decoders (variable node and check node decoders). At the component level, we first propose a soft optimal MIMO detector, which can generate soft log likelihood ratio (LLR) of each coded bit under the condition of unknown CSIR without forming any explicit channel estimate. Based on the proposed soft optimal detector, we develop two simplified sub-optimal MIMO detectors with polynomial and log polynomial decoding complexities. In addition, motivated by the EM-based detection algorithm in [24], we also propose in the MIMO context a modified EM-based detector that completely removes the positive feedback between the input and output extrinsic information and provides much better performance compared to the direct EM-based detector that has strong correlations. By analyzing the mutual information transfer characteristic [25] of the proposed soft MIMO detectors, system performance of different MIMO detection algorithms are analyzed and compared under various channel conditions. At the structural level, inspired by the turbo iterative principle [26], the LDPC-coded MIMO receiver is constructed in an unconventional manner where the soft MIMO detector and LDPC variable node decoder form one super soft-decoding unit and the LDPC check node decoder forms the other component of the iterative decoding scheme. Utilizing the proposed receiver structure, tractable extrinsic information transfer functions of the component soft decoders are obtained, which lead to a simple and efficient LDPC code degree profile optimization algorithm. This algorithm is shown to have global optimality and guaranteed convergence from any initialization, and is an improvement over the sub-optimal manual curve fitting technique proposed in [27]. Numerical and simulation results of the LDPC-coded MIMO system using the optimized degree profile further confirm the advantages of the proposed design approach for the coded MIMO system.

The rest of the paper is organized as follows. Section II describes the LDPC-coded MIMO system structure as well as the unknown block fading channel model. Section III proposes several different soft MIMO detectors that can be used as the building blocks for the turbo iterative MIMO receivers. In section IV, the receiver design of the coded MIMO systems is addressed in detail, which includes the overall receiver structure in Section IV-A, the extrinsic mutual information transfer characteristic analysis in Section IV-B, and the LDPC code degree profile optimization algorithm in Section IV-C. In Section V, the simulation results of the LDPC-coded MIMO system under various channel conditions are presented. Finally, conclusions are drawn in Section VI.

II. System Model

A. MIMO transmitter structure

We consider a MIMO system with M transmit antennas and N receive antennas signaling through a frequency flat fading channel with independent channel propagation coefficient between each transmit

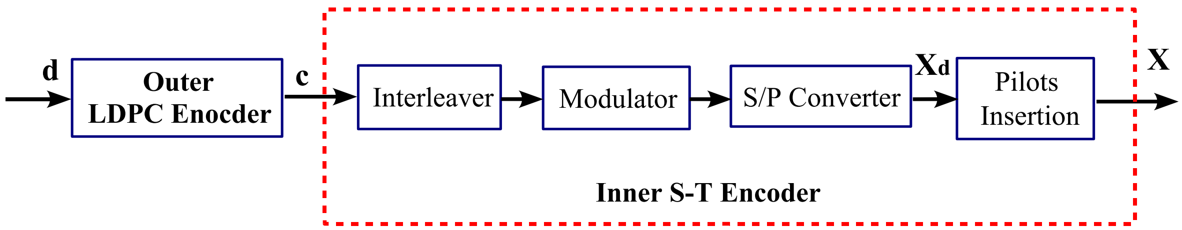

Fig. 1. Transmitter model of LDPC-coded MIMO systems

and receive antenna pair. As illustrated in Fig. 1, a block of k binary information bits denoted $\mathbf{d} = \{d_1, \cdots, d_k\}$ is first encoded by an outer LDPC encoder with code rate $R_{\text{outer}} = k/n$ into a codeword $\mathbf{c} = \{c_1, \cdots, c_n\}$ of length n. The codeword $\mathbf{c}$ is further segmented into L consecutive sub-blocks $\mathbf{C}_i$ of length K. Each sub-block $\mathbf{C}_i$ is then encoded by the inner space-time encoder into a coherent space-time *sub-frame* $\mathbf{X}_i$. This encoder is composed of an interleaver, modulator, serial-to-parallel converter, and a pilot insertion operator. The symbol structure of each sub-frame $\mathbf{X}_i$ is illustrated in Fig. 2, where the first

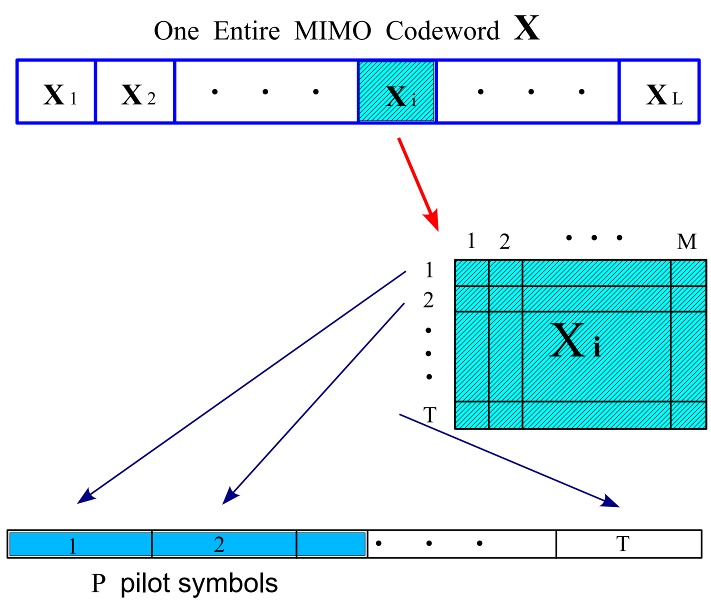

Fig. 2. Transmitted symbol structure of the coded MIMO system

p symbols are training pilots, followed by $(TM - p)$ data symbols. For the sake of simplicity, we only consider the case where both the number of pilot symbols $\left(p = T_\tau \times M\right)$ and the number of data symbols $\left(TM - p = T_d \times M\right)$ are multiples of the transmit antenna number M. We further denote the average signal to noise ratio (SNR) of pilot symbols by ρ_τ and data symbols by ρ_d. Hence, the transmitted signal

$\mathbf{X}_i$ can be partitioned into two sub-matrices: training followed by data, which is represented as

$$\mathbf{X}_i = \begin{bmatrix} (\rho_\tau/M)^{\frac{1}{2}} \cdot \mathbf{X}_\tau \\ (\rho_d/M)^{\frac{1}{2}} \cdot \mathbf{X}_{d,\,i} \end{bmatrix}, \tag{1}$$

where $\mathbf{X}_\tau \in \mathbb{C}^{T_\tau \times M}$ are the fixed pilot symbols sent over T_τ time intervals, and $\mathbf{X}_{d,\,i} \in \mathbb{C}^{T_d \times M}$ are the information bearing data symbols sent over T_d transmission intervals. Each element of the transmitted data signal $\mathbf{X}_{d,\,i}$ is a member of a finite complex alphabet $\mathcal{X}$ of size $|\mathcal{X}|$. One entire MIMO codeword $\mathbf{X}$ consists of $l = LTM$ complex symbols, which are transmitted from M transmit antennas and across L consecutive coherent sub-frames of length TM symbols.

It is assumed that the fading coefficient matrix $\mathbf{H}_i$ remains static within each coherent sub-block and varies independently from one sub-block to another. Hence, the signal model can be written as

$$\mathbf{Y}_i = \mathbf{X}_i \cdot \mathbf{H}_i + \mathbf{w}_i, \qquad 1 \le i \le L \ , \tag{2}$$

where $\mathbf{Y}_i$ is a $T \times N$ received complex signal matrix, $\mathbf{X}_i$ is a $T \times M$ transmitted complex signal matrix, $\mathbf{H}_i$ is an $M \times N$ complex channel matrix, and $\mathbf{w}_i$ is a $T \times N$ matrix of additive noise matrix. Both matrices $\mathbf{H}_i$ and $\mathbf{w}_i$ are assumed to have zero mean unit variance independent complex Gaussian entries. We also assume that the entries of the transmitted signal matrix $\mathbf{X}_i$ have, on average, the following power constraint,

$$\frac{1}{T} \cdot E\big[\mathbf{tr}\big(\mathbf{X}_i^H \mathbf{X}_i\big)\big] = \rho \ . \tag{3}$$

where ρ is the average signal to noise ratio at each receive antenna. Conservation of time and energy leads to the following constraints,

$$\begin{aligned} \mathbf{tr}\big(\mathbf{X}_\tau^H \cdot \mathbf{X}_\tau\big) &= MT_\tau, & E_{\mathbf{X}_{d,\,i}}\Big[\mathbf{tr}\Big(\mathbf{X}_{d,\,i}^H \cdot \mathbf{X}_{d,\,i}\Big)\Big] &= MT_d \ , \\ T &= T_\tau + T_d, & \rho T &= \rho_\tau T_\tau + \rho_d T_d \ . \end{aligned} \tag{4}$$

Due to the insignificant capacity gain resulting from using optimal power allocation between training and data symbols as reported in [28] [29], equal power allocation is assumed in this paper, with

$$\rho_\tau = \rho_d = \rho \ . \tag{5}$$

B. MIMO receiver structure

The MIMO receiver decodes the transmitted information bits $\mathbf{d}$ based on received signal matrices $\{\mathbf{Y}_i\}_{i=1}^L$ without knowing any instantaneous channel state information $\{\mathbf{H}_i\}_{i=1}^L$. The channel statistical distribution $p(\mathbf{H}_i)$ is assumed to be known both to the receiver and to the transmitter throughout the paper. We know that even with ideal CSI, the optimal decoding algorithm for this system has an exponential

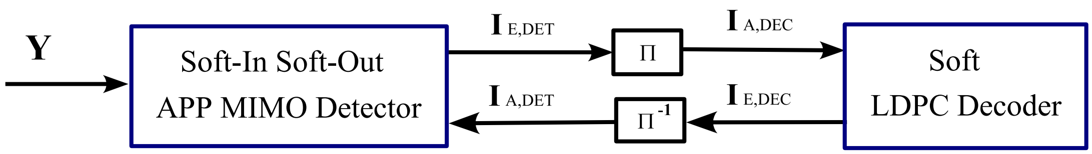

Fig. 3. Conventional receiver structure of LDPC-coded MIMO systems

complexity. Hence the near-optimal iterative receiver structure based on turbo principle [26] becomes a promising alternative.

As a standard iterative decoding procedure, the structure of the LDPC-coded MIMO receiver is demonstrated in Fig. 3. It consists of two important components, an inner soft-input soft-output MIMO detector (or MIMO demodulator) and an outer soft LDPC decoder, which form a bipartite graph structure [30]. Soft log likelihood ratio of each transmitted bit is passed forward and backward between these two soft decoders with increasing accuracy as the number of the iterations increase. At each iteration, the MIMO detector forms soft extrinsic information of each coded bit based on the received symbols $\{\mathbf{Y}_i\}_{i=1}^{L}$ and the apriori information coming from the soft LDPC decoder through proper interleaving, and serves as the apriori information for the LDPC decoder in the next iteration. Convergence is reached after certain number of the iterations and decoded bits are hence obtained.

Notice that the LDPC decoder in Fig. 3 is itself composed of two component soft decoders (variable node decoder and check node decoder), and the entire MIMO receiver can be viewed as a complicated graph code structure. Therefore, any bipartite separation other than the conventional structure can lead to an alternative iterative decoder. We utilize in Section IV-A an unconventional MIMO receiver structure, which combines the soft MIMO detector and LDPC variable node decoder together as a super component soft-decoder. The proposed receiver structure has great design advantages that can easily lead to an efficient LDPC code degree profile optimization algorithm as shown in Section IV-C.

III. Soft-Input Soft-Output MIMO Detector

As described in Section II-B, the soft-input soft-output MIMO detector is an important decoding component of the MIMO receiver, and plays an important role in determining the performance of the entire coded MIMO system. Regular communication systems with unknown channel state information typically employ a two-stage decoding procedure, which consists of channel estimation followed by coherent decoding based on the estimated channel parameters. However, conventional channel estimators perform estimation based only on the training pilots, thereby failing to make use of the channel information contained in the data symbols. Due to the mismatch between the actual and estimated channel, system performance suffers severe degradation especially in a communication environment with low signal to noise ratio, or limited training pilots in fast fading channels.

In this section, several better MIMO detectors which include the soft MIMO detector, EM-based MIMO detector, as well as their modified versions are proposed that offer an effective tradeoff between detection complexity and performance. The MIMO detection algorithms proposed in this section are block-based in the sense that the data detections are performed within each coherent fading block. By considering the channel coefficient correlations between adjacent coherent blocks, one could achieve even better performance by performing data detection on several adjacent coherent blocks together. In this case, the data detection algorithm has higher computational complexity and depends heavily on the correlations of the fading channel, and is beyond the scope of this paper. Therefore for simplicity, it is reasonable to use a block fading channel model in this situation and the performance penalty of the simple block-based MIMO detection algorithm would be small by properly tuning the channel coherence time T according to the actual channel correlations.

For the sake of simplicity, subscript (or time index) i, denoting the i^{th} coherent block, is dropped in this section while describing the block-wise soft MIMO detection algorithms. To be specific, we denote $\mathbf{X} = \left[\mathbf{X}_\tau^T, \mathbf{X}_d^T\right]^T$, $\mathbf{H}$, and $\mathbf{Y} = \left[\mathbf{Y}_\tau^T, \mathbf{Y}_d^T\right]^T$ as the transmitted signal, channel matrix, and received signal in each coherent block, respectively. Furthermore, sub-matrices $\mathbf{X}_\tau$, $\mathbf{X}_d$, $\mathbf{Y}_\tau$, and $\mathbf{Y}_d$ have the following structures, i.e.

$$\begin{aligned} \mathbf{X}_\tau &= \left[\mathbf{x}_{\tau,1}^T, \cdots, \mathbf{x}_{\tau,T_\tau}^T\right]^T, \qquad \mathbf{X}_d = \left[\mathbf{x}_{d,1}^T, \cdots, \mathbf{x}_{d,T_d}^T\right]^T, \qquad \mathbf{x}_{\tau,k}\Big|_{k=1}^{T_\tau}, \ \mathbf{x}_{d,k}\Big|_{k=1}^{T_d} \in \mathbb{C}^{1\times M}, \\ \mathbf{Y}_\tau &= \left[\mathbf{y}_{\tau,1}^T, \cdots, \mathbf{y}_{\tau,T_\tau}^T\right]^T, \qquad \mathbf{Y}_d = \left[\mathbf{y}_{d,1}^T, \cdots, \mathbf{y}_{d,T_d}^T\right]^T, \qquad \mathbf{y}_{\tau,k}\Big|_{k=1}^{T_\tau}, \ \mathbf{y}_{d,k}\Big|_{k=1}^{T_d} \in \mathbb{C}^{1\times M}. \end{aligned} \tag{6}$$

Similarly, the binary sub-codeword $\mathbf{C}$ that maps to the transmitted signal $\mathbf{X}$ can also be decomposed into

$$\mathbf{C} = \left[\mathbf{c}_1^T, \cdots, \mathbf{c}_{T_d}^T\right]^T, \qquad \mathbf{c}_k\Big|_{k=1}^{T_d} \in \mathbb{B}^{1\times M\cdot \log_2|\mathcal{X}|}, \tag{7}$$

where $\mathbb{B}$ is binary set $\{0,1\}$ and each row $\mathbf{c}_k$ represents the corresponding binary information that maps to $\mathbf{x}_{d,k}$.

A. Optimal soft MIMO detector

First, according to the channel model (2), the conditional probability density of the received signal matrix $\mathbf{Y}$ given the transmitted signal matrix $\mathbf{X}$ is given by [31]

$$p(\mathbf{Y}|\mathbf{X}) = \frac{\exp\left(-\mathbf{tr}\left\{\left[I_T + \mathbf{X}\mathbf{X}^H\right]^{-1}\cdot \mathbf{Y}\mathbf{Y}^H\right\}\right)}{\pi^{TN}\det^N\left[I_T + \mathbf{X}\mathbf{X}^H\right]}. \tag{8}$$

It is evident from the above transitional probability that the unknown MIMO channel is actually a memoryless vector channel and hence the optimal MIMO detector does not necessarily need to form a specific channel estimate.

In order to obtain the a posteriori probability of each coded bit, the a priori probability of the input signal matrix $\mathbf{X}$ is first calculated as

$$p\big(\mathbf{X}\big) = p\big(\mathbf{X}_d\big) = p(\mathbf{C}) = \prod_{k=1}^{T_d} p(\mathbf{x}_k) = \prod_{k=1}^{T_d} p(\mathbf{c}_k) = \prod_{k=1}^{T_d} \prod_{j=1}^{M \log_2 |\mathcal{X}|} p(c_{k,j}) \ , \tag{9}$$

where each element of matrix $\mathbf{X}_d$ is a member of a complex alphabet $\mathcal{X}$ of size $|\mathcal{X}|$, and corresponding to $\log_2 |\mathcal{X}|$ LDPC-coded bits. Therefore, the log likelihood ratio of each LDPC coded bit is given by

$$\begin{aligned} L_{\text{pos}}(c_{k,j}) &= \log\left(\frac{p\big(c_{k,j}=1\big|\mathbf{Y}\big)}{p\big(c_{k,j}=0\big|\mathbf{Y}\big)}\right) \\ &= \log\left(\frac{\Sigma_{\mathbf{X}\in\mathcal{D}^{+}_{k,j}}\ p\Big(\mathbf{Y}\Big|\mathbf{X}\Big)\cdot p\big(\mathbf{X}\big)}{\Sigma_{\mathbf{X}\in\mathcal{D}^{-}_{k,j}}\ p\Big(\mathbf{Y}\Big|\mathbf{X}\Big)\cdot p\big(\mathbf{X}\big)}\right), \qquad 1\le k\le T_d, \quad 1\le j\le M\cdot\log_2|\mathcal{X}| \ , \end{aligned} \tag{10}$$

where $\mathcal{D}^{+}_{k,j}$ ($\mathcal{D}^{-}_{k,j}$) is the set of $\mathbf{X}$ for which the $(k,j)^{\text{th}}$ bit $c_{k,j}$ of the LDPC coded sub-block $\mathbf{C}$ is "$+1$" ("-1"). Finally, by subtracting the input a priori information from the obtained a posterior log likelihood ratio, the soft extrinsic information of each coded bit is obtained as

$$L_{\text{ext}}(c_{k,j}) = L_{\text{pos}}(c_{k,j}) - L_{\text{app}}(c_{k,j}), \qquad L_{\text{app}}(c_{k,j}) = \log\left(\frac{p\big(c_{k,j}=1\big)}{p\big(c_{k,j}=0\big)}\right) \ , \tag{11}$$

where $L_{\text{app}}(c_{k,j})$ is the a priori information of the coded bit $c_{k,j}$ from the last iteration. Notice that there is no channel estimation stage in the soft MIMO detector described above, and therefore the proposed detection algorithm does not depend on the unknown channel state $\mathbf{H}$ but only on its underlying statistical distribution. Furthermore, the optimality of the proposed soft MIMO detection algorithm is restricted within the component level and does not depend on the overall receiver structure of the coded-MIMO system.

B. Sub-optimal soft MIMO detector

The optimal soft MIMO detection algorithm proposed in Section III-A provides the optimal extrinsic LLR values of each coded bit. However, the summation in both the numerator and the denominator of equation (10) consists of 2^{K-1} items, with $K \ \big(=T_d M \log_2 |\mathcal{X}|\big)$ increasing linearly with number of data slots T_d (or coherence time T). It has an unaffordable exponential complexity for practical communication systems, especially when the coherence time T is large. Hence, we propose a sub-optimal MIMO detector in this section with complexity increasing linearly with T_d.

Notice that the optimal extrinsic LLR value of bit $c_{k,j}$ depends on the input a priori information as well as the channel observations of the entire coherent block. Taking another point of view, the obtained

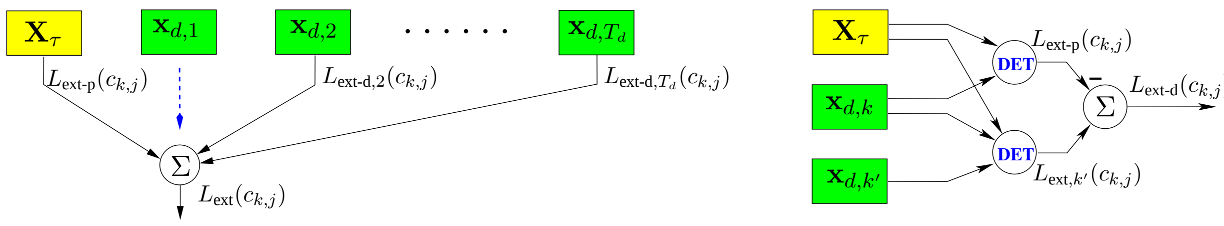

Fig. 4. Sub-optimal soft MIMO detector structure

extrinsic LLR is a combination of all the input information through the utilization of the proposed algorithm (10) in an implicit manner. Therefore, instead of performing soft MIMO detection in one operation, we can extract partial extrinsic information by processing only two rows of the data matrix $\mathbf{X}_d$ at a time, and then combining different partial extrinsic information to form the final extrinsic LLR. As illustrated in (the right side of) Fig. 4, in order to combine information from coded rows $\mathbf{x}_{d,k}$ and $\mathbf{x}_{d,k'}$, we first perform the optimal MIMO detection algorithm on the following reduced size *sub-coherent* block

$$\mathbf{X}_{[k,k']} = \left[{\mathbf{X}_\tau}^T, \mathbf{x}_{d,k}^T, \mathbf{x}_{d,k'}^T\right]^T, \qquad \mathbf{Y}_{[k,k']} = \left[{\mathbf{Y}_\tau}^T, \mathbf{y}_{d,k}^T, \mathbf{y}_{d,k'}^T\right]^T . \tag{12}$$

Therefore, the partial extrinsic LLR value $L_{\text{ext},\,k'}(c_{k,j})$ of bit $c_{k,j}$ obtained from the a priori information of row $\mathbf{c}_k$, $\mathbf{c}_{k'}$, and channel observation $\mathbf{Y}_{[k,k']}$ is given by

$$L_{\text{ext},\,k'}(c_{k,j}) = \log\left(\frac{\Sigma_{\mathbf{X}_{[k,k']}\in\mathcal{D}^+_{k,j}} p\left(\mathbf{Y}_{[k,k']}\middle|\mathbf{X}_{[k,k']}\right)\cdot p\left(\mathbf{X}_{[k,k']}\right)}{\Sigma_{\mathbf{X}_{[k,k']}\in\mathcal{D}^-_{k,j}} p\left(\mathbf{Y}_{[k,k']}\middle|\mathbf{X}_{[k,k']}\right)\cdot p\left(\mathbf{X}_{[k,k']}\right)}\right) - \log\left(\frac{p(c_{k,j}=1)}{p(c_{k,j}=0)}\right) , \tag{13}$$

where

$$1 \le k, k' \le T_d, \qquad 1 \le j \le M \log_2 |\mathcal{X}| \ ,$$

and $\mathcal{D}^+_{k,j}$ ($\mathcal{D}^-_{k,j}$) is the set of $\mathbf{X}_{[k,k']}$ for which bit $c_{k,j}$ is "$+1$" ("-1"). By the same reasoning, partial extrinsic information of bit $c_{k,j}$, related to (and contained in) the a priori information of $\mathbf{c}_k$ and channel observations $\mathbf{Y}_\tau$ and $\mathbf{y}_{d,k}$ can also be obtained by performing optimal detection on the following *sub-coherent* block

$$\mathbf{X}_{[k]} = \left[\mathbf{X}_\tau^T, \mathbf{x}_{d,k}^T\right]^T, \qquad \mathbf{Y}_{[k]} = \left[\mathbf{Y}_\tau^T, \mathbf{y}_{d,k}^T\right]^T , \tag{14}$$

with the corresponding extrinsic LLR value given by

$$L_{\text{ext-p}}(c_{k,j}) = \log\left(\frac{\Sigma_{\mathbf{X}_{[k]}\in\mathcal{D}^+_{k,j}} p\left(\mathbf{Y}_{[k]}\middle|\mathbf{X}_{[k]}\right)\cdot p\left(\mathbf{X}_{[k]}\right)}{\Sigma_{\mathbf{X}_{[k]}\in\mathcal{D}^-_{k,j}} p\left(\mathbf{Y}_{[k]}\middle|\mathbf{X}_{[k]}\right)\cdot p\left(\mathbf{X}_{[k]}\right)}\right) - \log\left(\frac{p(c_{k,j}=1)}{p(c_{k,j}=0)}\right) , \tag{15}$$

where

$$1 \le k \le T_d, \qquad 1 \le j \le M \log_2 |\mathcal{X}| \ .$$

Having obtained extrinsic information $L_{\text{ext},\,k'}\big(c_{k,j}\big)$ and $L_{\text{ext-p}}\big(c_{k,j}\big)$, one can obtain by the following substraction,

$$L_{\text{ext-d},\,k'}\big(c_{k,j}\big) = L_{\text{ext},\,k'}\big(c_{k,j}\big) - L_{\text{ext-p}}\big(c_{k,j}\big) \ , \tag{16}$$

the extrinsic information of bit $c_{k,j}$ extracted solely from the channel observation $\mathbf{y}_{d,k'}$ and the a priori information of $\mathbf{c}_{k'}$. In contrast to the situation of perfect channel state information at the receiver (CSIR) where $L_{\text{ext}}(c_{k,j})$ only depends on the a priori knowledge of $\mathbf{c}_k$ and observation $\mathbf{y}_{d,k}$, a non-zero extrinsic information of $c_{k,j}$ can be obtained from the a priori knowledge of $\mathbf{c}_{k'}$ and observation $\mathbf{y}_{k'}$ (with $k' \neq k$) in an unknown MIMO fading environment. An intuitive explanation of the above difference can be made by viewing $\mathbf{c}_{k'}$ as partially fixed pilots based on the input a priori information. Therefore, better channel knowledge is learned (although no explicit channel estimation exists), which translates into a better a posterior probability of $c_{k,j}$. Hence, a non-zero partial extrinsic information solely from the a priori probability of $\mathbf{c}_{k'}$ and the channel observation $\mathbf{y}_{k'}$ is obtained.

Due to the assumption that the input a priori information of different bits are independent, all the partial extrinsic information $L_{\text{ext-d},k'}(c_{k,j})$ and $L_{\text{ext-p}}(c_{k,j})$ can be viewed as being close to independent. As illustrated in (the left side of) Fig. 4, the final output extrinsic information $L_{\text{ext}}(c_{k,j})$ is obtained by summing all the independent partial extrinsic information obtained from different coded rows $\mathbf{c}_{k'}$ and pilot observations, i.e.

$$L_{\text{ext}}\big(c_{k,j}\big) = \sum_{\substack{k'=1\\k'\neq k}}^{T_d} L_{\text{ext-d},\,k'}\big(c_{k,j}\big) + L_{\text{ext-p}}\big(c_{k,j}\big) = \sum_{\substack{k'=1\\k'\neq k}}^{T_d} L_{\text{ext},\,k'}\big(c_{k,j}\big) - (T_d - 2)\cdot L_{\text{ext-p}}\big(c_{k,j}\big) \ , \tag{17}$$

where

$$1 \leq k \leq T_d, \quad 1 \leq j \leq M \log_2 |\mathcal{X}| \ .$$

A summation of $2^{2M\log_2|\mathcal{X}|}$ terms is required to extract the partial extrinsic information $L_{\text{ext},k'}(c_{k,j})$ in equation (13) and $2^{M\log_2|\mathcal{X}|}$ terms for $L_{\text{ext-p}}(c_{k,j})$ in equation (15). Therefore, in order to obtain the output soft extrinsic LLR values, a total number of $\big((T_d-1)\cdot 2^{2M\log_2|\mathcal{X}|}+2^{M\log_2|\mathcal{X}|}\big)$ terms of probability summation is required for each coded bit, as opposed to $2^{T_d M\log_2|\mathcal{X}|}$ terms in the original optimal soft MIMO detector. Furthermore, the proposed sub-optimal soft MIMO detection algorithm can be easily generalized by extracting partial extrinsic information through combining more than two (E in general) rows of the sub-codeword $\mathbf{C}$ together. By choosing different combination size of $2 \leq E \leq T_d$, a group of sub-optimal MIMO detectors can be constructed which offer a varying degree of detection complexity to system performance tradeoff.

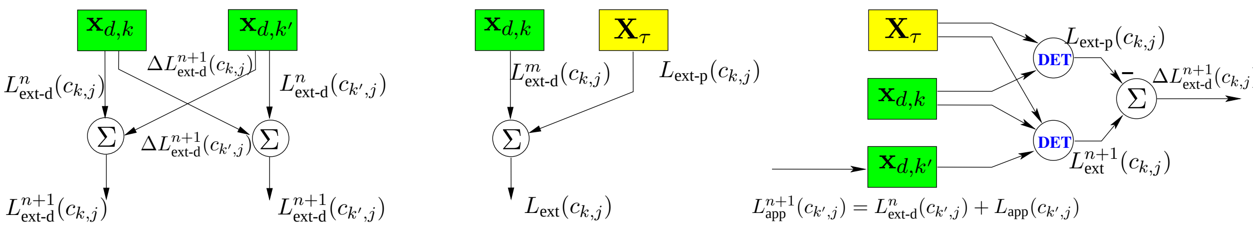

Fig. 5. Sub-optimal soft MIMO detector using butterfly structure

C. Sub-optimal butterfly soft MIMO detector

Motivated by the fast Fourier transform (FFT) algorithm, we can further reduce the complexity of the soft MIMO detector to $\left(\log_2 T_d \cdot 2^{2M \log_2 |\mathcal{X}|} + 2^{M \log_2 |\mathcal{X}|}\right)$ terms of summation per coded bit by using a sub-optimal butterfly MIMO detector structure as illustrated in Fig. 5. It is first assumed that the number of the data slots $T_d = 2^m$, a power of 2. If not, we can appropriately zero-pad the transmitted signal matrix $\mathbf{X}$. As demonstrated in (the left part of) Fig. 5, the sub-optimal butterfly detection algorithm obtains the extrinsic information through a multi-level structure similar to the fast Fourier transform, where the extrinsic information is accumulated from level to level. Specifically, if the partial extrinsic LLR value of coded bit $c_{k,j}$ at the n^{th} level is $L^{n}_{\text{ext-d}}(c_{k,j})$, then the extrinsic LLR value of the $(n+1)^{\text{th}}$ level can be updated by the following form, which is illustrated in (the right part of) of Fig. 5,

$$L^{n+1}_{\text{ext-d}}(c_{k,j}) = L^{n}_{\text{ext-d}}(c_{k,j}) + \Delta L^{n+1}_{\text{ext-d}}(c_{k,j}), \qquad 0 \le n \le m-1 \ , \tag{18}$$

where the second term $\Delta L^{n+1}_{\text{ext-d}}(c_{k,j})$ of equation (18) represents the additional partial extrinsic information obtained from the information of coded bits $\mathbf{c}_{k'}$, with sub-codeword row index k' given by

$$k' = \begin{cases} k + 2^{m-n-1} & \text{if} \quad k \pmod{2^{m-n}} < 2^{m-n-1} \\ k - 2^{m-n-1} & \text{if} \quad k \pmod{2^{m-n}} \ge 2^{m-n-1} \end{cases} . \tag{19}$$

Similar to the extraction algorithm provided in (16), $\Delta L^{n+1}_{\text{ext-d}}(c_{k,j})$ is given by the following form

$$\Delta L^{n+1}_{\text{ext-d}}(c_{k,j}) = L^{n+1}_{\text{ext}}(c_{k,j}) - L_{\text{ext-p}}(c_{k,j}) \ , \tag{20}$$

where $L_{\text{ext-p}}(c_{k,j})$ is given by equation (15), and partial extrinsic information $L^{n+1}_{\text{ext}}(c_{k,j})$ is obtained by performing optimal soft MIMO detection on the sub-coherent block $\mathbf{X}_{[k,k']}$ and $\mathbf{Y}_{[k,k']}$ with modified input a priori information, i.e.

$$L^{n+1}_{\text{ext}}(c_{k,j}) = \log\left(\frac{\sum_{\mathbf{X}_{[k,k']}\in\mathcal{D}^{+}_{k,j}} p\left(\mathbf{Y}_{[k,k']}\middle|\mathbf{X}_{[k,k']}\right)\cdot p^{n+1}_{\text{app}}\left(\mathbf{X}_{[k,k']}\right)}{\sum_{\mathbf{X}_{[k,k']}\in\mathcal{D}^{-}_{k,j}} p\left(\mathbf{Y}_{[k,k']}\middle|\mathbf{X}_{[k,k']}\right)\cdot \ p^{n+1}_{\text{app}}\left(\mathbf{X}_{[k,k']}\right)}\right) - \log\left(\frac{p(c_{k,j}=1)}{p(c_{k,j}=0)}\right) \ . \tag{21}$$

Furthermore, the modified a priori probability $p^{n+1}_{\text{app}}(\mathbf{X}_{[k,k']})$ in equation (21) is a combination of the a priori probability of $\mathbf{c}_k$ and $\mathbf{c}_{k'}$ as well as the n^{th} level extrinsic information of $\mathbf{c}_{k'}$, which can be represented as

$$p_{\text{app}}^{n+1}\big(\mathbf{X}_{[k,k']}\big) = p\big(\mathbf{c}_k\big) \cdot p\big(\mathbf{c}_{k'}\big) \cdot p_{\text{ext}}^{n}\big(\mathbf{c}_{k'}\big) = \prod_{j=1}^{M \log_2 |\mathcal{X}|} p\big(c_{k,j}\big) \cdot p\big(c_{k',j}\big) \cdot p_{\text{ext}}^{n}\big(c_{k',j}\big) \ , \tag{22}$$

where $p_{\text{ext}}^{n}\big(c_{k',j}\big)$ is given by

$$p_{\text{ext}}^{n}\big(c_{k',j}\big) \ = \ \frac{\exp\Big(c_{k',j} \cdot L_{\text{ext-d}}^{n}\big(c_{k',j}\big)\Big)}{1 + \exp\Big(L_{\text{ext-d}}^{n}\big(c_{k',j}\big)\Big)} \ . \tag{23}$$

Therefore, $\Delta L_{\text{ext-d}}^{n+1}\big(c_{k,j}\big)$ can be viewed as the partial extrinsic information obtained solely from the a priori information of $\mathbf{c}_{k'}$, channel observation $\mathbf{y}_{k'}$, and its extrinsic information at the n^{th} level.

Starting from the initial condition $L_{\text{ext-d}}^{0}(c_{k,j}) = 0$, the extrinsic information $L_{\text{ext-d}}^{n}(c_{k,j})$ of each coded bit is accumulated at each level by absorbing additional partial extrinsic information through the sub-coherent block combining process. As illustrated in (the middle part of) Fig. 5, the final soft extrinsic LLR value of each coded bit is formed by combining the extrinsic LLR information at the m^{th} (lowest) level with the extrinsic information obtained from pilot observations, which is given by

$$L_{\text{ext}}\big(c_{k,j}\big) = L_{\text{ext-d}}^{m}\big(c_{k,j}\big) + L_{\text{ext-p}}\big(c_{k,j}\big) \qquad 1 \le k \le T_d, \qquad 1 \le j \le M \log_2 |\mathcal{X}| \ . \tag{24}$$

Note that both the sub-optimal structure in Section III-B as well as the sub-optimal butterfly MIMO detector in the previous subsection are modifications of the optimal soft MIMO detection algorithm provided in Section III-A. The two sub-optimal MIMO detection algorithms provided in Section III-B and III-C have the following structural differences. First, the sub-optimal MIMO detector in Section III-B forms extrinsic information through a *linear* combining structure, where there are a total of $(T_d - 1)$ partial extrinsic information terms (each corresponding to the partial extrinsic LLR obtained from other rows k'); each term is computed by performing optimal detection on the sub-coherent block given by (13)-(16). On the other hand, the sub-optimal butterfly MIMO detector in Section III-C performs data detection by employing a multi-level structure, where the extrinsic information is distributed at succeeding levels until all the input a priori information and the channel observations are combined and exchanged between all different rows.

D. Modified EM-based MIMO detector

The soft MIMO detector and its two sub-optimal modifications proposed in previous sections perform data detection without forming any specific channel estimate. However, forming a channel estimation followed by coherent MIMO detection is in some cases a promising alternative especially when there are enough training pilots. Besides, the estimated channel state information $\widehat{\mathbf{H}}$ can be easily fedback to the transmitter for better power allocation and spectral shaping of the channel coding.

Recently, a lot of attention has been focused on turbo MAP EM estimators, which can take into account not only the training pilots but also the a priori information of the coded bits from the outer soft LDPC decoder. As reported in [23] [24], the proposed turbo EM estimator provides better performance than the conventional MMSE-based channel estimator and works well in an iterative decoding algorithm, especially when T_d is large. However, there exists positive feedback between the input and output soft LLR values which can cause severe performance degradation of the coded MIMO system. Therefore, we propose in this section a modified EM-based MIMO detector that avoids positive feedback and results in better performance than the direct EM-based detection algorithm. Mutual information transfer characteristic of the modified EM-based detector as well as the corresponding simulation results provided in Section IV and V further confirm our claims of superiority of the new detector.

To start with the detection algorithm, let us first look at the conventional MAP EM estimator, whose objective is to find the channel estimation $\widehat{\mathbf{H}}$ that maximizes a posterior probability

$$\widehat{\mathbf{H}} = \arg \max_{\mathbf{H}} \ p\big(\mathbf{H}\big|\mathbf{Y}\big) = \arg \max_{\mathbf{H}} \ p\big(\mathbf{Y}\,, \mathbf{H}\big) \ \ , \tag{25}$$

which is intractable by direct maximization. Hence by taking the transmitted data signal matrix $\mathbf{X}_d$ (or $\mathbf{X}$) as the unobserved (or missing) data, the following iterative expectation maximization (EM) algorithm (similar to [23]) is applied .

- **E-step:**

$$Q\big(\mathbf{H}\big|\widehat{\mathbf{H}}^{(n)}\big) = E_{\mathbf{X}\big|\widehat{\mathbf{H}}^{(n)},\mathbf{Y}}\Big[-\log p\big(\mathbf{H},\mathbf{Y}\big|\mathbf{X}\big)\Big] \ . \tag{26}$$

After some manipulations, we have the following concise form

$$\begin{aligned} Q\big(\mathbf{H}\big|\widehat{\mathbf{H}}^{(n)}\big) &= E_{\mathbf{X}\big|\widehat{\mathbf{H}}^{(n)},\mathbf{Y}}\Big[\mathbf{tr}\Big(\mathbf{H}^H\mathbf{H} + \big(\mathbf{Y}-\mathbf{X}\mathbf{H}\big)^H\big(\mathbf{Y}-\mathbf{X}\mathbf{H}\big)\Big)\Big] \\ &= \mathbf{tr}\Big(\mathbf{H}^H\mathbf{R}\ \mathbf{H} + \mathbf{Y}^H\mathbf{Y} - \big(\mathbf{Y}^H\mathbf{U}\mathbf{H} + \mathbf{H}^H\mathbf{U}^H\mathbf{Y}\big)\Big) \ \ , \end{aligned} \tag{27}$$

where $\mathbf{R}$ is given by

$$\begin{aligned} \mathbf{R} &= E_{\mathbf{X}\big|\widehat{\mathbf{H}}^{(n)},\mathbf{Y}}\big[\mathbf{X}^H\mathbf{X}\big] + I_M \\ &= \frac{\rho}{M}\cdot\Big(\sum_{j=1}^{T_d}\sum_{\mathbf{x}_{d,k}\in\mathcal{X}^M} p\big(\mathbf{x}_{d,k}\big|\widehat{\mathbf{H}}^{(n)},\mathbf{y}_{d,k}\big)\cdot\mathbf{x}_{d,k}^H\,\mathbf{x}_{d,k} + \mathbf{X}_\tau^H\mathbf{X}_\tau\Big) + I_M \ \ , \end{aligned} \tag{28}$$

and $\mathbf{U}$ is given by

$$\mathbf{U} = E_{\mathbf{X}\big|\widehat{\mathbf{H}}^{(n)},\mathbf{Y}}\big[\mathbf{X}\big] = \sqrt{\frac{\rho}{M}}\cdot\Big[\mathbf{X}_\tau^T,\, \boldsymbol{\mu}_1^T, \cdots, \boldsymbol{\mu}_{T_d}^T\Big]^T, \quad \boldsymbol{\mu}_k = \sum_{\mathbf{x}_{d,k}\in\mathcal{X}^M} p\big(\mathbf{x}_{d,k}\big|\widehat{\mathbf{H}}^{(n)},\mathbf{y}_{d,k}\big)\cdot\mathbf{x}_{d,k} \ \ . \tag{29}$$

The a posterior probability $p\big(\mathbf{x}_{d,k}\big|\widehat{\mathbf{H}}^{(n)},\mathbf{y}_{d,k}\big)$ is given by

$$p\big(\mathbf{x}_{d,k}\big|\widehat{\mathbf{H}}^{(n)},\mathbf{y}_{d,k}\big) = \frac{p\big(\mathbf{y}_{d,k}\big|\widehat{\mathbf{H}}^{(n)},\mathbf{x}_{d,k}\big)\cdot p\big(\mathbf{x}_{d,k}\big)}{\sum_{\mathbf{x}_{d,k}\in\mathcal{X}^M} p\big(\mathbf{y}_{d,k}\big|\widehat{\mathbf{H}}^{(n)},\mathbf{x}_{d,k}\big)\cdot p\big(\mathbf{x}_{d,k}\big)} \ \ , \tag{30}$$

with $p\big(\mathbf{y}_{d,k}\big|\widehat{\mathbf{H}}^{(n)},\mathbf{x}_{d,k}\big)$ and $p\big(\mathbf{x}_{d,k}\big)$ given as

$$p\big(\mathbf{y}_{d,k}\big|\widehat{\mathbf{H}}^{(n)},\mathbf{x}_{d,k}\big)=\frac{1}{\pi^N}\exp\Big(-\big\|\mathbf{y}_{d,k}-\sqrt{\frac{\rho}{M}}\cdot\widehat{\mathbf{H}}^{(n)}\mathbf{x}_{d,k})\big\|^2\Big)\ , \tag{31}$$

where $p\big(\mathbf{x}_{d,k}\big)$ is given by

$$p\big(\mathbf{x}_{d,k}\big)=\prod_{j=1}^{M\log_2|\mathcal{X}|}p\big(c_{k,j}\big)\ . \tag{32}$$

- **M-step:**

$$\widehat{\mathbf{H}}^{(n+1)}=\arg\ \min_{\mathbf{H}}\ Q\big(\mathbf{H}\big|\widehat{\mathbf{H}}^{(n)}\big)\ . \tag{33}$$

  After some manipulations, the updated channel estimation $\widehat{\mathbf{H}}^{(n+1)}$ is obtained as

$$\widehat{\mathbf{H}}^{(n+1)}=\mathbf{R}^{-1}\cdot\mathbf{U}^H\cdot\mathbf{Y}\ . \tag{34}$$

- **Initialization:**

  We use the conventional minimal mean square error (MMSE) channel estimator for initialization, which is given by

$$\widehat{\mathbf{H}}^{(0)}=\sqrt{\frac{\rho}{M}}\cdot\mathbf{X}_\tau^H\cdot\big(\frac{\rho}{M}\mathbf{X}_\tau\mathbf{X}_\tau^H+I_{T_\tau}\big)^{-1}\cdot\mathbf{Y}_\tau\ . \tag{35}$$

  Since the EM iteration is embedded within the large iterative decoding loop of the soft MIMO receiver, we can also take the estimated channel $\mathbf{H}$ from the last decoding iteration as an EM initialization. Compared with the simple MMSE estimator, the obtained estimation from the last decoding iteration (through an EM algorithm) provides a better initialization since additional a priori information of the coded bits is used. Therefore, by using the alternative initialization, EM algorithm is able to begin at a better starting point and hence results in smaller number of EM iterations.

Maximum a posterior channel estimation is obtained when the MAP EM algorithm converge to $\widehat{\mathbf{H}}$ after certain number of iterations. Hence the soft extrinsic information of each coded bit $c_{k,j}$ is provided by taking $\widehat{\mathbf{H}}$ as the true channel coefficients followed by coherent MIMO detection,

$$L_{\text{ext}}\big(c_{k,j}\big)=\log\left(\frac{\Sigma_{\mathbf{x}_{d,k}\in\mathcal{D}^+_{k,j}}\ p\Big(\mathbf{y}_{d,k}\Big|\widehat{\mathbf{H}},\mathbf{x}_{d,k}\Big)\cdot p\big(\mathbf{x}_{d,k}\big)}{\Sigma_{\mathbf{x}_{d,k}\in\mathcal{D}^-_{k,j}}\ p\Big(\mathbf{y}_{d,k}\Big|\widehat{\mathbf{H}},\mathbf{x}_{d,k}\Big)\cdot p\big(\mathbf{x}_{d,k}\big)}\right)-\log\left(\frac{p\big(c_{k,j}=1\big)}{p\big(c_{k,j}=0\big)}\right)\ , \tag{36}$$

where

$$1\le k\le T_d,\qquad 1\le j\le M\log_2|\mathcal{X}|\ .$$

$\mathcal{D}^+_{k,j}$ ($\mathcal{D}^-_{k,j}$) is the set of $\mathbf{x}_{d,k}$ for which bit $c_{k,j}$ is "$+1$" ("-1"), and probabilities $p\Big(\mathbf{y}_{d,k}\Big|\widehat{\mathbf{H}},\mathbf{x}_{d,k}\Big)$ and $p\big(\mathbf{x}_{d,k}\big)$ are given by (31) and (32) respectively.

It is well known that short girth in the LDPC Tanner graph is one of the major performance bottleneck for short length LDPC code design [32] [33], where positive feedback of the iterative LLR values generated by the existing short length loops directly affects the iterative message passing algorithm. Similarly, positive feedback caused by the correlations between input and output extrinsic information of the MIMO detector will also cause severe system performance degradation. Therefore, considerable effort has been made in various detection algorithms to avoid the same information from counting twice, or to avoid the output extrinsic LLR values from containing any input a priori information.

Unfortunately, if we study the conventional direct EM-based soft MIMO detection algorithm carefully, we will find that the estimated channel coefficient does depend on the a priori information of the entire sub-codeword $\mathbf{C}$. To be specific, channel estimation $\widehat{\mathbf{H}}$ can be represented as a function given by

$$\widehat{\mathbf{H}} = \widehat{\mathbf{H}}\Big(\mathbf{Y}, \mathbf{A}_{\text{L-app}}\Big) \ , \tag{37}$$

where $\mathbf{A}_{\text{L-app}} \in \mathbb{R}^{T_d \times M \log_2 |\mathcal{X}|}$ is the a priori information matrix with each element $a_{k,j}$ equal to the a priori LLR value $L_{\text{app}}(c_{k,j})$. Therefore, the extrinsic information obtained by equation (36) contains the input a priori information through $\widehat{\mathbf{H}}$, in a sense that $L_{\text{ext}}(c_{k,j})$ depends on $L_{\text{app}}(c_{k,j})$, even though the a priori LLR value is already subtracted from the log a posterior value as demonstrated by the second term. In order to eliminate input-output correlations introduced by the direct channel estimation $\widehat{\mathbf{H}}$, which is a function of $L_{\text{app}}\big(c_{k,j}\big)$, we propose a modified EM channel estimation algorithm that uses only part of the a priori information (a subset of matrix $\mathbf{A}_{\text{L-app}}$) of the sub-codeword $\mathbf{C}$. If we denote $\mathcal{E}$ as a subset of $\{1, 2, \cdots, T_d\}$ that includes k, the partial a priori information matrix can be formed by the following weighting operation

$$\mathbf{A}^{\mathcal{E}}_{\text{L-app}} = \mathbf{diag}\big(\mathbf{s}\big) \cdot \mathbf{A}_{\text{L-app}} \ , \tag{38}$$

where the selecting vector $\mathbf{s}$ of size $1 \times T_d$ is given by

$$\mathbf{s} = \big[s_1, s_2, \cdots, s_{T_d}\big], \qquad s_j = \begin{cases} 1 & \text{if } j \notin \mathcal{E} \\ 0 & \text{if } j \in \mathcal{E} \end{cases} \ . \tag{39}$$

The modified channel estimation is hence obtained by applying the same APP EM algorithm by using $\mathbf{A}^{\mathcal{E}}_{\text{L-app}}$ as the input a priori information matrix instead, i.e.

$$\widehat{\mathbf{H}}_{\mathcal{E}} = \widehat{\mathbf{H}}\Big(\mathbf{Y}, \mathbf{A}^{\mathcal{E}}_{\text{L-app}}\Big) \ . \tag{40}$$

The modified estimation $\widehat{\mathbf{H}}_{\mathcal{E}}$ can therefore be used to perform coherent detections for coded rows $\mathbf{c}_k$ with index $k \in \mathcal{E}$,

$$L_{\text{ext}}\big(c_{k,j}\big) = \log\left(\frac{\Sigma_{\mathbf{x}_{d,k} \in \mathcal{D}^+_{k,j}} \ p\Big(\mathbf{y}_{d,k} \big| \widehat{\mathbf{H}}_{\mathcal{E}}, \mathbf{x}_{d,k}\Big) \cdot p\big(\mathbf{x}_{d,k}\big)}{\Sigma_{\mathbf{x}_{d,k} \in \mathcal{D}^-_{k,j}} \ p\Big(\mathbf{y}_{d,k} \big| \widehat{\mathbf{H}}_{\mathcal{E}}, \mathbf{x}_{d,k}\Big) \cdot p\big(\mathbf{x}_{d,k}\big)}\right) - \log\left(\frac{p\big(c_{k,j}=1\big)}{p\big(c_{k,j}=0\big)}\right), \quad k \in \mathcal{E} \ . \tag{41}$$

Let us further assume that the entire set $\{1, 2, \cdots, T_d\}$ can be decomposed into the following disjoint sets with the same size, i.e.

$$\{1, 2, \cdots, T_d\} = \bigcup_n \mathcal{E}_n, \qquad \mathcal{E}_n \bigcap \mathcal{E}_{n'} = \phi, \qquad \left|\mathcal{E}_n\right| = S_E \ . \tag{42}$$

Instead of having only one EM estimation in the direct EM-based detector, $\lceil T_d/S_E \rceil$ separate EM estimations are to be completed during one entire soft decoding iteration in the modified EM-based detector. Note that $\mathcal{E}_n = \{n\}$ and $\mathcal{E}_n = \phi$ correspond to special cases: $\mathcal{E}_n = \{n\}$ has the maximum detection complexity, but takes into account all available a priori information from the outer soft decoder, while on the other hand $\mathcal{E}_n = \phi$ corresponds to the case of conventional direct EM-based detection algorithm. For a short complexity analysis, we know that within each EM estimation, there are total $N_{sum}^{EM} = \bar{I} \cdot \left(3T_d \cdot 2^{M \log_2 |\mathcal{X}|}\right)$ summation operations, where $\bar{I}$ is the average number of iterations required by the convergence of the EM algorithm. Therefore, the average number of the summations for each coded bit in the modified EM-based MIMO detector is

$$N_{sum}^{DEC} = \left(3\bar{I} \cdot \lceil T_d/S_E \rceil + 1\right) \cdot 2^{M \log_2 |\mathcal{X}|} \ . \tag{43}$$

The EM channel estimation algorithms proposed in this section can make full use of the soft a priori information of the coded bits from the outer LDPC decoder, and hence provide better (and more accurate) channel estimations. From another point of view, the MAP EM estimator is generally equivalent to extending the pilot structure to the entire transmitted signal matrix $\mathbf{X}$. Instead of limiting the pilots to $\mathbf{X}_\tau$, the receiver treats $\mathbf{X}_d$ as partially fixed pilots as well especially when the LLR ratios are getting significantly improved as a result of the messages being updated constantly through the iterations.

Finally, a brief comparison of the pilots size required by the different MIMO detection algorithms is as follow. First, we note that the proposed optimal soft MIMO detector as well as its two sub-optimal modifications are able to provide soft data detections with arbitrary number of pilot symbols and only need a small number of pilots in order to remove detection ambiguity (in the first decoding iteration). The modified EM-based detector only requires a small number ($T_\tau \geq M$) of pilots for the initialization of the EM estimation. Therefore, these four soft MIMO detection algorithms provide a wide range of trade-offs between complexity and performance and can work in different MIMO fading environments and support various training sizes.

IV. Design of LDPC-coded MIMO Systems

Conventionally the coded MIMO receiver is obtained by connecting the inner soft MIMO detector and the outer LDPC decoder to form one large iterative decoding loop. As evident from Fig. 6, the overall MIMO receiver actually consists of two iterative decoding loops. In the outer loop, the soft MIMO detector forms extrinsic information of each coded bit $\{\mathbf{C}_i\}_{i=1}^{L}$ based on the received signal $\{\mathbf{Y}_i\}_{i=1}^{L}$ as well as the input a priori knowledge from the LDPC decoder, and serves as the input a priori information for the LDPC decoder in the next iteration. The soft LDPC decoder has an inner iterative decoding loop that is composed of a variable node decoder, a check node decoder, and two connecting edge interleavers. The soft extrinsic information, which describes the uncertainty of each coded bit, is iteratively exchanged in the outer loop between the MIMO detector and LDPC decoder as well as in the inner loop between variable node and check node decoders inside the LDPC decoder.

A. Receiver structure of the LDPC-coded MIMO systems

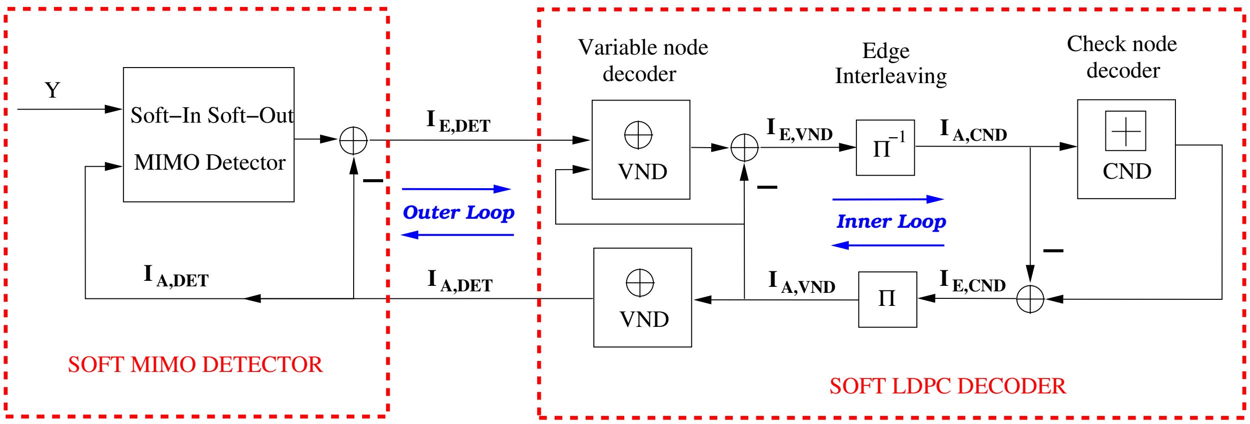

Fig. 6. Conventional receiver structure of LDPC-coded MIMO systems

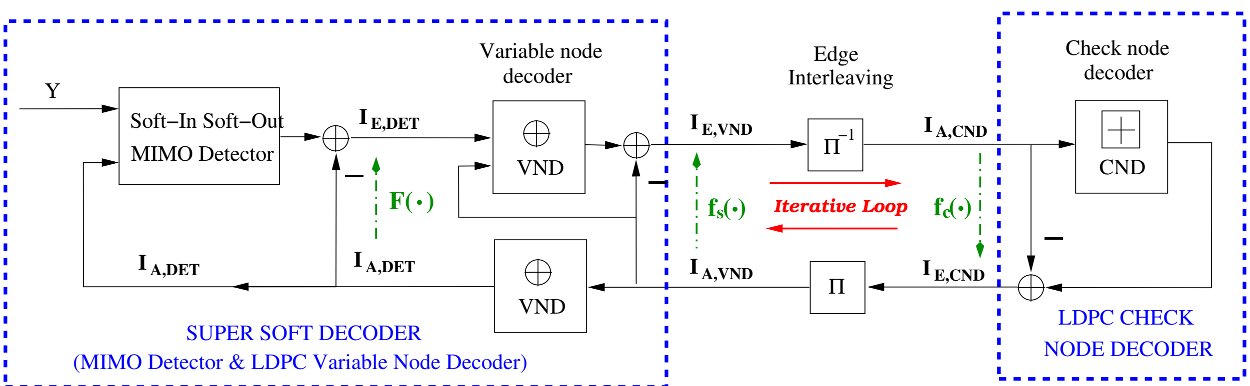

Fig. 7. New receiver structure of LDPC-coded MIMO systems

In this paper, we structure the MIMO receiver differently by combining the soft MIMO detector and LDPC variable node decoder as a super soft-decoder, a form also utilized in [27]. As illustrated in Fig. 7, the decoding loop is formed by exchanging extrinsic information between the super decoder and the

LDPC check node decoder iteratively. Compared with the conventional iterative MIMO receiver (named as bit-interleaved coded modulation with iterative decoding (BICM-ID) algorithm) shown in Fig. 6, the new receiver structure has two advantages. First, the proposed receiver structure has only one iterative decoding loop and hence achieves smaller decoding complexity compared to the two iterative loops (inner LDPC decoder loop and outer “MIMO detector $\rightleftarrows$ LDPC decoder” loop) in the conventional BICM-ID structure. Second, the proposed structure has the advantage of enabling the extrinsic information transfer characteristic function of the soft component decoders to have tractable forms. By fully exploiting the closed form EXIT functions, a simple and efficient LDPC code degree profile optimization algorithm with proven global optimality and guaranteed convergence is proposed in Section IV-C, which is superior to the sub-optimal manual curve fitting technique [25] [27].

B. Analysis of extrinsic information transfer characteristics

In order to understand as well as design the iterative decoding systems having bipartite graph structures, we use the extrinsic information transfer characteristic of the soft MIMO detector and LDPC decoder, which was proposed by Brink in [25], to analyze the convergence behavior of the iterative decoding schemes of the coded MIMO system.

1) Brief introduction on EXIT-chart

We briefly describe in this section the EXIT-chart technique proposed in [25]. For readers who are familiar with the topic, please skip to Section IV-B-2 directly. The extrinsic information transfer (EXIT) function is used to the describe the input-output (a priori information versus extrinsic information) relations of the soft component decoders from an information theoretical perspective. Taking the component soft decoders in Fig. 7 as an example, the corresponding EXIT functions of the super soft-decoder and LDPC check node decoder can be described by the following mapping (also depicted in Fig. 7 accordingly),

$$I_{\text{E,VND}} = f_s\big(I_{\text{A,VND}}\big)\,, \qquad I_{\text{E,CND}} = f_c\big(I_{\text{A,CND}}\big)\,, \tag{44}$$

where $I_{\text{A,VND}}$ represents the mutual information between the coded bit x and the input a priori information of the super soft-decoder, and $I_{\text{E,VND}}$, $I_{\text{E,CND}}$, as well as $I_{\text{A,CND}}$ are similarly defined. According to the iterative decoding structure, where the output extrinsic information from one component decoder is treated as a priori input to the other one, the mutual information between the extrinsic LLR values and the coded bits is updated through the following evolution,

$$I^{k}_{\text{E,CND}} = f_c \circ f_s\Big(I^{k-1}_{\text{E,CND}}\Big)\,, \qquad f_c \circ f_s(\cdot) = f_c\big(f_s(\cdot)\big)\ , \tag{45}$$

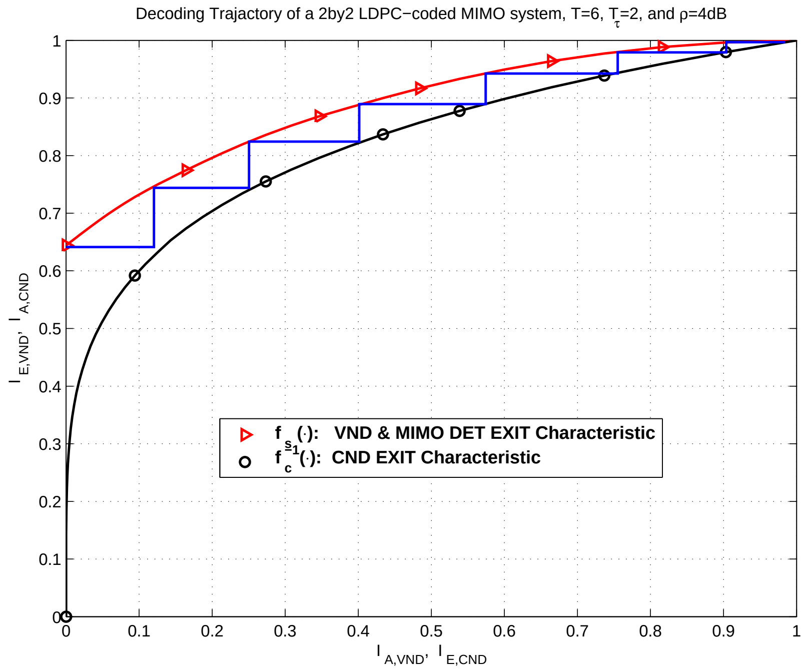

Fig. 8. Decoding trajectory of a regular $(3, 6)$ LDPC-coded MIMO system with optimal soft MIMO detector over a 2×2 unknown MIMO channel with coherence time $T = 6$, training length $T_\tau = 2$, and signal to noise ratio $\rho = 4dB$.

with index k indicating the k^{th} decoding iteration and the initialization is given by $I^0_{\text{E,CND}} = 0$. As an example, we demonstrate in Fig. 8 the EXIT functions f_s and f_c (with x and y axis flipped) of the component soft decoders as well as the decoding trajectory of an LDPC-coded MIMO system. The 2×2 MIMO system considered has BPSK modulation, uses optimal soft MIMO detector at the receiver, and transmits over a fading channel with coherence time $T = 6$, training number $T_\tau = 2$, and signal to noise ratio $\rho = 4dB$. The outer LDPC code is a regular $(3, 6)$ code with codeword length 8×10^4. We can observe from Fig. 8 that, as long as the EXIT chart curve f_s is above curve f_c^{-1} (with the x and y axis flipped), i.e.

$$f_s(x) \geq f_c^{-1}(x), \qquad 0 \leq x \leq 1 \ , \tag{46}$$

the decoding trajectory is able to make its zigzag way until reaching the successful decoding point $(1, 1)$. Therefore, it can serve as a convergence criterion of the iterative decoding algorithm for the LDPC code design purpose.

2) EXIT characteristic of the soft MIMO detector

According to the results provided in [25] [27], the extrinsic mutual information $I_{\text{E,DET}}$ between the transmitted bit x and the output LLR values $L_{\text{ext}}(x)$, which measures the information contents of the output extrinsic LLR values, can be represented as

$$I_{\text{E,DET}}(\rho;\sigma_A^2) = I\Big(L_{\text{ext}}(x)\,;x\Big) = \frac{1}{2}\sum_{x=\pm 1}\int_{-\infty}^{\infty}\log_2\Big(\frac{2p_{_E}(\xi|x)}{p_{_E}(\xi|x=+1)+p_{_E}(\xi|x=-1)}\Big)\cdot p_{_E}(\xi|x)\,d\xi, \tag{47}$$

where distribution $p_{_E}(\xi|x=\pm 1)$ is obtained through Monte Carlo simulation (histogram measurements) by setting the system SNR equal to ρ and the input a priori LLR values conditioned on the transmitted bit x have a Gaussian distribution given by

$$L_{\text{app}}(x) = x\cdot n, \qquad x\in\{+1,-1\}, \qquad n\sim\mathcal{N}\Big(\frac{2}{\sigma_A^2}\ .\frac{4}{\sigma_A^2}\Big) \tag{48}$$

Therefore, the extrinsic mutual information $I_{\text{E,DET}}$ depends both on the system SNR ρ and the noise variance level σ_A^2 of the input a priori information. By viewing ρ as an index parameter, the EXIT function of the soft MIMO detector is given by the following form

$$I_{\text{E,DET}} = I_{\text{E,DET}}\Big(\rho;\ \sigma_A^2 = J^{-1}(I_{\text{A,DET}})\Big) \triangleq F\big|_{\rho}(I_{\text{A,DET}})\ , \tag{49}$$

where function $J(\cdot)$ is given by (equivalent to equation (24) in the Appendix of [27]),

$$J\big(\sigma_A^2\big) = I\Big(I_{\text{app}}(x)\,;\,x\Big) = \frac{1}{\ln 2}\Bigg(\frac{1}{\sigma_A^2} - \int_{-\infty}^{\infty}\frac{\sigma_A}{\sqrt{2\pi}}\cdot\ln\cosh(y)\cdot\exp\Big(-\frac{(\sigma_A^2\cdot y-1)^2}{2\sigma_A^2}\Big)\,dy\Bigg)\ . \tag{50}$$

Furthermore, input mutual information $I_{\text{A,DET}}$ of the soft MIMO detector is related to mutual information $I_{\text{A,VND}}$ through the following equation for a variable node of degree d_v

$$I_{\text{A,DET}} = J\Big(d_v\cdot J^{-1}\big(I_{\text{A,VND}}\big)\Big)\ . \tag{51}$$

3) EXIT characteristics of the LDPC variable node and check node decoders

Following the same reasoning as given in [27], the extrinsic mutual information transfer characteristic of a variable node of degree d_v is given by the following form

$$I_{\text{E,VND}}\Big(I_{\text{A,VND}},\,d_v\Big) = J\Big((d_v-1)\cdot J^{-1}(I_{\text{A,VND}}) + J^{-1}(I_{\text{E,DET}})\Big)\ . \tag{52}$$

According to the duality properties [34] of the EXIT curves between single parity check codes and repetition codes over binary erasure channels, the mutual information transfer characteristic of a degree d_c check node over binary input Gaussian output channels can be well approximated as

$$I_{\text{E,CND}}(I_{\text{A,VND}}) \approx 1 - I_{\text{E,REP}}(1-I_{\text{A,VND}}) = 1 - J\Big((d_c-1)\cdot J^{-1}\big(1-I_{\text{A,VND}}\big)\Big)\ . \tag{53}$$

C. LDPC code optimization

Following the methodology given in [25] [27], the EXIT functions of the super MIMO soft-decoder (combination of the LDPC variable node decoder and soft MIMO detector) can be obtained as

$$\begin{aligned} I_{\text{E,VND}} &= f_s\big(I_{\text{A,VND}}\big) = \sum_{i=1}^{D_v} \lambda_i \cdot I_{\text{E,VND}}\Big(I_{\text{A,VND}}, d_{v,i}\Big) \\ &= \sum_{i=1}^{D_v} \lambda_i \cdot J\Bigg(\big(d_{v,i}-1\big)\cdot J^{-1}\big(I_{\text{A,VND}}\big) + J^{-1}\Bigg(F\big|_{\rho}\Big(J\Big(d_{v,i}\cdot J^{-1}\big(I_{\text{A,VND}}\big)\Big)\Big)\Bigg)\Bigg) \ . \end{aligned} \tag{54}$$

where λ_i is the fraction of the variable nodes having edge degree $d_{v,i}$, and D_v is the number of different variable node degrees. Similarly according to (53), the check nodes of the LDPC code have a transfer characteristic given by the following form

$$I_{\text{E,CND}} = f_c\big(I_{\text{A,CND}}\big) \approx 1 - \sum_{i=1}^{D_c} \rho_i \cdot J\Big(\big(d_{c,i}-1\big)\cdot J^{-1}\big(1-I_{\text{A,CND}}\big)\Big) \ , \tag{55}$$

where ρ_i is the fraction of the check nodes having edge degree $d_{c,i}$, and D_c is the number of different check node degrees.

Following the successful decoding (convergence) criterion provided in [25], the degree profile optimization problem can be reduced to the following maximization problem by taking the LDPC code rate R_{outer} as the objective

$$\max_{\{\lambda_i,\rho_i\}} R_{\text{outer}} = \max_{\{\lambda_i,\rho_i\}} \left(1 - \frac{\sum_{i=1}^{D_c} \rho_i/d_{c,i}}{\sum_{i=1}^{D_v} \lambda_i/d_{v,i}}\right) \ , \tag{56}$$

under linear constraints given by

$$\begin{aligned} & I_{\text{E,VND}}(I_{\text{A,VND}}) \geq I_{\text{A,CND}}(I_{\text{E,CND}}) = I_{\text{A,CND}}(I_{\text{A,VND}}), \\ & \sum_{i=1}^{D_v} \lambda_i = 1, \quad \sum_{i=1}^{D_c} \rho_i = 1, \qquad 0 \leq \lambda_i, \rho_i \leq 1 \ . \end{aligned} \tag{57}$$

Utilizing the closed form EXIT functions of the component soft decoders given by (54) and (55), we propose an efficient LDPC code degree profile optimization algorithm in the following, which is composed of two simple linear optimization steps.

- **Variable node degree profile optimization:**

  For a fixed check node degree profile $\{\rho_i^k\}$ from the k^{th} iteration, the optimal variable node degree profile $\{\lambda_i^{k+1}\}$ is given by

$$\{\lambda_i^{k+1}\} = \arg\ \max_{\{\lambda_i\}} \sum_{i=1}^{D_v} \lambda_i/d_{v,i} \ , \tag{58}$$

  under the constraints

$$\begin{aligned} & f_s\Big(f_c(a_n)\Big) \geq a_n \ , \\ & \sum_{i=1}^{D_v} \lambda_i = 1, \quad 0 \leq \lambda_i \leq 1, \qquad 1 \leq n \leq N, \end{aligned} \tag{59}$$

where $\{a_n | a_n \in [0,1]\}$ is a set of specified constraint points, and N is the total number of constraints on the curve.

- **Check node degree profile optimization:**

  For a fixed variable node degree profile $\{\lambda_i^{k+1}\}$ from the $(k+1)^{\text{th}}$ iteration, the optimal check node degree profile $\{\rho_i^{k+1}\}$ is given by

$$\{\rho_i^{k+1}\} = \arg \min_{\{\rho_i\}} \sum_{i=1}^{D_c} \rho_i / d_{c,i} \ , \tag{60}$$

  under the constraints

$$f_c\Big(f_s(a_n)\Big) \geq a_n \ ,$$
$$\sum_{i=1}^{D_v} \rho_i = 1, \quad 0 \leq \rho_i \leq 1, \qquad 1 \leq n \leq N, \tag{61}$$

  where a_n and N are similarly defined as before.

- **Initialization:**

  In general, we can start with any feasible degree profiles. Based on our experience from numerical simulations, we find that it is always a good choice to start with a regular check node degree d_c.

If we stack the LDPC code degree profile $\{\lambda_i, \rho_i\}$ into a super vector $\eta = [\lambda_1, \cdots, \lambda_{D_v}, \rho_1, \cdots, \rho_{D_c}]^T$. We can see that the objective R_{outer} given in equation (56) is a concave function with respect to η and that all the constraints given in (57) are linear. Hence, the above degree optimization problem has only one unique optimal solution. Due to the non-decreasing property of the proposed iterative maximization algorithm, it is guaranteed to converge to the global maximum solution $\eta^\star$ from any initialization point. Therefore, in contrast to the sub-optimal manual curving fitting technique proposed in [27], the above iterative LDPC optimization algorithm provides much better performance and can serve as an efficient tool for coded MIMO system design.

V. Numerical and Simulation Results

A. *Elimination of positive feedback in EM-based MIMO detectors*

We demonstrate in Fig. 9 the extrinsic information transfer functions of the EM-based and modified EM-based soft MIMO detectors over an unknown 2×2 MIMO channel with coherence time interval $T = 6$ and 18, training length $T_\tau = 2$, and signal to noise ratio $\rho = 4dB$. BPSK modulation is assumed for all the simulation results in this section unless explicitly mentioned. For comparison purpose, the mutual information transfer characteristics of the simple MMSE-based MIMO detector and detector with ideal CSIR are also included in the plot.

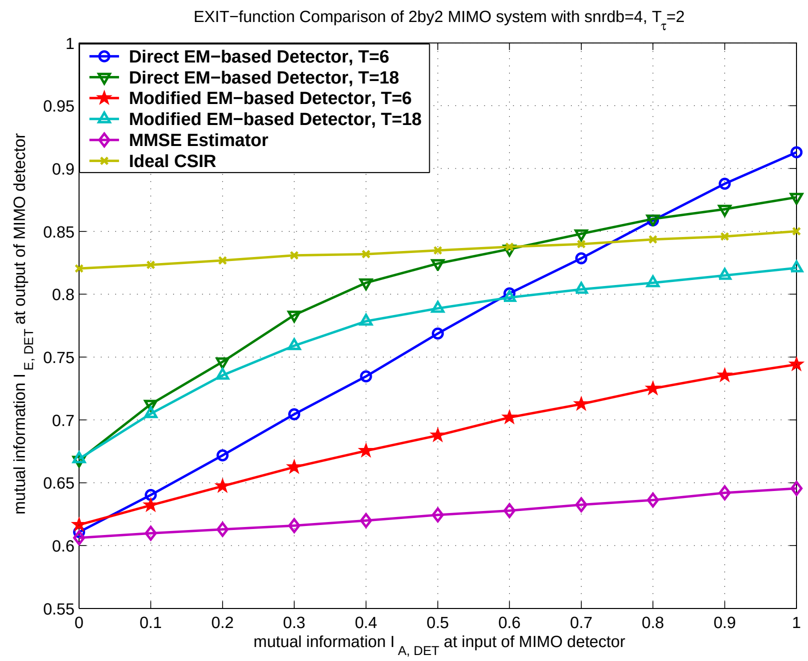

Fig. 9. Extrinsic information transfer characteristic of the EM-based MIMO detectors over a 2×2 unknown MIMO channel with training length $T_\tau = 2$ and signal to noise ratio $\rho = 4dB$.

We can observe from the plot that for direct EM-based MIMO detector, output extrinsic mutual information $I_{\text{E,DET}}$ is even greater than that of the detector with ideal CSIR in high $I_{\text{A,DET}}$ ranges, which directly indicates the existing positive feedback between the input and output extrinsic information. Such strong correlations between the output extrinsic information $L_{\text{ext}}(x)$ and the input a priori information $L_{\text{app}}(x)$ will cause a severe performance degradation as verified by the simulations results provided in Section V-D. As expected, the modified EM-based MIMO detectors successfully eliminate the correlation and achieve significant performance gain compared to the simple MMSE-based detector especially when the coherence time interval T is large.

B. EXIT function comparison between different soft MIMO detectors

As reported in [34], the area below the transfer function $I_{\text{E,DET}} = F(I_{\text{A,DET}})$ well approximates the maximum achievable rate of the outer LDPC encoder. Hence, the extrinsic information transfer characteristic $F(\cdot)$ can be easily used to compare and evaluate the performance of different soft MIMO detectors. We demonstrate in Fig. 10 the extrinsic information transfer functions of different MIMO detectors described in Section III under the same 2×2 unknown MIMO channel with coherence time $T = 6$, training length $T_\tau = 2$, and system SNR $\rho = 4dB$. For comparison purpose again, the mutual information transfer characteristics of the simple MMSE-based MIMO detector and detector with ideal

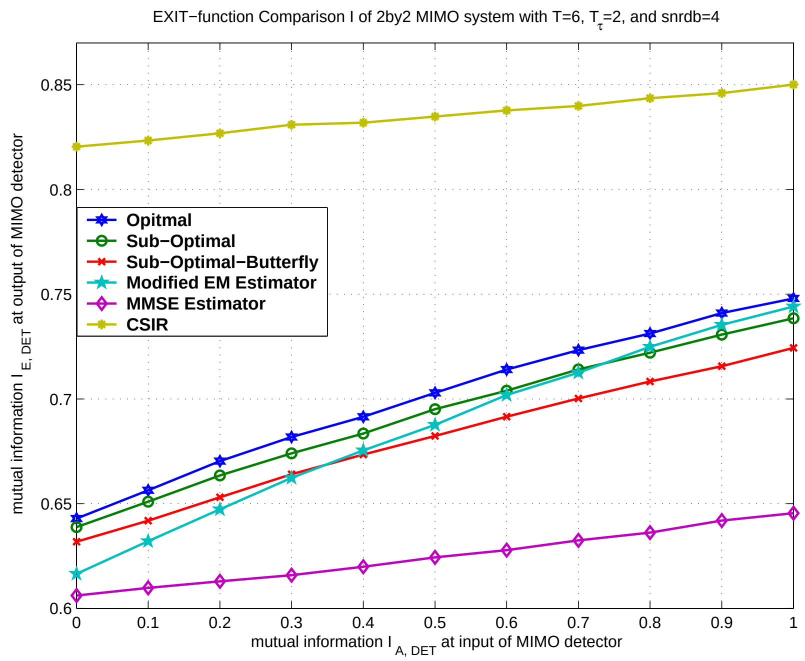

Fig. 10. Comparison of the extrinsic information transfer characteristic of different MIMO detectors in a 2×2 unknown MIMO system with coherence time $T = 6$, training length $T_\tau = 2$, and signal to noise ratio $\rho = 4dB$.

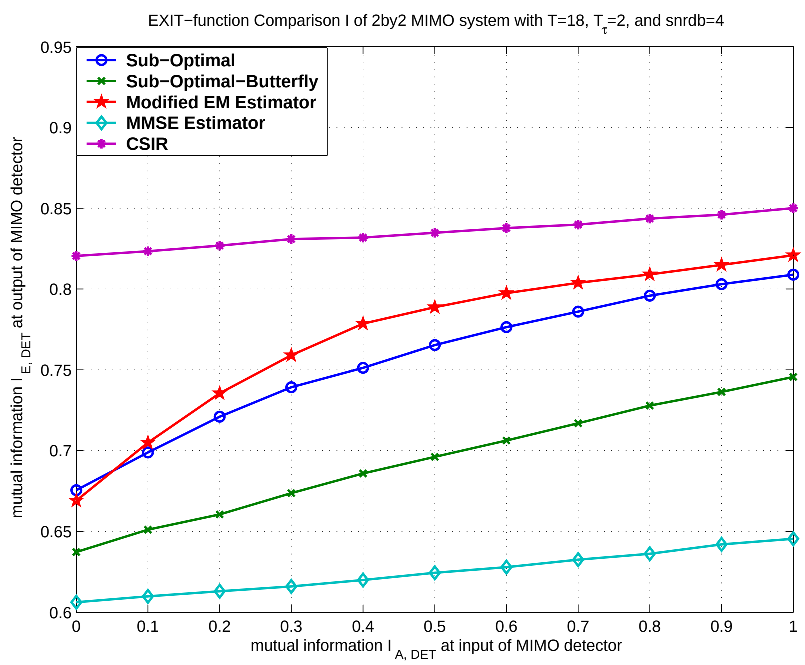

Fig. 11. Comparison of the extrinsic information transfer characteristic of different MIMO detectors in a 2×2 unknown MIMO system with coherence time $T = 18$, training length $T_\tau = 2$, and signal to noise ratio $\rho = 4dB$.

CSIR are also included. As can be observed from the plot, all four soft MIMO detectors have comparable

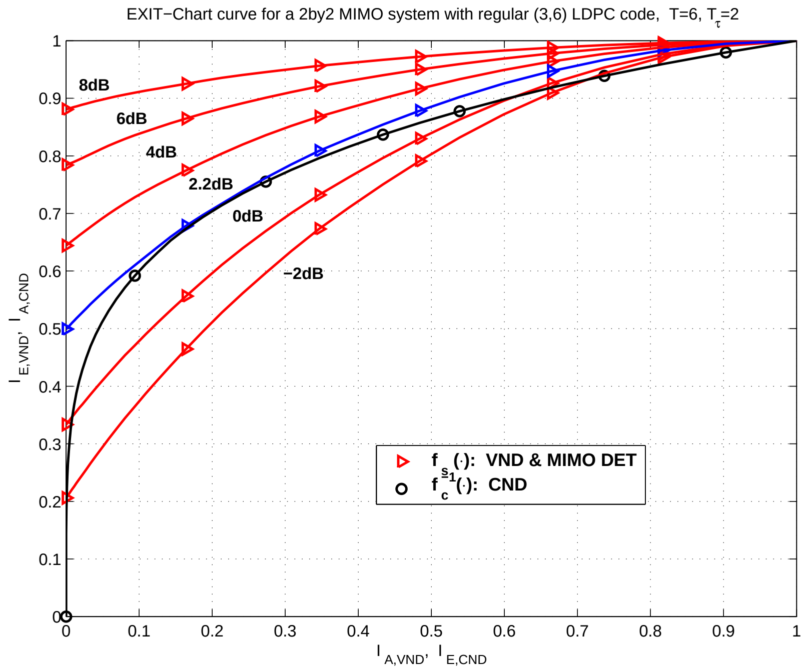

Fig. 12. EXIT-chart curve of a 2×2 regular $(3, 6)$ LDPC-coded MIMO system over a unknown fading channel with coherence time $T = 6$, and training number $T_\tau = 2$ using optimal soft MIMO detectors, under different signal to noise ratios $\rho = -2, 0, 2.2, 4, 6, 8\ dB$.

performance in a small coherence time T channel environment, i.e. $T \leq 10$. All of them achieve significant performance gain over the simple MMSE-based detector but are far away from the MIMO detector with ideal CSIR.

Furthermore, although the optimal soft MIMO detector is the best among all MIMO detectors under the same channel condition, it is not always affordable for practical communications systems due to its complexity especially when the coherence time T is large. Therefore, sub-optimal soft MIMO detectors as well as the EM-based detectors turn out to be promising alternatives for their excellent trade-offs between complexity and performance over moderate to slow fading channels. As illustrated in Fig. 11, the extrinsic information transfer functions of these sub-optimal MIMO detectors are compared over the same 2×2 unknown MIMO channel with large coherence time $T = 18$. In this case (with large T), the modified EM-based MIMO detector outperforms other sub-optimal detectors and tends to approach the performance of the MIMO detector with ideal CSIR.

C. LDPC code degree profile optimization

The analysis of the mutual information transfer characteristic provided in Section IV-B not only enables us to analyze the system performance, but also provides a powerful design approach for the LDPC code

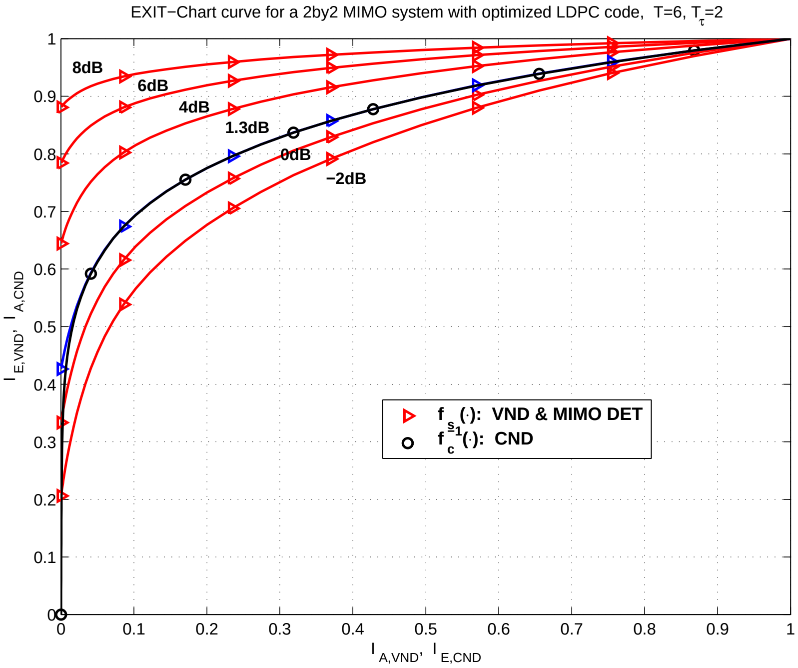

Fig. 13. EXIT-chart curve of a 2×2 optimized LDPC-coded MIMO system over a unknown fading channel with coherence time $T = 6$, and training number $T_\tau = 2$ using optimal soft MIMO detectors, under different signal to noise ratios $\rho = -2, 0, 1.3, 4, 6, 8\ dB$.

optimization. We demonstrate in Fig. 12 the EXIT-chart curves of a 2×2 regular $(3, 6)$ LDPC-coded MIMO system with codeword length 8×10^4. The simulation is carried out over an unknown fading channel with coherence time $T = 6$ and training number $T_\tau = 2$, using optimal soft MIMO detectors, and under several different system SNRs. From the plot, we can observe that $2.2dB$ is the minimal SNR that can avoid curve intersection and hence leads to successful decoding, which is further confirmed by the cliff region shown in the real simulation result of Fig. 14. We also illustrate in Fig. 13 the EXIT-chart curves for the optimized LDPC-coded MIMO system with outer code rate $R_{\text{outer}} = 1/2$ and codeword length 8×10^4 under the same system settings. It can be observed from the plot that after applying LDPC code optimization, the two mutual information transfer functions match each other perfectly (almost fall on top of each other), and achieve about $0.9dB$ gain in performance.

D. Overall coded MIMO system performance

We demonstrate in Fig. 14 the bit error rate of an LDPC-coded MIMO system over unknown fading channels. For the sake of simulation simplicity, we consider a small 2×2 MIMO system over a relatively fast unknown fading channel with coherence time $T = 6$. According to the non-coherent MIMO capacity analysis provided in [28] [29], the number of training symbols T_τ is set equal to the number of transmit antennas M, in a sense to maximizes the system capacity (or mutual information rate). The outer LDPC code is a regular $(3, 6)$ code with code rate $R_{\text{outer}} = 1/2$, and codeword length 8×10^4. By taking into account the pilots cost, the overall system coding rate is $R_{\text{overall}} = 1/3$ bits per transmission. We can observe from Fig. 14 that over 1.5dB performance gain can be achieved by using optimal soft MIMO detectors than the simple MMSE-based detector. The two sub-optimal MIMO detectors as well as the modified EM-based soft MIMO detector also provide significant performance gain, and at the same time maintain affordable decoding complexity. On the other hand, due to the existing positive feedback, the direct EM-based MIMO detector has a 2dB performance degradation compared to the modified EM-based detector and performs even worse than the simple MMSE-based detector.

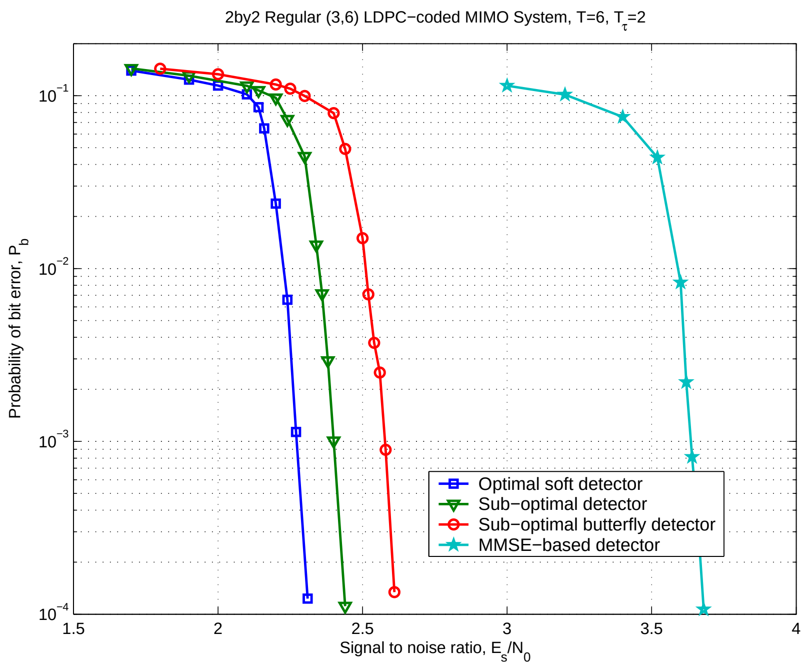

Fig. 14. Probability of bit error of a 2×2 regular $(3, 6)$ LDPC-coded MIMO system over a unknown fading channel with coherence time $T = 6$ and training number $T_\tau = 2$ using several different soft MIMO detectors.

Using the optimization algorithm provided in Section IV-C, the optimal LDPC code degree profiles (with outer code rate $R_{\text{outer}} = 1/2$ and codeword length 8×10^4) for the coded MIMO system using the different soft MIMO detection algorithms are obtained and used in the overall performance simulation. We consider the same 2×2 coded MIMO system used in Fig. 14 that transmits over the same unknown fading channel with coherence time $T = 6$ and pilot number $T_\tau = 2$ for simulations. The probability of bit error of the LDPC-coded MIMO system with optimized LDPC code degree profile is shown in Fig. 15. Compared with Fig. 14, we can achieve about 0.6dB performance gain by using the optimized LDPC degree profile as opposed to the simple regular $(3, 6)$ LDPC code. Additional simulation results,

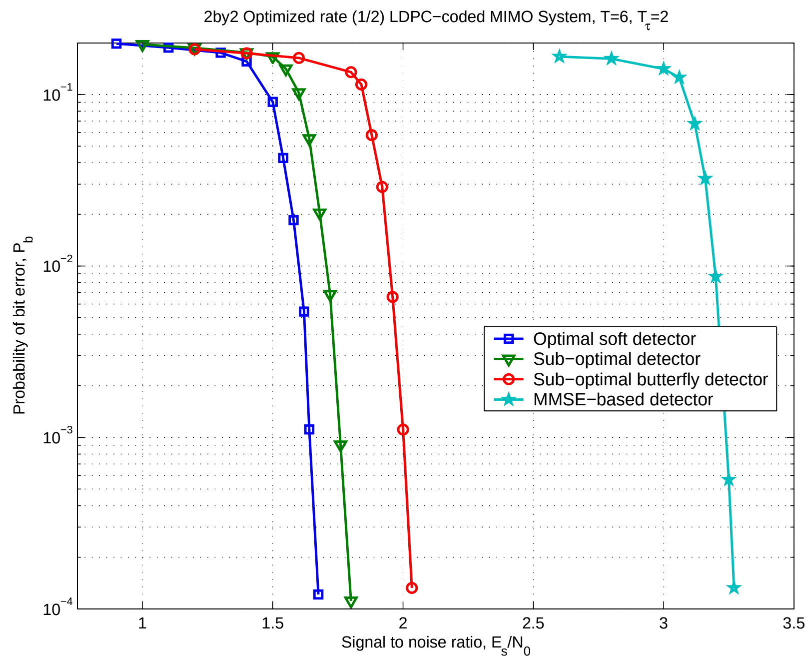

Fig. 15. Probability of bit error of a 2×2 optimized LDPC-coded MIMO system over a unknown fading channel with coherence time $T = 6$ and training number $T_\tau = 2$ using several different soft MIMO detectors.

not shown here, indicate that an even more significant performance gain can be achieved by the proposed LDPC code optimization approach if higher modulation format (such as QPSK or 16-QAM) is used, or if the coherence time interval T is larger. Under these channel conditions, the extrinsic information transfer functions (54) and (55) of a regular LDPC-coded MIMO system are very dissimilar to each other and this emphasizes the importance of the proposed curve fitting technique.

VI. Conclusion

In this paper, we developed a practical LDPC-coded MIMO system over a flat fading wireless environment with channel state information unavailable both at the transmitter and the receiver. We first proposed several soft-input soft-output MIMO detectors, including one optimal soft MIMO detector, two sub-optimal soft detectors, and a modified EM-based MIMO detector, whose performances are much better than the conventional MMSE-based detectors and offer an effective tradeoff between complexity and performance. By analyzing the extrinsic information transfer characteristic of the soft MIMO detectors, performance of the coded MIMO system using different MIMO detection algorithms are analyzed and compared under various channel conditions. Motivated by the turbo iterative principle, the LDPC-coded MIMO receiver is constructed in an unconventional manner where the soft MIMO detector and LDPC variable node decoder form one super soft-decoding unit, and the LDPC check node decoder forms the other component of the iterative decoding scheme. The proposed receiver structure has lower decoding complexity and further leads to tractable EXIT functions of the component soft decoders. Based on the obtained closed form EXIT functions, a simple and efficient LDPC code degree profile optimization algorithm is developed with proven global optimality and guaranteed convergence from any initialization. Finally, numerical and simulation results of the LDPC-coded MIMO system using the optimized degree profile further confirm the advantage of using the proposed design approach for the coded MIMO system.

REFERENCES

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecomm. (ETT)*, vol. 10, no. 6, pp. 585–595, Nov 1999.

[2] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[3] J. H. Kotecha and A. M. Sayeed, "Transmit signal design for optimal estimation of correlated MIMO channels," *IEEE Trans. on Signal Processing*, vol. 52, pp. 546–557, Feb 2004.

[4] D. Samardzija and N. Mandayam, "Pilot-assisted estimation of MIMO fading channel response and achievable data rates," *IEEE Trans. on Signal Processing*, vol. 51, pp. 2882–2890, Nov 2003.

[5] I. Bradaric, A. P. Petropulu, and K. I. Diamantaras, "Blind MIMO FIR channel identification based on second-order spectra correlations," *IEEE Trans. on Signal Processing*, vol. 51, pp. 1668–1674, June 2003.

[6] H. Sahlin and H. Broman, "MIMO signal separation for FIR channels: A criterion and performance analysis," *IEEE Trans. on Signal Processing*, vol. 48, pp. 642–649, Mar 2000.

[7] S. Amari and J. F. Cardoso, "Blind source separation-semiparametric statistical approach," *IEEE Trans. on Signal Processing*, vol. 45, pp. 2692–2700, Nov 1997.

[8] C. Jutten and J. Heuralt, "Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture," *Signal Processing*, July 1991.

[9] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 1169–1174, July 2000.

[10] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. on Information Theory*, vol. 46, pp. 2567–2578, Nov 2000.

[11] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. on Communications*, vol. 48, pp. 2041–2052, Dec 2000.

[12] B. Hochwald and T. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. on Information Theory*, vol. 46, pp. 543–564, Mar 2000.

[13] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. on Information Theory*, vol. 46, pp. 1962–1973, Sept 2000.

[14] D. Agrawal, T. J. Richardson, and R. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. on Information Theory*, vol. 47, pp. 2618–2626, Sept 2001.

[15] M. L. McCloud, M. Brehler, and M. K. Varanasi, "Signal design and convolutional coding for noncoherent space-time communication on the block-Rayleigh-fading channel," *IEEE Trans. on Information Theory*, vol. 48, pp. 1186–1194, May 2002.

[16] V. Tarokh and I.-M. Kim, "Existence and construction of noncoherent unitary space-time codes," *IEEE Trans. on Information Theory*, vol. 48, pp. 3112–3117, Dec 2002.

[17] M. J. Borran, A. Sabharwal, and B. Aazhang, "On design criteria and construction of noncoherent space-time constellations," *IEEE Trans. on Information Theory*, vol. 49, pp. 2332–2351, Oct 2003.

[18] W. Zhao, G. Leus, and G. B. Giannakis, "Algebraic design of unitary space-time constellations," in *IEEE International Conference on Communications 2003, Anchorage, AK*, May 2003, pp. 3180–3184.

[19] C. N. Georghiades and J. C. Han, "Sequence estimation in the presence of random parameters via the EM algorithm," *IEEE Trans. on Communications*, vol. 45, pp. 300–308, March 1997.

[20] G. Kaleh, "Joint parameter estimation and symbol detection for linear and nonlinear unknown channels," *IEEE Trans. on Communications*, vol. 42, pp. 2406–2413, July 1994.

[21] C. Cozzo and B. L. Hughes, "Joint channel estimation and data detection in space-time communications," *IEEE Trans. on Communications*, vol. 51, pp. 1266–1270, Aug 2003.

[22] J. J. Boutros, F. Boixadera, and C. Lamy, "Bit-interleaved coded modulations for multiple-input multiple-output channels," in *IEEE International Symposium on Spread Spectrum Techniques and Applications*, Sept 2000, vol. 1, pp. 123–126.

[23] M. Gonzalez-Lopez, J. Miguez, and L. Castedo, "Turbo aided maximum likelihood channel estimation for space-time coded systems," in *IEEE International Symposium on PIMRC 2002*, Sept 2002, pp. 364–368.

[24] H. Wymeersch, F. Simoens, and M. Moeneclaey, "Code-aided joint channel estimation and frame synchronization for MIMO systems," in *Workshop on Signal Processing Advances in Wireless Communications (SPAWC'04)*, July 2004.

[25] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. on Communications*, vol. 49, pp. 1727–1737, Oct 2001.

[26] J. Hagenauer, "The turbo principle: Tutorial introduction and state of the art," in *IEEE International Symposium on Turbo Codes and Related Topics*, Sept 1997.

[27] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Trans. on Communications*, vol. 52, pp. 670–678, Apr 2004.

[28] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. on Information Theory*, vol. 49, pp. 951–963, Apr 2003.

[29] J. Zheng and B. D.Rao, "Capacity analysis of MIMO systems with unknown channel state information," in *IEEE Information Theory Workshop 2004*, San Antonio, Oct 2004.

[30] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. on Information Theory*, vol. 27, pp. 533–547, Sept 1981.

[31] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. on Information Theory*, vol. 45, pp. 139–157, Jan 1999.

[32] L. Wei, "Several properties of short LDPC codes," *IEEE Trans. on Communications*, vol. 52, pp. 721–727, May 2004.

[33] Y. J. Ko and J. H. Kim, "Girth conditioning for construction of short block length irregular LDPC codes," *Electronics Letters*, vol. 40, pp. 187–188, Feb 2004.

[34] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: model and erasure channel properties," *to appear in IEEE Trans. on Information Theory*, 2004.