Reliability-Based Turbo Detection

Jun Tao, *Member, IEEE*, Jingxian Wu, *Member, IEEE*, and Yahong Rosa Zheng, *Senior Member, IEEE*

Abstract—This paper proposes a reliability-based turbo detection scheme for multiple-input multiple-output (MIMO) systems with frequency-selective fading. The proposed scheme performs iterative successive soft interference cancellation (SSIC) with a block decision-feedback equalizer (BDFE). To minimize the negative impacts of error propagation in SSIC, we propose a new group-wise *reliability-based* ordering scheme, where neighboring symbols that severely interfere with each other are clustered into the same group, and for each group, more "reliable" symbols are detected before less "reliable" ones. The symbol reliability is measured by using the symbol *a priori* probability, which is a unique byproduct of the turbo detection and can be obtained with little overhead. The reliability information is updated iteratively as the turbo detection progresses, and this leads to a dynamic ordering scheme that is unavailable in conventional ordered successive interference cancellation (OSIC) schemes. Simulation results show that extra performance gain is obtained at a very small ordering cost, and the reliability-based turbo detection can achieve a performance that is only 0.5 dB away from the optimum maximum *a posteriori* probability (MAP) detection.

Index Terms—Block decision-feedback equalizer (BDFE), multiple-input multiple-output (MIMO), reliability-based turbo detection, symbol reliability.

I. INTRODUCTION

TURBO detection is a powerful receiver technique that improves system performance by iteratively exchanging extrinsic soft information between a soft-decision equalizer and a soft-decision channel decoder. The optimum turbo detection schemes [1], [2] adopt the soft-output Viterbi algorithm (SOVA) [3] or the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [4] for channel equalization and decoding. Both SOVA and BCJR algorithms are trellis-based methods, and their computational complexities increase exponentially as $Q^{N(L-1)}$ [5] with Q, N and L being the modulation level, the number of transmit antennas and the channel length, respectively. The complexity becomes prohibitively high when L and Q are large. This necessitates the study of low-complexity sub-optimal turbo detection (or equalization) algorithms.

A large number of low-complexity turbo equalization algorithms have been developed during the last decade to trade-off complexity with performance for both single-input single-output (SISO) systems [6]–[13] and multiple-input multiple-output (MIMO) systems [14]–[16]. In [6]–[8], linear minimum

mean square error (LMMSE) equalizers with soft interference cancellation have been used to replace the BCJR-based optimum maximum *a posteriori* probability (MAP) equalizer for SISO systems. The variations of [6] can be found in [9]–[11], where a decision-aided equalization (DAE) is proposed in [9] and an adaptive equalizer is adopted in [10] by taking into account the sub-optimality of the previously-estimated symbols. The result in [10] has also been extended to infinite-length LMMSE equalizer in [11]. In [12], a joint coding and decision-feedback equalization (DFE) scheme is proposed, and the soft information from the output of both the equalizer and the channel decoder are fed back for interference cancellation. A soft feedback equalizer is proposed in [13], where the soft output of the equalizer is used for inter-symbol interference (ISI) cancellation. The scheme in [13] is different from the turbo DFE with hard decision feedback as proposed in [7]. The turbo detection scheme proposed in [6] has been extended to MIMO systems [14], [15]. In [16], pre-filtering is employed to reduce the number of channel trellis states so that the BCJR-based equalization can be performed with reduced complexity for MIMO systems. Recently, a turbo detection structure employing block decision feedback equalizer (BDFE) was proposed in [17] for SISO systems. The BDFE can achieve a better performance than the conventional linear equalizer (LE) and DFE [18]. The turbo equalization in [17] has enabled a low-complexity sequence-based log-likelihood ratio (LLR) calculation for performance improvement.

In this paper, we propose a reliability-based turbo detection scheme for MIMO block-transmission systems with frequency-selective fading. In a spatially-multiplexed MIMO system with frequency-selective fading, the interference among the transmitted symbols arises from two sources: the ISI due to the time dispersion of the fading, and the spatial multiplexing interference (MI) among the concurrent data streams sent by multiple transmit antennas. In the proposed turbo detection scheme, successive soft interference cancellation (SSIC) is performed in both the time domain on ISI, and the space domain on MI. The space-time SSIC subtracts the interference of previously-detected symbols, in the form of soft decisions, during the detection of the current symbol. Due to the possible error propagation in SSIC, the order in which the symbols are detected is critical to the detection performance. An unreliable soft decision will negatively affect the detection of the subsequent symbols. To combat error propagation and improve the SSIC performance, we propose a low-complexity group-wise *reliability-based* ordering scheme. In the proposed scheme, an entire block of symbols are first divided into multiple groups, such that adjacent symbols severely interfering each other are clustered into the same group. Then, within each group, symbols are ordered according to a new *a priori* reliability metric, which is calculated
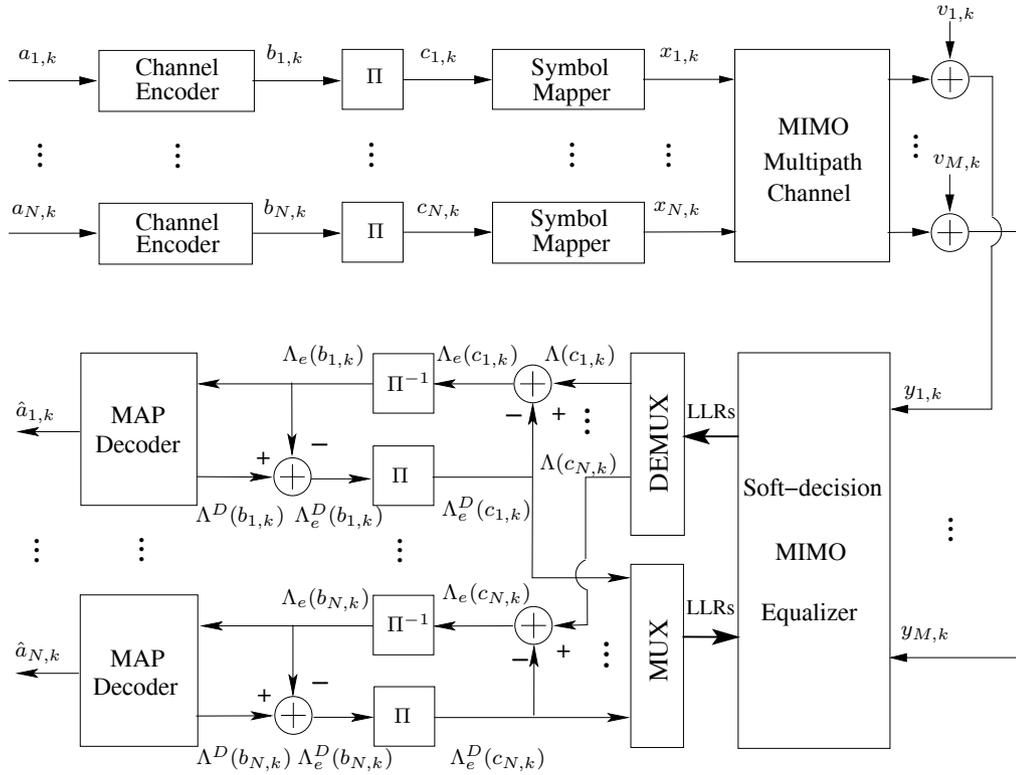
Fig. 1. A MIMO communication system with turbo detection.

from the symbol *a priori* probability and is a unique byproduct of turbo detection. Symbols with higher reliability will be detected before those with lower reliability.

Compared with existing ordered successive interference cancellation (OSIC) schemes in the literature [19]–[22], the proposed *reliability-based* ordering scheme contributes to performance improvement from three perspectives. First, the ordering is based on the symbol *a priori* probability which can be obtained at a very low cost. On the other hand, most existing OSIC schemes rely on the channel condition in terms of mean square error (MSE) or signal-to-noise ratio (SNR), and the ordering process usually involves matrix pseudo-inverse operations thus is computationally intensive [21]. An LLR-based ordering method is proposed in [23]. However, the reliability information therein is in the form of symbol *a posteriori* LLR, and the calculation of which still requires matrix inversion for each symbol. Second, with the proposed scheme, the ordering is performed in the two-dimensional (2-D) space-time domain, while existing OSIC schemes perform ordering only in the one-dimensional (1-D) space domain. Third, since the symbol *a priori* knowledge changes as the iteration progresses, the proposed ordering scheme is inherently *dynamic*, compared with the *static* ordering of the existing OSIC schemes.

The reliability-based turbo detection is developed by using the block decision feedback equalizer [18] as a construction tool. The model adopted in this work is slightly different from that used in [18]. The adopted model uses a zero-padding transmission block, and all the received samples from one block are used during the detection. Such a model has two advantages: first, the detection performance is improved since more useful information is incorporated during the equaliza-

tion process; second, it leads to a block-Toeplitz structure of the equivalent channel matrix, which enables the adoption of a fast algorithm for the equalizer matrices calculation.

The rest of this paper is organized as follows. Section II presents the MIMO block transmission model and reviews the basics of turbo detection. The development of the reliability-based MIMO turbo detection using BDFE is detailed in Section III, where a low-complexity implementation of the soft-decision BDFE is also proposed. Section IV analyzes the complexity of the proposed detection scheme, and compares it with that of a linear turbo equalization scheme. Simulation results are presented in Section V, and Section VI concludes the paper.

Notation: The superscripts, $(\cdot)^t$ and $(\cdot)^h$, represent the matrix transpose and conjugate transpose, respectively. The $K \times P$ complex matrix space is denoted by $\mathcal{C}^{K \times P}$. An identity matrix of size K is denoted as $\mathbf{I}_K$, and a $K \times K$ diagonal matrix with diagonal elements $d_1, d_2, \cdots, d_K$ is denoted as $\mathrm{diag}\{d_1, d_2, \cdots, d_K\}$. The operator $\mathbb{E}(\cdot)$ represents the mathematical expectation, and $\Pi(\cdot)$ and $\Pi^{-1}(\cdot)$ denote interleaving and de-interleaving operations, respectively. The operators $\min(\cdot)$, $\max(\cdot)$, and $\mathrm{mod}(\cdot)$ perform minimization, maximization, and modulus operations, respectively. $\lceil x \rceil$ returns the smallest integer that is larger than x, trace $(\mathbf{A})$ returns the trace of the matrix $\mathbf{A}$, and $\mathbf{a} \ll k$ performs a left circular shift operation on the vector $\mathbf{a}$ by k locations.

II. SYSTEM DESCRIPTION

Consider an $N \times M$ MIMO communication system as shown in Fig. 1, with N and M being the number of transmit antennas and receive antennas, respectively. At the transmitter,

N information bit streams, with $a_{n,k}$ being the k-th information bit of the n-th stream, are multiplexed onto the N transmit branches. The n-th bit stream is encoded, interleaved, and modulated. The outputs of the encoder, the interleaver (Π), and the symbol mapper (modulator) are denoted by $b_{n,k}$, $c_{n,k}$, and $x_{n,k}$, respectively. For a Q-ary modulation constellation set $\mathcal{S} = \{\chi_q\}_{q=1}^{Q}$, every $\log_2 Q$ coded bits are mapped onto one modulation symbol, *i.e.*, the group of bits, $\{c_{q,p}\}_{p=1}^{\log_2 Q}$, are mapped to the modulation symbol χ_q. The modulation symbols on the n-th branch are transmitted via the n-th transmit antenna in the form of a length-N_b block as, $\mathbf{x}^{(n)} = [x_{n,1}, x_{n,2}, \cdots, x_{n,N_b}] \in \mathcal{S}^{1 \times N_b}$. To avoid inter-block interference (IBI), guard intervals (GI) are inserted among the transmitted blocks. The GI is implemented as zeros padded at the end of each block in this paper. It is noted that to maintain a high transmission efficiency in practical systems, the GI is usually not necessary provided that IBI can be properly reconstructed and canceled with either pilot symbols or already detected symbols.

On the receiver side, the received samples of one block at the m-th receive antenna can be written as

$$y_{m,k} = \sum_{l=0}^{L-1} \sum_{n=1}^{N} f_{m,n}(l) x_{n,k-l} + v_{m,k} \qquad (1)$$

for $k = 1, 2, \cdots, N_b + L - 1$, where $f_{m,n}(l)$ is the l-th tap of the equivalent discrete-time channel between the n-th transmit antenna and the m-th receive antenna, L is the channel length, and $v_{m,k}$ is the zero-mean additive white Gaussian noise (AWGN) with variance σ_v^2. With zero-padding block transmission, $x_{n,k} = 0$ for $k \leq 0$. It has been assumed that the channel is constant over one block.

Stacking the received samples from all M receive antennas at time k into a vector as, $\mathbf{y}_k = [y_{1,k}, y_{2,k}, \cdots, y_{M,k}]^t \in \mathcal{C}^{M \times 1}$, we have

$$\mathbf{y}_k = \sum_{l=0}^{L-1} \mathbf{F}_l \mathbf{x}_{k-l} + \mathbf{v}_k \qquad (2)$$

where $\mathbf{x}_{k-l} = [x_{1,k-l}, x_{2,k-l}, \cdots, x_{N,k-l}]^t \in \mathcal{S}^{N \times 1}$, $\mathbf{v}_k = [v_{1,k}, v_{2,k}, \cdots, v_{M,k}]^t \in \mathcal{C}^{M \times 1}$, and the (m,n)-th element of the channel matrix, $\mathbf{F}_l \in \mathcal{C}^{M \times N}$, is $f_{m,n}(l)$. The received block is then obtained by stacking $\{\mathbf{y}_k\}_{k=1}^{N_b + L - 1}$ into a column vector as $\mathbf{y} = [\mathbf{y}_1^t, \mathbf{y}_2^t, \cdots, \mathbf{y}_{N_b+L-1}^t]^t \in \mathcal{C}^{M(N_b+L-1) \times 1}$, which can be denoted as

$$\mathbf{y} = \mathbf{F}\mathbf{x} + \mathbf{v} \qquad (3)$$

where $\mathbf{x} = [\mathbf{x}_1^t, \mathbf{x}_2^t, \cdots, \mathbf{x}_{N_b}^t]^t \in \mathcal{S}^{NN_b \times 1}$, $\mathbf{v} = [\mathbf{v}_1^t, \mathbf{v}_2^t, \cdots, \mathbf{v}_{N_b+L-1}^t]^t \in \mathcal{C}^{M(N_b+L-1) \times 1}$, and the block channel matrix, $\mathbf{F} \in \mathcal{C}^{M(N_b+L-1) \times NN_b}$, is given as

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_0 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{F}_{L-1} & \cdots & \cdots & \mathbf{F}_0 & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{F}_{L-1} & \cdots & \cdots & \mathbf{F}_0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{F}_{L-1} \end{bmatrix}. \qquad (4)$$

The channel matrix in (4) possesses a block-Toeplitz structure that is desirable for fast algorithm implementation.

For a MAP-based detection, the detector calculates the LLR of the information bit as follows

$$\begin{aligned} \Lambda[a_{n,k}] &= \ln \frac{P[a_{n,k}=0|\mathbf{y}]}{P[a_{n,k}=1|\mathbf{y}]} \\ &= \ln \underbrace{\frac{\sum\limits_{\forall \mathbf{a}: a_{n,k}=0} p(\mathbf{y}|\mathbf{a}) \prod\limits_{\forall n', k': k' \neq k \text{ if } n'=n} P(a_{n',k'})}{\sum\limits_{\forall \mathbf{a}: a_{n,k}=1} p(\mathbf{y}|\mathbf{a}) \prod\limits_{\forall n', k': k' \neq k \text{ if } n'=n} P(a_{n',k'})}}_{\Lambda_e[a_{n,k}]} + \Lambda_a[a_{n,k}] \quad (5) \end{aligned}$$

where $\Lambda_a[a_{n,k}] = \ln \frac{P[a_{n,k}=0]}{P[a_{n,k}=1]}$ is the *a priori* LLR of $a_{n,k}$. The computation complexity of (5) increases exponentially with the block size N_b [24], and it is prohibitively high for practical applications.

A common way to reduce the detection complexity is to separate the equalization and decoding operations [6]-[16], as shown at the bottom part of Fig. 1. The output LLRs from the soft-decision MIMO equalizer is de-multiplexed, de-interleaved, and then delivered as the *a priori* input to the N channel decoders. The decoders then generate new LLRs, which are interleaved, multiplexed, and input as the *a priori* knowledge for the equalizer to start the next iteration. To avoid early limit-cycle behavior [7], only the extrinsic LLR $\Lambda_e[c_{n,k}] = \Lambda[c_{n,k}] - \Lambda_e^D[c_{n,k}]$ of the equalizer and the extrinsic LLR $\Lambda_e^D[b_{n,k}] = \Lambda^D[b_{n,k}] - \Lambda_e[b_{n,k}]$ of the decoder, are exchanged. Successive iterations are performed until the detection converges. In the final iteration, the MAP decoders calculate the LLRs, $\{\Lambda[a_{n,k}]\}_{n=1}^{N}$, of the N information streams and output the corresponding hard decisions, $\{\hat{a}_{n,k}\}_{n=1}^{N}$. In practice, the MAP decoder is commonly adopted due to its moderate number of trellis states. The MAP-based equalizer, however, may incur a prohibitive complexity on the order of $\mathcal{O}(Q^{N(L-1)})$, for highly dispersive channels with large L. We next develop a sub-optimum soft-decision MIMO equalizer that can approach the performance of the optimum MAP equalizer, but at a much lower complexity.

III. RELIABILITY-BASED MIMO TURBO DETECTION USING BDFE

In this section, a soft-decision MIMO BDFE employing *reliability-based* ordering is developed.

A. Reliability-Based Ordering

Define the *a priori* mean and the *a priori* variance of the symbol, $x_{n,k}$, as

$$\bar{x}_{n,k} = \sum_{q=1}^{Q} \chi_q P(x_{n,k} = \chi_q), \qquad (6a)$$

$$\sigma_{n,k}^2 = \sum_{q=1}^{Q} |\chi_q - \bar{x}_{n,k}|^2 P(x_{n,k} = \chi_q) \qquad (6b)$$

where $P(x_{n,k} = \chi_q)$ is the *a priori* probability of the symbol $x_{n,k}$. It can be obtained from the *a priori* probabilities of its demapping bits, $\{c_{n,(k-1)\log_2 Q+p}\}_{p=1}^{\log_2 Q}$, as $P(s_{n,k} = \chi_q) = \prod_{p=1}^{\log_2 Q} P[c_{n,(k-1)\log_2 Q+p} = c_{q,p}]$. Details are referred to [8].

The reliability information of a symbol is extracted from its *a priori* statistics. We propose to measure the symbol reliability by using its *a priori* variance defined in (6b). The variance of a random variable (RV) measures the expected squared deviation of the RV from its mean. A lower *a priori* variance means less uncertainty, thus a higher reliability. Therefore, we define the *a priori* reliability measure of the symbol $x_{n,k}$ as

$$\rho_{n,k} = \frac{1}{\sigma_{n,k}^2}. \tag{7}$$

With the reliability information, we are able to determine the detection ordering. Intuitively, the symbols shall be detected in a fashion that more reliable symbols are detected before less reliable ones, so that the potential negative effects of error propagation in the SSIC can be reduced. As a result, we need to order the reliability measures over the entire block. The complexity of such a global ordering (for example, using the classic bubble sorting algorithm) is proportional to $N^2 N_b^2$, and it could be high when N_b is large.

We propose a group-wise ordering scheme to minimize the ordering overhead during the turbo detection. The motivation for the group-wise ordering comes from the fact that a given symbol mainly interferes a group of its neighboring symbols. Specifically, for a causal channel with length L, a symbol transmitted at the time instant k will only interfere those transmitted between the time instants k to $k + L - 1$. Define the reliability vector at the time instant k as $\boldsymbol{\rho}_k \triangleq [\rho_{1,k}, \cdots, \rho_{N,k}]^t \in \mathcal{C}^{N \times 1}$, and the group-wise *reliability-based* ordering scheme is presented as follows.

Step 1: Divide the reliability information of the entire block into N_g groups as

$$\mathcal{G}_1 = \{\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \cdots, \boldsymbol{\rho}_{L_g}\}, \mathcal{G}_2 = \{\boldsymbol{\rho}_{L_g+1}, \boldsymbol{\rho}_{L_g+2}, \cdots, \boldsymbol{\rho}_{2L_g}\},$$
$$\cdots, \mathcal{G}_{N_g} = \{\boldsymbol{\rho}_{(N_g-1)L_g+1}, \boldsymbol{\rho}_{(N_g-1)L_g+2}, \cdots, \boldsymbol{\rho}_{N_g L_g}\} \tag{8}$$

where $N_g L_g = N_b$, and it is obvious that each group contains L_g reliability vectors or equivalently, $N L_g$ reliability measures.

Step 2: Sort the reliability measures within each group in an ascending order as

$$\mathcal{G}'_i = \{\rho_{\langle O'_{(i-1)NL_g+1}\rangle}, \rho_{\langle O'_{(i-1)NL_g+2}\rangle}, \cdots, \rho_{\langle O'_{iNL_g}\rangle}\} \tag{9}$$

for $i = 1, 2, \cdots, N_g$, where the ordered index set, $\{O'_{(i-1)NL_g+1}, O'_{(i-1)NL_g+2}, \cdots, O'_{iNL_g}\}$, is a permutation of the index set, $\{(i-1)NL_g+1, (i-1)NL_g+2, \cdots, iNL_g\}$. An index mapping operator, $\langle \cdot \rangle$, has been introduced and it is defined as

$$\langle \cdot \rangle : (n, k) = \langle j \rangle \text{ with } k = \left\lceil \frac{j}{N} \right\rceil, n = \begin{cases} \text{mod}(j, N) \neq 0 \\ N, \text{ o.w.} \end{cases} \tag{10}$$

Step 3: Obtain the ordering, $\boldsymbol{O}$, by assembling the reliability indices of all the ordered groups in (9) as

$$\boldsymbol{O} = \{\langle O'_1\rangle, \langle O'_{NL_g+1}\rangle, \cdots, \langle O'_{(N_g-1)NL_g+1}\rangle | \langle O'_2\rangle, \langle O'_{NL_g+2}\rangle,$$
$$\cdots, \langle O'_{(N_g-1)NL_g+2}\rangle | \cdots | \langle O'_{NL_g}\rangle, \cdots, \langle O'_{N_g NL_g}\rangle\} \tag{11}$$

For the convenience of analysis, we simplify the notation of (11) as follows

$$\boldsymbol{O} = \{\langle O_1\rangle, \langle O_2\rangle, \cdots, \langle O_{N_g}\rangle | \langle O_{N_g+1}\rangle, \langle O_{N_g+2}\rangle, \cdots, \langle O_{2N_g}\rangle |$$
$$\cdots | \langle O_{(NL_g-1)N_g+1}\rangle, \langle O_{(NL_g-1)N_g+2}\rangle, \cdots, \langle O_{NL_g N_g}\rangle\} \tag{12}$$

So far, the ordering process is completed. Before moving into the development of the soft-decision BDFE based on the ordering, we provide two remarks on the group-wise *reliability-based* ordering scheme.

Remark 1: As mentioned above, the group-wise ordering is motivated by the finite memory of the channel. Therefore, a natural choice of the group size is $L_g = L$. Increasing or decreasing the group size only incurs marginal performance gain or loss, as will be shown in the simulation results.

Remark 2: The sorting operation for the N_g groups incurs an overall complexity on the order of $\mathcal{O}(N^2 N_b L_g)$, which is proportional to the group size L_g. When $L_g = N_b$, which corresponds to the global ordering, the ordering complexity is on the order of $\mathcal{O}(N^2 N_b^2)$ as mentioned before. In practical applications, a tradeoff between the ordering complexity and the detection performance can be achieved by selecting a proper group size.

B. Development of The Soft-Decision BDFE

With the ordering in (12), an alternative system representation of (3) can be obtained as

$$\mathbf{y} = \mathbf{H}\mathbf{x}' + \mathbf{v} \tag{13}$$

where $\mathbf{x}' = [x_{\langle O_1\rangle}, x_{\langle O_2\rangle}, \cdots, x_{\langle O_{NN_b}\rangle}]^t$, $\mathbf{H} = [\mathbf{f}_{O_1}, \mathbf{f}_{O_2}, \cdots, \mathbf{f}_{O_{NN_b}}]$, with $\mathbf{f}_j$ being the j-th column of the channel matrix $\mathbf{F}$. Based on the ordered system model in (13), the soft-decision BDFE as shown in Fig. 2 is developed in this subsection.

From the figure, to detect the symbol $x_{\langle O_g\rangle}$, a feedforward filter, $\mathbf{C}_g \in \mathcal{C}^{NN_b \times M(N_b+L-1)}$, and a feedback filter, $\mathbf{D}_g \in \mathcal{C}^{NN_b \times NN_b}$, are adopted, leading to the output vector

$$\tilde{\mathbf{x}}'_g = \mathbf{C}_g \left(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}'_g\right) - \mathbf{D}_g \left(\hat{\mathbf{x}}' - \bar{\mathbf{x}}'_g\right) + \bar{\mathbf{x}}'_g \tag{14}$$

where $\bar{\mathbf{x}}'_g = [\bar{x}_{\langle O_1\rangle}, \cdots, \bar{x}_{\langle O_{g-1}\rangle}, 0, \bar{x}_{\langle O_{g+1}\rangle}, \cdots, \bar{x}_{\langle O_{NN_b}\rangle}]^t$ is the ordered *a priori* mean vector with the g-th element being zero, $\hat{\mathbf{x}}' = [\hat{x}_{\langle O_1\rangle}, \cdots, \hat{x}_{\langle O_{NN_b}\rangle}]^t$ is the ordered tentative soft decision vector. The tentative soft symbol decision is defined as the symbol *a posteriori* mean as

$$\hat{x}_{\langle O_g\rangle} = \sum_{q=1}^{Q} \chi_q P(x_{\langle O_g\rangle} = \chi_q | \mathbf{y}) \tag{15}$$

where $P(x_{\langle O_g\rangle} = \chi_q | \mathbf{y})$ is the symbol *a posteriori* probability (APP). The calculation of $P(x_{\langle O_g\rangle} = \chi_q | \mathbf{y})$ is discussed in Subsection III-C. Using soft instead of hard decisions during the successive interference cancellation (SIC) will partly reduce the negative effects of error propagation. To enable SIC, the feedback filter $\mathbf{D}_g$ is designed as a zero-diagonal strict upper triangular matrix, and the symbol detection is performed successively by following the reverse order of the elements in $\mathbf{x}'$, *i.e.*, the last element in $\mathbf{x}'$ is detected first and the first element in $\mathbf{x}'$ is detected last.
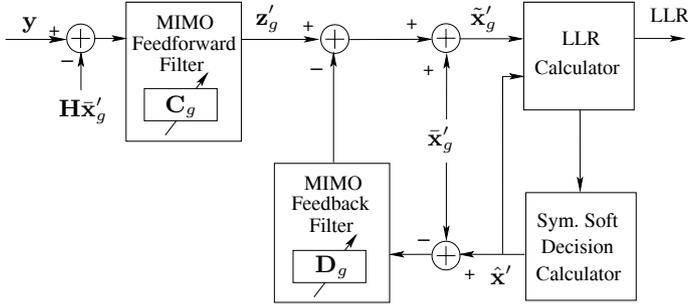
Fig. 2.   Soft-decision MIMO BDFE (when $x_{\langle O_g \rangle}$ is equalized).

With the common assumption of perfect decision feedback [25], *i.e.*, $\hat{\mathbf{x}}' = \mathbf{x}'$, the error vector of BDFE can be written as

$$\mathbf{e}_g = \mathbf{C}_g \left(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}'_g \right) - \left(\mathbf{D}_g + \mathbf{I}_{NN_b} \right) \left(\mathbf{x}' - \bar{\mathbf{x}}'_g \right) \qquad (16)$$

Minimizing the mean square error, $\sigma^2_{e_g} \triangleq \mathbb{E}\left[\mathbf{e}^h_g \mathbf{e}_g \right]$, with the orthogonality principle, $\mathbb{E}\left[\mathbf{e}_g \left(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}'_g \right)^h \right] = \mathbf{0}$, leads to

$$\mathbf{C}_g = \mathbf{R}_g \mathbf{\Sigma}'_g \mathbf{H}^h \left[\mathbf{H}\mathbf{\Sigma}'_g \mathbf{H}^h + \sigma^2_v \mathbf{I}_{M(N_b+L-1)} \right]^{-1} \qquad (17)$$

where $\mathbf{R}_g = \mathbf{D}_g + \mathbf{I}_{NN_b}$, and $\mathbf{\Sigma}'_g = \text{diag}\{\sigma^2_{\langle O_1 \rangle}, \cdots, \sigma^2_{\langle O_{g-1} \rangle}, 1, \sigma^2_{\langle O_{g+1} \rangle}, \cdots, \sigma^2_{\langle O_{NN_b} \rangle}\}$ is the ordered *a priori* covariance matrix of $\mathbf{x}'$ with its g-th element set as 1. Combining (16) and (17) leads to

$$\mathbf{e}_g = \mathbf{R}_g \boldsymbol{\epsilon}_g \qquad (18)$$

with

$$\boldsymbol{\epsilon}_g = \mathbf{\Sigma}'_g \mathbf{H}^h \left[\mathbf{H}\mathbf{\Sigma}'_g \mathbf{H}^h + \sigma^2_v \mathbf{I}_{M(N_b+L-1)} \right]^{-1} \left(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}'_g \right) - \left(\mathbf{x}' - \bar{\mathbf{x}}'_g \right) \qquad (19)$$

The autocorrelation matrix of $\boldsymbol{\epsilon}_g$ can be calculated as

$$\mathbf{\Sigma}_{\epsilon_g} = \mathbf{\Sigma}'_g - \mathbf{\Sigma}'_g \mathbf{H}^h \left[\mathbf{H}\mathbf{\Sigma}'_g \mathbf{H}^h + \sigma^2_v \mathbf{I}_{M(N_b+L-1)} \right]^{-1} \mathbf{H}\mathbf{\Sigma}'_g$$

$$= \left[\mathbf{\Sigma}'^{-1}_g + \frac{1}{\sigma^2_v} \mathbf{H}^h \mathbf{H} \right]^{-1} \qquad (20)$$

where the second equality is obtained with the identity $(\mathbf{A} + \mathbf{C}\mathbf{B}\mathbf{C}^h)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}(\mathbf{B}^{-1} + \mathbf{C}^h\mathbf{A}^{-1}\mathbf{C})^{-1}\mathbf{C}^h\mathbf{A}^{-1}$. Based on (18) and (20), the autocorrelation matrix of $\mathbf{e}_g$ can be expressed by

$$\mathbf{\Sigma}_{e_g} = \mathbf{R}_g \left[\mathbf{\Sigma}'^{-1}_g + \frac{1}{\sigma^2_v} \mathbf{H}^h \mathbf{H} \right]^{-1} \mathbf{R}^h_g. \qquad (21)$$

Since $\sigma^2_{e_g} = \text{trace}(\mathbf{\Sigma}_{e_g})$, the minimization of the mean square error is equivalent to minimizing $\text{trace}(\mathbf{\Sigma}_{e_g})$. In addition, the solution that minimizes $\text{trace}(\mathbf{\Sigma}_{e_g})$ should also satisfy the successive interference cancellation constraint, *i.e.*, $\mathbf{R}_g = \mathbf{D}_g + \mathbf{I}_{NN_b}$ needs to be an upper triangular matrix with unit diagonal elements. The solution satisfying the above conditions is found as [26]

$$\mathbf{R}_g = \mathbf{U}^h_g \qquad (22)$$

where $\mathbf{U}_g \in \mathcal{C}^{NN_b \times NN_b}$ is a lower triangular matrix with unit diagonal elements, and it is obtained from the Cholesky decomposition of $\mathbf{\Sigma}'^{-1}_g + \frac{1}{\sigma^2_v}\mathbf{H}^h\mathbf{H}$ as

$$\mathbf{\Sigma}'^{-1}_g + \frac{1}{\sigma^2_v}\mathbf{H}^h\mathbf{H} = \mathbf{U}_g \mathbf{\Delta}_g \mathbf{U}^h_g \qquad (23)$$

with $\mathbf{\Delta}_g \in \mathcal{C}^{NN_b \times NN_b}$ being a diagonal matrix. The solution for the feedforward and feedback filters are

$$\mathbf{C}_g = \mathbf{U}^h_g \mathbf{\Sigma}'_g \mathbf{H}^h \left[\mathbf{H}\mathbf{\Sigma}'_g \mathbf{H}^h + \sigma^2_v \mathbf{I}_{M(N_b+L-1)} \right]^{-1} \quad (24a)$$

$$\mathbf{D}_g = \mathbf{U}^h_g - \mathbf{I}_{NN_b} \qquad (24b)$$

C. LLR Calculation Based on Equalized Symbols

The calculation of the bit LLRs based on the equalized symbols is discussed in this subsection. Based on (16), an equivalent equalization model is obtained as

$$\mathbf{z}_g \triangleq \mathbf{C}_g(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}'_g) = \mathbf{R}_g \left(\hat{\mathbf{x}}' - \bar{\mathbf{x}}'_g \right) + \mathbf{e}_g. \qquad (25)$$

The g-th element in $\mathbf{z}_g$ is expressed by

$$z_g = x_{\langle O_g \rangle} + \sum_{l=g+1}^{NN_b} r_{g,l} \left[\hat{x}_{\langle O_l \rangle} - \bar{x}_{\langle O_l \rangle} \right] + e_g \qquad (26)$$

where $r_{g,l}$ is the (g,l)-element of $\mathbf{R}_g$ with $r_{g,g} = 1$. The APP of $x_{\langle O_g \rangle}$ conditioned on z_g is

$$P\left(x_{\langle O_g \rangle} | z_g \right) = \frac{p\left(z_g | x_{\langle O_g \rangle} \right) P\left(x_{\langle O_g \rangle} \right)}{p\left(z_g \right)}. \qquad (27)$$

With the common approximation that e_g in (26) is a zero-mean complex Gaussian random variable, we determine the conditional probability as

$$p\left(z_g | x_{\langle O_g \rangle} \right) = \frac{1}{\pi \alpha^2_g} \exp\left\{ \frac{-|\beta_g|^2}{\alpha^2_g} \right\} \qquad (28)$$

where $\beta_g = z_g - x_{\langle O_g \rangle} - \sum_{l=g+1}^{NN_b} r_{g,l} \left[\hat{x}_{\langle O_l \rangle} - \bar{x}_{\langle O_l \rangle} \right]$. Since $\mathbf{\Sigma}_{e_g} = \mathbf{\Delta}^{-1}_g$, the variance of e_g is $\alpha^2_g = \delta^{-1}_{g,g}$ with $\delta_{g,g}$ being the g-th diagonal element of $\mathbf{\Delta}_g$. The symbol *a priori* probability, $P\left(x_{\langle O_g \rangle} \right)$, in (27) can be calculated based on the bit *a priori* probabilities as shown before, and it is initialized as $P\left(x_{\langle O_g \rangle} = \chi_q \right) = \frac{1}{Q}$ during the first iteration. The value of $p(z_g)$ can be obtained by using the normalization $\sum_{q=1}^Q P\left(x_{\langle O_g \rangle} = \chi_q | z_g \right) = 1$.

Once the symbol APP $p\left(x_{\langle O_g \rangle} | z_g \right)$ is obtained, the APP of the demapping bits $\left\{ c_{n_g,(k_g-1)\log_2 Q+p} \right\}_{p=1}^{\log_2 Q}$ of the symbol $x_{\langle O_g \rangle}$, can be calculated as follows

$$P\left[c_{n_g,(k_g-1)\log_2 Q+p} = b | z_g \right] = \sum_{x_{\langle O_g \rangle} \in \mathcal{S}^{(b)}_p} P\left(x_{\langle O_g \rangle} | z_g \right) \quad (29)$$

for $p = 1, 2, \cdots, \log_2 Q$, where $(n_g, k_g) = \langle O_g \rangle$ and $\mathcal{S}^{(b)}_p \triangleq \{\chi_q | \chi_q \in \mathcal{S} : c_{q,p} = b\}$ with $b \in \{0,1\}$. The LLR of the code bit can then be calculated as

$$\Lambda\left[c_{n_g,(k_g-1)\log_2 Q+p} \right] = \ln \frac{\sum\limits_{x_{\langle O_g \rangle} \in \mathcal{S}^{(0)}_p} P\left(x_{\langle O_g \rangle} | z_g \right)}{\sum\limits_{x_{\langle O_g \rangle} \in \mathcal{S}^{(1)}_p} P\left(x_{\langle O_g \rangle} | z_g \right)}. \qquad (30)$$

for $p = 1, 2, \cdots, \log_2 Q$.

Finally, since it is computationally expensive to calculate the sequence-based symbol APP $P(x_{\langle O_g \rangle}|\mathbf{y})$, the APP $P\left(x_{\langle O_g \rangle}|z_g\right)$ in (27) has been adopted for calculating the tentative soft decision, $\hat{x}_{\langle O_g \rangle}$, in (15).

D. Low-Complexity Implementation of The Soft-Decision BDFE

One of the main computational burdens of the soft-decision BDFE comes from the calculation of the equalizer matrices in (24a) and (24b), which need to be updated for each symbol. The calculation of the matrices $\mathbf{C}_g$ and $\mathbf{D}_g$ involves one Cholesky decomposition in (23), and one matrix inversion and three matrix multiplications in (24a), resulting in a computational complexity on the order of $\mathcal{O}((NN_b)^3)$. The overall complexity over the entire block is then on the order of $\mathcal{O}((NN_b)^4)$, which is prohibitively high when N_b gets large. It is desirable to develop a low-complexity implementation. We propose to reduce the computation complexity of the MIMO BDFE from two perspectives described as follows.

First, the symbol-dependent equalizer matrices, $\mathbf{C}_g$ and $\mathbf{D}_g$, are replaced with constant symbol-independent matrices, $\mathbf{C}$ and $\mathbf{D}$, to achieve a low-complexity approximation of the soft-decision MIMO BDFE. A close observation on (24) reveals that the symbol-wise filter update is solely due to the dependence on the second-order *a priori* information $\mathbf{\Sigma}_g'$. Therefore, computational complexity can be significantly reduced by replacing $\mathbf{\Sigma}'$ with the constant autocovariance matrix $\mathbf{\Sigma}' = \mathrm{diag}\{\sigma_{\langle O_1 \rangle}^2, \sigma_{\langle O_2 \rangle}^2, \cdots, \sigma_{\langle O_{NN_b} \rangle}^2\}$, which leads to the constant filter matrices, $\mathbf{C}$ and $\mathbf{D}$, for all the symbols to be equalized. It is demonstrated through simulations that employing constant equalizer matrices for all the symbols does not apparently degrade the equalization performance. It is noted that with the constant filter matrices $\mathbf{C}$ and $\mathbf{D}$, the computations of (25) and (26) over the entire block is also simplified significantly, as shown in the next section on the complexity analysis.

Second, fast algorithm is employed to further reduce the computation complexity. Based on (24a), the calculation of the constant feedforward filter $\mathbf{C}$ is given as

$$\mathbf{C} = \mathbf{R}\mathbf{\Sigma}'\mathbf{H}^h \left[\mathbf{H}\mathbf{\Sigma}'\mathbf{H}^h + \sigma_v^2 \mathbf{I}_{M(N_b+L-1)}\right]^{-1}$$
$$= \frac{1}{\sigma_v^2}\mathbf{R}\left[\mathbf{\Sigma}'^{-1} + \frac{1}{\sigma_v^2}\mathbf{H}^h\mathbf{H}\right]^{-1}\mathbf{H}^h. \quad (31)$$

With the Cholesky decomposition, $\mathbf{\Sigma}'^{-1} + \frac{1}{\sigma_v^2}\mathbf{H}^h\mathbf{H} = \mathbf{U}\mathbf{\Delta}\mathbf{U}^h$, (31) can be simplified to

$$\mathbf{C} = \frac{1}{\sigma_v^2}\mathbf{\Delta}^{-1}\mathbf{U}^{-1}\mathbf{H}^h \quad (32)$$

Since $\mathbf{\Delta}$ is a diagonal matrix, it is easy to calculate $\mathbf{\Delta}^{-1}$. As of $\mathbf{U}^{-1}$, there is no need to evaluate it explicitly. Instead, based on the fact that the matrix $\mathbf{U}$ is a lower triangular matrix with unit diagonal, the back-substitution method can be employed to solve the following linear system [27]

$$\mathbf{U}\mathbf{A} = \frac{1}{\sigma_v^2}\mathbf{H}^h \quad (33)$$

where $\mathbf{A} \in \mathcal{C}^{NN_b \times M(N_b+L-1)}$ is the unknown matrix to be solved. Once $\mathbf{A}$ is obtained, the feedforward filter $\mathbf{C}$ can be calculated as

$$\mathbf{C} = \mathbf{\Delta}^{-1}\mathbf{A} \quad (34)$$

The low-complexity implementation of the proposed soft-decision MIMO BDFE is summarized in Table I.

TABLE I
LOW-COMPLEXITY SOFT-DECISION MIMO BDFE WITH
RELIABILITY-BASED ORDERING

INPUT:
- Received sample vector $\mathbf{y}$ of an entire block as in (3);
- *A priori* bit LLRs $\left\{\Lambda_a[c_{n,(k-1)\log_2 Q+p}]\right\}$ for the entire block.

INITIALIZATION:
- Compute the *a priori* mean $\{\bar{x}_{n,k}\}$ and variance $\{\sigma_{n,k}^2\}$ for the entire block according to (6a) and (6b);
- Compute the *a priori* reliability $\{\rho_{n,k} = 1/\sigma_{n,k}^2\}$ for the block, and obtain the ordering $\mathbf{O}$ by following the procedure in (8)–(12);
- Perform Cholesky decomposition as $\mathbf{\Sigma}'^{-1} + \frac{1}{\sigma_v^2}\mathbf{H}^h\hat{\mathbf{H}} = \mathbf{U}\mathbf{\Delta}\mathbf{U}^h$ to obtain the feedback filter matrix $\mathbf{D} = \mathbf{U} - \mathbf{I}_{NN_b}$, then obtain the feedforward equalizer matrix $\mathbf{C}$ according to (32)–(34);
- Compute $\mathbf{w} = \mathbf{C}(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}')$, and initialize the soft symbol decision vector $\hat{\mathbf{x}}'$ to be zero.

SOFT-DECISION MIMO BDFE ALGORITHM:
FOR $g = NN_b$ TO 1 DO
- Compute $z_g = w_g + \bar{x}_{\langle O_g \rangle}\mathbf{c}_g^h\mathbf{h}_g$, with w_g, $\mathbf{c}_g^h$, and $\mathbf{h}_g$ as the g-th element, row, and column of $\mathbf{w}$, $\mathbf{C}$, and $\mathbf{H}$, respectively;
- Compute the APP's $\left\{P\left(x_{\langle O_g \rangle} = \chi_q|z_g\right)\right\}_{q=1}^{Q}$ in (27)–(28);
- Compute the bit LLRs $\left\{\Lambda[c_{n_g,(k_g-1)\log_2 Q+p}]\right\}_{p=1}^{\log_2 Q}$ in (30);
- Compute $\hat{x}_{\langle O_g \rangle}$ based on the above APP's, and use it to update the g-th element in $\hat{\mathbf{x}}'$.
END

OUTPUT:
- Bit LLRs $\left\{\Lambda\left[c_{n,(k-1)\log_2 Q+p}\right]\right\}$ of the entire block.

IV. COMPLEXITY ANALYSIS

In this section, the complexity of the proposed soft-decision BDFE with reliability-based ordering is discussed and compared with a soft-decision LE, as a benchmark of low-complexity equalization. An $N \times M$ MIMO block transmission system with a block size N_b is studied. It is assumed that $M \geq N$ and all the MN subchannels have the same length L. For LE, it is assumed that the equalizer has $K = K_1 + K_2 + 1$ tap vectors (each with length M), spanning over the index range $[-K_1, K_2]$. Without loss of generality, let $K_1 = K_2 = \alpha L$ such that $K = 2\alpha L + 1$, where α is a scaling factor reflecting the size of the sliding window. For the fairness of comparison, constant equalizer taps [8] are also adopted for LE. The complexity is measured by the number of complex multiplications (CM).

For both equalization methods, the main complexity arises from three sources: the calculation of the equalizer taps, the filtering operation to obtain symbol estimations, and the calculation of bit LLRs based on the estimated symbols. For the LLR calculation, the procedures in (29)–(30) are the same for both equalizations. The difference only lies in the calculation of the conditional probability in (28), or

equivalently, the conditional mean and variance of the complex Gaussian distribution. For BDFE, the conditional mean and variance have already been obtained during the Cholesky decomposition and the filtering operation. For LE, the calculation of the conditional mean and variance incurs a complexity on the order of $\mathcal{O}(MNKN_b)$. This complexity, however, is small compared to other operations. We next focus on the first two computational sources for both BDFE and LE.

The complexity of calculating the constant feedforward and feedback matrices $\mathbf{C}$ and $\mathbf{D}$ for BDFE is dominated by the Cholesky decomposition of $\boldsymbol{\Sigma}'^{-1} + \frac{1}{\sigma_v^2}\mathbf{H}^h\mathbf{H}$ in (31) and the back-substitution operation in (33), which incur complexities on the order of $\mathcal{O}(N^3N_b^3)$ and $\mathcal{O}(MN^2N_b^2(N_b + L - 1))$, respectively. It is noted that when $\mathbf{H} = \mathbf{F}$ (refer to (3) and (13)) in the first iteration, the block-Toeplitz structure shown in (4) can be utilized to reduce the complexity of Cholesky decomposition to the order of $\mathcal{O}((NN_b)^2)$ [27]. The filtering operation in (25)–(26) incurs a complexity on the order of $\mathcal{O}(MNN_b(N_b + L - 1))$. Combining the two computational complexities together, the overall complexity for the BDFE is dominated by the filter matrices calculation and is on the order of $\mathcal{O}(MN^2N_b^2(N_b + L - 1))$.

For LE, the calculation of the equalizer tap vector involves an inversion of a matrix with size $\min(MK, N(K + L - 1))$, where the minimum operation is due to the use of matrix inversion lemma during the equalizer tap vector calculation [28]. Let $\lambda = \min(MK, N(K + L - 1))$, and the complexity is thus on the order of $\mathcal{O}(\lambda^3)$. The filtering operation over the entire block causes a complexity on the order of $\mathcal{O}(MN^2N_bK(K + L - 1))$. Combining the two complexities together, an overall complexity is on the order of $\mathcal{O}(\max(\lambda^3, MN^2N_bK(K + L - 1)))$.

From the analysis above, the complexity of BDFE is mainly determined by the block size N_b, while the complexity of LE is a function of the channel length L, the parameter K (or α), and the block size N_b. The complexity comparison between BDFE and LE is then determined by the system parameters N_b, L and K. We next compare the complexity between LE and BDFE with three examples.

In the first example, the block size N_b and the scaling factor α are fixed as $N_b = 100$ and $\alpha = 2$, while the channel length L changes over the range $[5, 35]$. In the second example, the block size N_b and the channel length L are fixed as $N_b = 100$ and $L = 15$, while $\alpha \in [1, 5]$. In both examples, a 2×2 MIMO system has been studied. The comparison results for examples 1 and 2 are demonstrated in Fig. 3. In the top subfigure (example 1), the channel length L has negligible effect on the complexity of BDFE, yet the complexity of LE increases by about two orders of magnitude from $L = 5$ to $L = 35$. At $L = 5$, the BDFE complexity is more than one order of magnitude higher than the LE complexity. However, when $L \geq 10$, the two complexities are comparable to each other. When $L \geq 25$, LE has a higher complexity than BDFE. From the bottom subfigure (example 2), the BDFE complexity is independent of α, while the LE complexity increases with the size of the sliding window.

In the third example, the block size N_b changes in the range of $[50, 150]$ and the channel length has two choices: $L = 15$ and $L = 20$. The scaling factor for LE is fixed as $\alpha = 2$. A
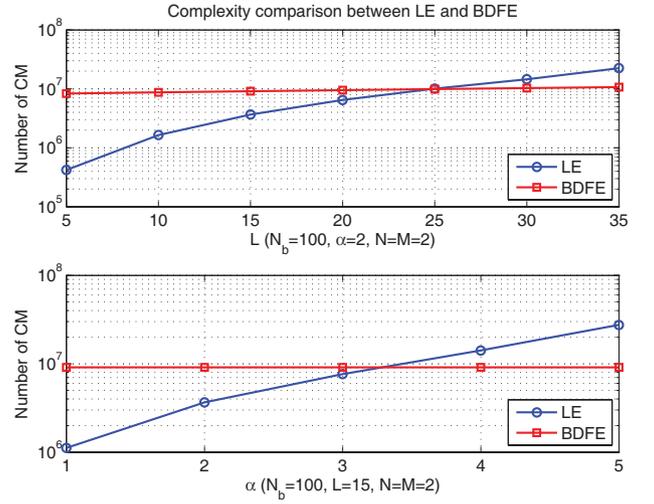

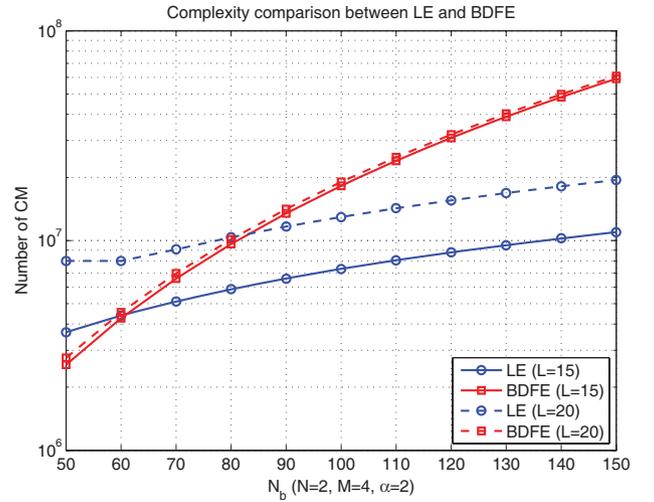Fig. 3.   Complexity comparison between LE and BDFE: examples 1 and 2.

Fig. 4.   Complexity comparison between LE and BDFE: example 3.

2×4 MIMO system is investigated. The comparison results are demonstrated in Fig. 4.

Obviously, the BDFE complexity increases much faster than the LE complexity with the block size N_b. This observation is expected, because the BDFE complexity is cubic in N_b, while the LE complexity is linear in N_b. However, the largest gap between the LE and BDFE complexities is within one order of magnitude over the range $N_b \in [50, 150]$. Since block transmission is primarily designed for fast-fading environment [18], the choice of block size N_b shall not be too large so as to guarantee the time-invariant assumption of the channel. As a result, the proposed BDFE has a comparable complexity to LE. Finally, it is again observed that the channel length L has negligible effects on the BDFE complexity, while the LE complexity increases considerably with L.

V. SIMULATION RESULTS

Simulation results are presented in this section to demonstrate the performance of the proposed reliability-based turbo detection scheme. The data is transmitted in packets. Each packet is independently encoded by a rate 1/2 non-systematic convolutional code with a constraint length 4 and a generator
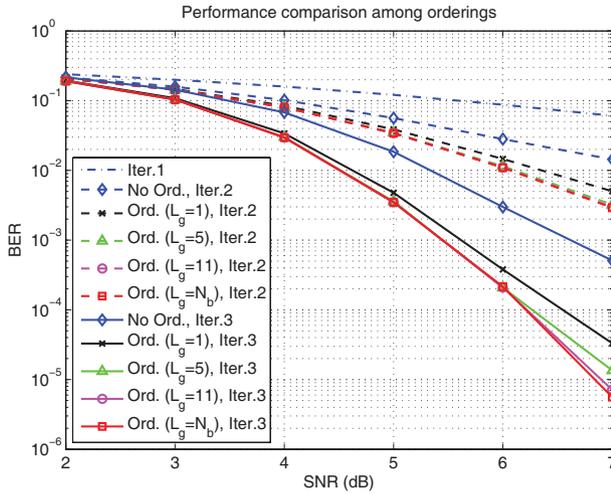
Fig. 5. Performance comparison among the different orderings (2×2 MIMO, Proakis Type A Channel, QPSK).

Fig. 6. Performance comparison between LE and BDFE (2×2 MIMO, uniform PDP with $L = 10$, 8PSK).

polynomial $[G_1, G_2] = [17, 13]_{oct}$. The encoded packet is further divided into blocks for transmission. The soft-decision MIMO BDFE is operated over blocks, and the MAP channel decoding is performed over an entire packet.

First we demonstrate the performance gain contributed by the reliability-based ordering for the BDFE. A 2×2 MIMO system using QPSK modulation is studied. Each packet carries 2,000 symbols, which are divided into 20 blocks with size $N_b = 100$. The MIMO channel is generated with the Proakis Type A channel [25] as $\mathbf{f}_{1,1} = [0.04, -0.05, 0.07, -0.21, -0.5, 0.72, 0.36, 0, 0.21, 0.03, 0.07]$, $\mathbf{f}_{1,2} = \mathbf{f}_{1,1} \ll 3$, $\mathbf{f}_{2,1} = \mathbf{f}_{1,1} \ll 6$, $\mathbf{f}_{2,2} = \mathbf{f}_{1,1} \ll 9$, where $\mathbf{f}_{m,n}$ denotes the subchannel between the n-th transmit antenna and the m-th receive antenna. For the group-wise ordering, different choices of group sizes $L_g = 1$, $L_g = 5$, $L_g = L = 11$, and $L_g = N_b = 100$ have been selected leading to $N_g = 100$, $N_g = 20$, $N_g = 10$ and $N_g = 1$, respectively. The choice of $L_g = N_b$ corresponds to the global ordering over the entire block. It is also noted that for the choice of $L_g = 11$, the last group contains $N \times \mod(N_b, L_g) = 2$ symbols. The simulation results are shown in Fig. 5.

Since there is no *a priori* reliability information available in the first iteration and the bit error rate (BER) curves in all cases overlap, only one curve is plotted for the first iteration. We have two observations from the figure. First, the reliability-based ordering leads to apparent performance gain over the non-ordering case. At the BER level of 10^{-3}, more than 1 dB performance gain is achieved for the choice of $L_g > 1$. When $L_g = 1$, which corresponds to the ordering in the 1-D space domain only, there is still performance gain over the non-ordering case, yet it is inferior to the case that the ordering is performed in the 2-D space-time domain. Second, when $L_g > 1$, different choices of group sizes lead to almost identical performance. Therefore, the group-wise ordering scheme has almost the same performance as the global ordering, but with a lower complexity.

Next we compare the performance of the proposed turbo detection scheme with that of the linear turbo detection scheme [14]. A 2×2 MIMO system with 8PSK modulation
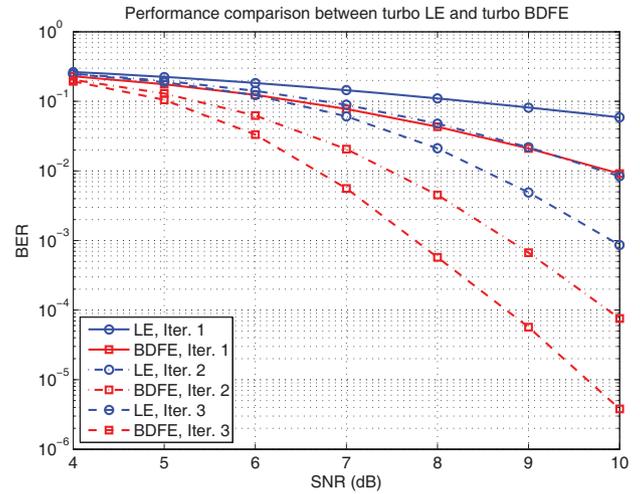
is considered, and the MIMO channel has a uniform PDP with a channel length of $L = 10$. Each packet carries 6,000 symbols, which are divided into 60 blocks with a block size $N_b = 100$. The MIMO channel remains constant over each packet while changes from packet to packet. In the simulation, 6,000 packets have been simulated. A group size $L_g = L = 10$ is used for the reliability-based ordering. The parameters K_1 and K_2 are set as $K_1 = K_2 = L$ for LE. The BER results of the two detection schemes are compared in Fig. 6. From the figure, the proposed detection significantly outperforms the linear turbo detection, and has achieved more than 2 dB gain in all the three iterations.

Fig. 7 compares the performance of the proposed BDFE with the optimum MAP equalizer in a 2×2 MIMO system with BPSK modulation. Each packet carries 2,000 symbols, and the block size is $N_b = 100$. The MIMO channel has a length of $L = 3$, and is generated with the Proakis type B [25] channel as $\mathbf{f}_{1,1} = [0.407, 0.815, 0.407]$, $\mathbf{f}_{1,2} = [0.815, 0.407, 0.407]$, $\mathbf{f}_{2,1} = [0.407, 0.407, 0.815]$, and $\mathbf{f}_{2,2} = [0.407, -0.407, 0.815]$. With a BPSK modulation, there are $Q^{N(L-1)} = 16$ trellis states for the MAP equalization. The result shown in Fig. 7 is simulated with 1,000 packets. It is observed that when the BER equals to 10^{-3}, the proposed BDFE is only about 0.5 dB from the optimum MAP equalizer. In addition, only marginal performance improvement is observed by increasing the number of iterations from 4 to 5. Therefore, the proposed BDFE achieves a BER performance close to the optimum MAP equalizer.

VI. CONCLUSION

A reliability-based turbo detection scheme with successive soft interference cancellation was proposed for MIMO systems with frequency-selective fading. Compared with existing low-complexity turbo detection schemes, the proposed scheme has two advantages. First, it utilized the non-linear structure of the BDFE, which enabled SSIC thus led to a better performance than the linear equalization structure with a similar complexity. Second, it introduced a novel group-wise *reliability-based*
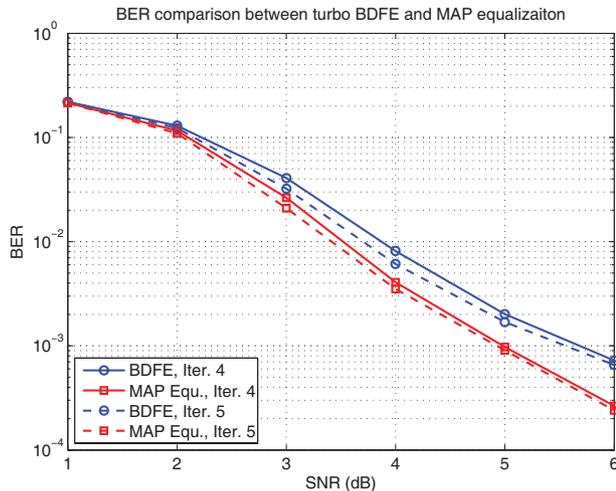
Fig. 7. Performance comparison between BDFE and MAP equalization (2×2 MIMO, Proakis Type B Channel, BPSK).

ordering scheme to minimize the effect of error propagation in SSIC. Compared with conventional ordering schemes, the proposed ordering method has a much lower complexity due to the low cost associated with the calculation of the reliability information, and the low complexity of a group-wise ordering operation. Moreover, the ordering is updated dynamically as the turbo iterations progress. It was demonstrated through simulations that performance gain was achieved with only a small ordering overhead, and the performance of the proposed scheme approached the optimum performance of the MAP detection.

REFERENCES

[1] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: turbo-equalization," *Eur. Trans. Telecommun.*, vol. 6, no. 5, pp. 507-511, Sep. 1995.
[2] G. Bauch, H. Khorram, and J. Hagenauer, "Iterative equalization and decoding in mobile communications systems," in *Proc. EPMCC'97*, Sep. 1997, pp. 307-312.
[3] J. Hagenauer and P. Hoher, "A Viterbi algorithm with soft-decision outputs and its applications," in *Proc. Globecom'89*, Nov. 1989, pp. 1680-1686.
[4] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284-287, Mar. 1974.
[5] F. K. H. Lee and P. J. Mclane, "Parallel-trellis turbo equalizers for sparse-coded transmission over SISO and MIMO sparse multipath channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3568-3578, Dec. 2006.
[6] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046-1061, July 1999.
[7] M. Tuchler, R. Koetter, and A. C. Singer, "Turbo equalization: principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754-767, May 2002.
[8] M. Tuchler, A. C. Singer, and R. Koetter, "Minimum mean square error equalization using *a priori* information," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 673-683, Mar. 2002.
[9] Z. Wu and J. Cioffi, "Low complexity iterative decoding with decision-aided equalization for magnetic recording channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 4, pp. 699-708, Apr. 2001.
[10] C. Laot, A. Glavieux, and J. Labat, "Turbo equalization: adaptive equalization and channel decoding jointly optimized," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1744-1752, Sep. 2001.
[11] C. Laot, R. L. Bidan, and D. Leroux, "Low complexity MMSE turbo equalization: a possible solution for EDGE," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 965-974, May 2005.
[12] S. L. Ariyavisitakful and Y. Li, "Joint coding and decision feedback equalization for broadband wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 9, pp. 1670-1678, Dec. 1998.
[13] R. R. Lopes and J. R. Barry, "The soft-feedback equalier for turbo equalization of highly dispersive channels," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 783-788, May 2006.
[14] T. Abe and T. Matsumoto, "Space-time turbo equalization in frequency selective MIMO channels," *IEEE Trans. Veh. Technol.*, vol. 52, no. 3, pp. 469-475, May 2003.
[15] T. Abe, S. Tomisato, and T. Matsumoto, "A MIMO turbo equalizer for frequency-selective channels with unknown interference," *IEEE Trans. Veh. Technol.*, vol. 52, no. 3, pp. 476-482, May 2003.
[16] G. Bauch and N. Al-dhahir, "Reduced-complexity space-time turbo-equalization for frequency-selective MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 819-828, Oct. 2002.
[17] J. Wu and Y. R. Zheng, "Low complexity soft-input soft-output block decision feedback equalization," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 281-289, Feb. 2008.
[18] G. K. Kaleh, "Channel equalization for block transmission systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 1, pp. 110-121, Jan. 1995.
[19] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Lab. Tech. J.*, vol. 1, pp. 41-59, 1996.
[20] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. IEEE ISSSE'98*, Sep. 1998, pp. 295-300.
[21] Z. Luo, S. Liu, M. Zhao, and Y. Liu, "Generalized parallel interference cancellation algorithm for V-BLAST systems," in *Proc. IEEE Int. Conf. Commun.*, June 2006, vol. 7, pp. 3207-3212.
[22] A. Lozano and C. Papadias, "Layered space-time receivers for frequency-selective wireless channels," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 65-73, Jan. 2002.
[23] S. W. Kim and K. P. Kim, "Log-likelihood-ratio-based detection ordering in V-BLAST," *IEEE Trans. Commun.*, vol. 54, no. 2, pp. 302-307, Feb. 2006.
[24] R. Koetter, A. C. Singer, and M. Tuchler, "Turbo equalization," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 67-80, Jan. 2004.
[25] J.G. Proakis, *Digital Communication*, 4th edition. McGraw-Hill, 2001
[26] A. Stamoulis, G. B. Giannakis, and A. Scaglione, "Block FIR decision-feedback equalizers for filterbank precoded transmissions with blind channel estimation capabilities," *IEEE Trans. Commun.*, vol. 49, pp. 69-83, Jan. 2001.
[27] N. Al-Dhahir and A. H. Sayed, "The finite-length multi-input multi-output MMSE-DFE," *IEEE Trans. Signal Process.*, vol. 48, no. 10, pp. 2921-2936, Oct. 2000.
[28] J. Tao, Y. R. Zheng, C. Xiao, T. C. Yang, and W. B. Yang, "Channel equalization for single carrier MIMO underwater acoustic communications," *EURASIP J. Advances in Signal Process*, vol. 2010, article ID 281769, 17 pages, 2010 (doi:10.1155/2010/281769).

Jun Tao (S'10-M'10) received the B.S. and M.S. degrees in electrical engineering from the Department of Radio Engineering, Southeast University, Nanjing, China, in 2001 and 2004, respectively, and the Ph.D. degree in electrical engineering from the Department of Electrical and Computer Engineering, University of Missouri, Columbia, in 2010.

From 2004 to 2006, he was a System Design Engineer with Realsil Microelectronics Inc. (a subsidiary of Realtek), Suzhou, China. He is currently a Postdoctoral Fellow at the Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla. His research interests include channel modeling, channel estimation, and turbo equalization for wireless radio-frequency communications, and robust signal detection for underwater acoustic communications. He is the holder of one IC chip design patent, which has been commercialized.

Jingxian Wu (S'02-M'06) received the B.S. degree in electronic engineering from Beijing University of Aeronautics and Astronautics, Beijing, China, in 1998, the M.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2001, and the Ph.D. degree in electrical engineering from the University of Missouri, Columbia, in 2005.

He is currently an Assistant Professor with the Department of Electrical Engineering, University of Arkansas, Fayetteville. His research interests mainly focus on wireless communications and wireless networks, including ultra-low power communications, cooperative communications, cognitive radio, and cross-layer optimization, etc. He is currently an Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He served as a Cochair for the 2012 Wireless Communication Symposium of the IEEE International Conference on Communication, and a Cochair for the 2009 Wireless Communication Symposium of the IEEE Global Telecommunications Conference. Since 2006, he has served as a Technical Program Committee Member for a number of international conferences, including the IEEE Global Telecommunications Conference, the IEEE Wireless Communications and Networking Conference, the IEEE Vehicular Technology Conference, and the IEEE International Conference on Communications.

Yahong Rosa Zheng (S'99-M'03-SM'07) received the B.S. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 1987, and the M.S. degree from Tsinghua University, Beijing, China, in 1989, both in electrical engineering. She received the Ph.D. degree from the Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada, in 2002. She was an NSERC Postdoctoral Fellow from 2003 to 2005 at the University of Missouri-Columbia. Since 2005, she has been a faculty member with the Department of Electrical and Computer Engineering at Missouri University of Science and Technology. Her research interests include array signal processing, wireless communications, and wireless sensor networks. She has served as a Technical Program Committee (TPC) member for many IEEE international conferences, including VTC, GlobeCom, ICC, WCNC, and Aerospace Conference, etc. She served as an Associate Editor for IEEE Transactions on Wireless Communications for 2006-2008. She has been a senior member of the IEEE since 2007. She is the recipient of an NSF CAREER award in 2009.