Research Article

Main-Branch Structure Iterative Detection Using Approximate Message Passing for Uplink Large-Scale Multiuser MIMO Systems

Qifeng Zou, Xuezhai Tan, Mei Liu, and Lin Ma

School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150080, China

Correspondence should be addressed to Qifeng Zou; zouqifeng163@163.com

Received 31 October 2015; Revised 4 February 2016; Accepted 7 February 2016

Academic Editor: Wei Liu

Copyright © 2016 Qifeng Zou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The emerging large-scale/massive multi-input multioutput (MIMO) system combined with orthogonal frequency division multiplexing (OFDM) is considered a key technology for its advantage of improving the spectral efficiency. In this paper, we introduce an iterative detection algorithm for uplink large-scale multiuser MIMO-OFDM communication systems. We design a Main-Branch structure iterative turbo detector using the Approximate Message Passing algorithm simplified by linear approximation (AMP-LA) and using the Mean Square Error (MSE) criterion to calculate the correlation coefficients between main detector and branch detector for the given iteration. The complexity of our method is compared with other detection algorithms. The simulation results show that our scheme can achieve better performance than the conventional detection methods and have the acceptable complexity.

1. Introduction

Multi-input multioutput systems with multiple antennas employed at both the transmitter and receiver got a lot of attention due to their multiplexing and diversity capabilities which can offer much higher data rates and enhance the system capacity [1]. OFDM has been shown to be an attractive scheme for mitigating the effects of intersymbol interference (ISI) by broadband wireless channels which have long response memory [2]. In recent years, large-scale (or massive) multiuser MIMO systems have become a hot topic for the next generation 5G wireless communication, which equips tens to hundreds of antennas at the base station as shown in Figure 1. The large-scale multiuser MIMO technology combined with OFDM promises significant improvements in terms of spectral efficiency, link reliability, and coverage compared to conventional small-scale MIMO systems [3]. Unfortunately, the promised benefits of large-scale MIMO come at the cost of significantly increased computational complexity in the BS, especially for the antenna array design and signal processing. A major challenge in uplink communications of broadband multiuser MIMO systems

is to create a receiver algorithm that can efficiently detect the multiple signals, transmitted from multiple antennas of different uplink users [4].

Turbo detection which performs detection/equalization and decoding in an iterative manner in coded data transmission over ISI channels have been widely studied [5]. The maximum a posteriori probability (MAP) can achieve optimal performance but the complexities are exponential in N_T , where N_T is the number of transmit antennas [6]. To reduce the complexity, the detector using soft interference cancellation (SIC) based on the Minimum Mean Square Error (MMSE) criterion for MIMO system was proposed in [7, 8]. MMSE-SIC has lower complexity than MAP algorithm but the accuracy is poor because of its structure without the feedback design. Moreover, the complexity of these suboptimal detectors remains prohibitive for the large-scale antenna array. Recently in ten years, message passing based on factor graphs models [9, 10] has been studied for detection/equalization on ISI channels. In [11], Kaynak et al. proposed a belief propagation (BP) algorithm, which has better performance than MMSE-SIC; however the computational complexity is still very large because of the

FIGURE 1: Large-scale multiuser MIMO communication system model.

marginalization operations over discrete symbols. In [12], a variant BP algorithm based on Gaussian tree approximation has been proposed recently, which approximates the dense factor graph of the MIMO system into a tree and passes exact messages over the resultant tree. More recently, Wu et al. proposed a relatively low complexity iterative detection algorithm for large-scale multiuser MIMO-OFDM systems using Approximate Message Passing in [13], but the performance is lower than conventional MMSE detection algorithm when the number of iterations is small. Besides those algorithms derived from factor graph, iterative BP algorithm based on Markov random field (MRF) was also investigated in [14].

In this paper, we propose a new iterative detection algorithm by using several signal processing methods for uplink multiuser MIMO systems. First, we design a Main-Branch structure for iterative signal detection. The soft information is parallel iterating between main detector and branch detector when the LLRs iterate between main detector and decoders. The Main-Branch structure detectors are allowed to exchange soft information in the absence of interleavers and promote the performance of suboptimal constituent detectors. Second, the principle of expectation propagation and linear approximation method are applied to obtain the symbol belief and approximate the Gaussian messages for reducing computing complexity. Finally, we employ AMP-LA algorithm as the constituent detector of the Main-Branch structure, and using the Mean Square Error (MSE) criterion to calculate the

correlation coefficients between main detector and branch detector for the given iteration, we also give the LLRs computing between Main-Branch structure detector and decoders. The proposed scheme not only gains better performance by canceling the residual ISI and CCI but also simplifies the computational complexity of the detectors for avoiding the problem of matrix inversion on each frequency bin involved in the conventional MMSE detection algorithm. Simulation results show that the proposed algorithm can achieve better tradeoffs between performance and complexity than existing turbo detection algorithms and approaches the optimal performance with a small number of iterations.

The rest of the paper is organized as follows. The system model is described in Section 2, and the proposed Main-Branch iterative structure with AMP algorithm is discussed in Section 3. In Section 4, novel iterative detection based Main-Branch structure joint AMP-LA algorithm is proposed. We will show the simulation results along with a discussion in Section 5. And Section 6 concludes the paper.

Notation. The operators $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ denote transpose, complex conjugate, and complex conjugate transpose, respectively. $\mathbf{I}_N$ represents the identity matrix of size $N \times N$, and $\ln(\cdot)$ denotes the natural logarithm; $E_{P(X)}(\cdot)$ is the expectation operation with respect to the probability distribution $P(X)$. The operation $f(X) \propto g(X)$ denotes the relation of $f(X) = cg(X)$ for some positive c . $\mathbf{X} \setminus X_n$ denotes the symbols in $\mathbf{X}$

FIGURE 2: Block diagram of the coded large-scale multiuser MIMO-OFDM system with iterative detector at the receiver.

with X_n excluded; $\mathcal{N}_C(X; \hat{X}, \hat{\tau}) \triangleq (\pi\hat{\tau})^{-1} \exp(-|X - \hat{X}|^2/\hat{\tau})$ denotes the circular Gaussian pdf with mean $\hat{X}$ and variance $\hat{\tau}$.

2. System Model and Iterative Receiver

We consider the uplink of a large-scale multiuser MIMO system with N_T single antenna users, and the receiver is equipped with N_R antennas at the base station (BS). At the transmitter, the information bits are first encoded into a code sequence and interleaved; each Q interleaved coded bit is mapped into one symbol. Let the transmitted frequency domain symbols vector by the n th user be $\mathbf{X}_n(k) = [X_n(1), X_n(2), \dots, X_n(K)]^T$, where $X_n(k) \in \mathcal{A}$ is the frequency domain symbol transmitted at the k th subcarrier and K is the number of the subcarriers in the OFDM system; here $\mathcal{A} = \{\alpha_1, \alpha_2, \dots, \alpha_s, \dots, \alpha_{2Q}\}$ denotes the modulation constellation set. The large-scale multiuser MIMO system uses the OFDM modulation technique at the transmitters. Suppose the maximum taps number of MIMO intersymbol interference (ISI) channel is L . We make K point IFFT to symbol sequence $\mathbf{X}_n(k)$, and the cyclic prefix (CP) with length $\nu \geq L$ is inserted at the beginning of the OFDM signals to remove the interblock interference and make the linear convolution to a circular convolution, as shown in Figure 2. Then the OFDM modulated signals are sent through the MIMO-ISI channel. After removing CP, the FFT technique is used to transform received signal from time domain to frequency domain at the receiver. Then the channel estimation algorithm in frequency domain provides the current channel state knowledge to the detector. As a core

part of this large-scale multiuser MIMO system, the proposed iteration detector at the receiver can suppress cochannel interference (CCI).

Hence the received N_R -dimensional base band signal vector at the k th subcarrier in frequency domain can be written as

$$\mathbf{Y}(k) = \mathbf{H}(k) \mathbf{X}(k) + \mathbf{W}(k), \quad (1)$$

where $\mathbf{X}(k) \in \mathbb{C}^{N_T \times 1}$ denotes all the transmitted frequency domain symbols associated with the k th subcarrier and $\mathbf{W}(k)$ denotes a N_R -dimensional stationary Additive White Gaussian Noise vector in frequency domain with zero-mean and covariance matrix $\sigma_W^2 \mathbf{I}$. $\mathbf{H}(k)$ is a $N_R \times N_T$ matrix representing the MIMO channel frequency response at the k th subcarrier, which is given by

$$\mathbf{H}(k) = \sum_{l=0}^{L-1} \mathbf{h}_l \exp\left(\frac{-j2\pi l k}{K}\right), \quad (2)$$

where $\mathbf{h}_l$ is a $N_R \times N_T$ time-domain block circulant channel matrix with the entry $h_l^{(m,n)}$ being the l th channel impulse response taps from the n th user transmit antenna to the m th receiving antenna.

The matrix form of $\mathbf{H}(k)$ can be denoted as

$$\mathbf{H}(k) = \begin{bmatrix} H_{1,1}(k) & \cdots & H_{1,N_T}(k) \\ \vdots & \ddots & \vdots \\ H_{N_R,1}(k) & \cdots & H_{N_R,N_T}(k) \end{bmatrix}. \quad (3)$$

3. Iterative Detection Receiver and Approximate Message Passing Algorithm

3.1. Turbo Main-Branch Iterating Detection Structure. The iterative process by Main-Branch structure soft detector in the system is shown in Figure 2. Turbo MIMO detection exchanges extrinsic log likelihood ratios (LLRs) of the coded bits between a Main-Branch iterating structure part and a bank of channel decoders. The Main-Branch iterative part consists of one soft-input soft-output (SISO) main detector and one SISO branch detector; the soft information is parallel iterating between main detector and branch detector when the LLRs iterate between main detector and decoders. We will omit the subcarrier index k for notational simplicity.

Unlike the extrinsic information exchanged between main detector and channel decoders, the extrinsic information between the main detector and the branch detector has significant correlation because no interleaving can be applied. On the other hand, detection algorithm can perform iteratively by Main-Branch structure soft detector without the decoder, called “self-iteration,” so it can be used in uncoded systems also. For the coded systems, the proposed turbo detection algorithm is established in Figure 2 to adapt common circumstances. At the i th iteration, the received signal $\mathbf{Y}$ is detected by the main detector and its extrinsic LLR $L_{e,M}(c_n^q)$ of each coded bit c_n^q corresponding to the symbol X_n is passed to the branch detector, to be used to generate its *a priori* information $L_{a,b}(c_n^q)$. At the same time the extrinsic information generated by the main detector is passed to channel decoders as their *a priori* information directly. The branch detector in turn generates its own extrinsic information $L_{e,b}(c_n^q)$ with the given *a priori* information $L_{a,b}(c_n^q)$. For the significant correlation between $L_{e,b}(c_n^q)$ and $L_{a,b}(c_n^q)$, they combine together to generate the *a priori* information $L_{a,M}(c_n^q)$. $L_{a,M}(c_n^q)$ combined with the *a priori* information $L_a(c_n^q)$ fed back from the decoders generate extrinsic LLRs of the main detector for the next detection process. The extrinsic LLRs of the main detector are given by

$$L_{e,M}(c_n^q) = \ln \frac{P(c_n^q = 1 | \mathbf{Y})}{P(c_n^q = 0 | \mathbf{Y})} - L(c_n^q), \quad (4)$$

where $\ln(P(c_n^q = 1 | \mathbf{Y})/P(c_n^q = 0 | \mathbf{Y}))$ is the *a posteriori* LLR at the output of the main detector and $L(c_n^q)$ represents the total *a priori* information at the input of the main detector, given as

$$L(c_n^q) = L_{a,M}(c_n^q) + L_a(c_n^q), \quad (5)$$

where $L_a(c_n^q) \triangleq \ln(P(c_n^q = 1)/P(c_n^q = 0))$ represents the *a priori* LLRs passed down from the decoders, and $L_{a,M}(c_n^q)$ represents the *a priori* LLRs passed from the branch detector.

The extrinsic LLRs of the branch detector are given by

$$L_{e,b}(c_n^q) = \ln \frac{P'(c_n^q = 1 | \mathbf{Y})}{P'(c_n^q = 0 | \mathbf{Y})} - L_{a,b}(c_n^q), \quad (6)$$

where $\ln(P'(c_n^q = 1 | \mathbf{Y})/P'(c_n^q = 0 | \mathbf{Y}))$ is the *a posteriori* LLR at the output of the branch detector and $L_{a,b}(c_n^q)$ represents the *a priori* information at the input of the branch detector.

The problem is how to adopt extrinsic information $L_{e,M}(c_n^q)$ to generate $L_{a,b}(c_n^q)$ and adopt $L_{e,b}(c_n^q)$ to generate $L_{a,M}(c_n^q)$. We can regard the *a priori* LLRs and the extrinsic LLRs as the output of an equivalent AWGN channel by the method of [15]. The unbiased versions of these LLRs associate with the branch detector as the transmitted symbol $\mathbf{X}$ corrupted by AWGN:

$$\begin{aligned} \mathbf{Y}_{a,b} &= \mathbf{X} + \mathbf{u}_{a,b}, \\ \mathbf{Y}_{e,b} &= \mathbf{X} + \mathbf{u}_{e,b}, \end{aligned} \quad (7)$$

where $\mathbf{u}_{a,b}$ and $\mathbf{u}_{e,b}$ are assumed to be zero-mean Gaussian random vector with zero-mean and covariance matrix $\sigma_{a,b}^2 \mathbf{I}$ and $\sigma_{e,b}^2 \mathbf{I}$, which are independent of the transmitted data $\mathbf{X}$, but correlated with each other with correlation coefficient ρ_b . The *a priori* information corresponding to the symbol X_n passing to the main detector can be defined as

$$\begin{aligned} L_{a,M}(c_n^q) &= \ln \frac{P(\mathbf{Y}_{e,b}, \mathbf{Y}_{a,b} | c_n^q = 1)}{P(\mathbf{Y}_{e,b}, \mathbf{Y}_{a,b} | c_n^q = 0)} \\ &\quad - \ln \frac{P(\mathbf{Y}_{a,b} | c_n^q = 1)}{P(\mathbf{Y}_{a,b} | c_n^q = 0)} \\ &= \left(\frac{1 - \rho_b \sigma_{e,b}/\sigma_{a,b}}{1 - \rho_b^2} \right) L_{e,b}(c_n^q) \\ &\quad + \left(\frac{\rho_b^2 - \rho_b \sigma_{a,b}/\sigma_{e,b}}{1 - \rho_b^2} \right) L_{a,b}(c_n^q). \end{aligned} \quad (8)$$

We assume that the variance $\sigma_{a,b}^2 = \sigma_{e,b}^2$, and then we can construct the *a priori* LLRs for the main equalizer based on two correlated LLRs $L_{e,b}(c_n^q)$ and $L_{a,b}(c_n^q)$ as

$$L_{a,M}(c_n^q) = \left(\frac{1}{1 + \rho_b} \right) L_{e,b}(c_n^q) + \left(\frac{\rho_b}{1 + \rho_b} \right) L_{a,b}(c_n^q). \quad (9)$$

In order to accurately measure the correlation of the *a priori* LLRs $L_{a,M}(c_n^q)$ and the extrinsic information $L_{e,b}(c_n^q)$, we force a linear relationship $\hat{L}_{a,M}(c_n^q) = \lambda_1 L_{e,b}(c_n^q)$, where λ_1 is a positive scaling factor, and another linear relationship $\hat{L}_{a,b}(c_n^q) = \lambda_2 L_{e,M}(c_n^q)$ is also forced to measure the correlation of the *a priori* LLRs $L_{a,b}(c_n^q)$ and the extrinsic information $L_{e,M}(c_n^q)$.

The aim of coefficients λ_1 computing is to minimize the error between $L_{a,M}(c_n^q)$ and $\hat{L}_{a,M}(c_n^q)$. So we can denote the Mean Square Error (MSE) function as

$$\text{MSE} = E \left[\left| L_{a,M}(c_n^q) - \hat{L}_{a,M}(c_n^q) \right|^2 \right]. \quad (10)$$

We also can get the MSE between $L_{a,b}(c_n^q)$ and $\hat{L}_{a,b}(c_n^q)$ like (10), evaluating the derivative of Mean Square Error function with respect to the coefficients λ and setting it to zero; according to the theory in [16], we can get $\lambda_1 \approx (1 - \rho_b)/(1 + \rho_b)$ and $\lambda_2 \approx (1 - \rho_M)/(1 + \rho_M)$, and two linear correlation evaluation equations are given as

$$L_{a,M}(c_n^q) = \left(\frac{1 - \rho_b}{1 + \rho_b} \right) L_{e,b}(c_n^q), \quad (11)$$

$$L_{a,b}(c_n^q) = \left(\frac{1 - \rho_M}{1 + \rho_M} \right) L_{e,M}(c_n^q), \quad (12)$$

where ρ_b is the noise correlation parameter between $L_{e,b}(c_n^q)$ and $L_{a,b}(c_n^q)$ and ρ_M represents the other noise correlation parameter between $L_{e,M}(c_n^q)$ and $L_{a,M}(c_n^q)$, so the noise correlation parameters ρ_b and ρ_M can be estimated as [17, 18]

$$\hat{\rho}_b = \frac{\sum_{n=1}^N \{(L_{a,b}(c_n^q) - \text{sign}[L_{a,b}(c_n^q)] \theta_{a,b}) (L_{e,b}(c_n^q) - \text{sign}[L_{e,b}(c_n^q)] \theta_{e,b})\}}{\sqrt{\sum_{n=1}^N (L_{a,b}(c_n^q) - \text{sign}[L_{a,b}(c_n^q)] \theta_{a,b})^2} \sqrt{\sum_{n=1}^N (L_{e,b}(c_n^q) - \text{sign}[L_{e,b}(c_n^q)] \theta_{e,b})^2}}, \quad (13)$$

$$\hat{\rho}_M = \frac{\sum_{n=1}^N \{(L_{a,M}(c_n^q) - \text{sign}[L_{a,M}(c_n^q)] \theta_{a,M}) (L_{e,M}(c_n^q) - \text{sign}[L_{e,M}(c_n^q)] \theta_{e,M})\}}{\sqrt{\sum_{n=1}^N (L_{a,M}(c_n^q) - \text{sign}[L_{a,M}(c_n^q)] \theta_{a,M})^2} \sqrt{\sum_{n=1}^N (L_{e,M}(c_n^q) - \text{sign}[L_{e,M}(c_n^q)] \theta_{e,M})^2}}, \quad (14)$$

where $\theta_{a,b} = E[L_{a,b}(c_n^q) | c_n^q = 1]$, $\theta_{e,b} = E[L_{e,b}(c_n^q) | c_n^q = 1]$, $\theta_{a,M} = E[L_{a,M}(c_n^q) | c_n^q = 1]$, $\theta_{e,M} = E[L_{e,M}(c_n^q) | c_n^q = 1]$, and $\text{sign}(\cdot)$ is the Sign function, which can be defined as

$$\text{sign}(t) = \begin{cases} 1, & t \geq 0, \\ -1, & t < 0. \end{cases} \quad (15)$$

3.2. Factor Graph and Approximate Message Passing Algorithm. The detector generates the extrinsic LLRs of c_n^q based on the received signal $\mathbf{Y}$ and the *a priori* LLRs can be written as

$$\begin{aligned} L_e(c_n^q) &= \ln \frac{\sum_{X_n \in \mathcal{A}_q^1} P(X_n | \mathbf{Y})}{\sum_{X_n \in \mathcal{A}_q^0} P(X_n | \mathbf{Y})} - L(c_n^q) \\ &= \ln \frac{\sum_{X_n \in \mathcal{A}_q^1} P(X_n) P(\mathbf{Y} | X_n)}{\sum_{X_n \in \mathcal{A}_q^0} P(X_n) P(\mathbf{Y} | X_n)} - L(c_n^q). \end{aligned} \quad (16)$$

For the presentation of factor and message passing algorithm in [9, 10], we obtain the joint distribution probability as

$$\begin{aligned} P(\mathbf{X}, \mathbf{Y}) &= P(\mathbf{X}) P(\mathbf{Y} | \mathbf{X}) \\ &= \prod_n P(X_n) \prod_m f_m(Y_m | \mathbf{X}), \end{aligned} \quad (17)$$

where $f_m(Y_m | \mathbf{X})$ denotes the channel transition function, which can be defined as

$$f_m(Y_m | \mathbf{X}) = \frac{1}{\pi \sigma_W^2} \exp \left(-\frac{|Y_m - \sum_m H_{m,n} X_n|^2}{\sigma_W^2} \right), \quad (18)$$

where $H_{m,n}$ is the component of $\mathbf{H}$ in the m row and n column.

Let us investigate the message passing algorithm on the factor graph model in Figure 3, where ϕ_n represents the mapping constraint function $\delta[\phi_n(c_n), x_n]$ and $\delta[\cdot]$ is the Kronecker delta function. “=” denotes the cloning node of variable, and we must note that the code constraints cross all the subcarriers. We assume that the pass messages are from the top to the bottom of Figure 3 and back immediately, so we can avoid inner iteration. When extrinsic information is updated and passed downward, the new iteration will start.

Using the sum-product updating rule, for the i th turbo iteration, the message passed from the channel transition node f_m to the cloning node of X_n is given by

$$\mu_{f_m \rightarrow X_n}^{(i)}(X_n) = \sum_{\mathbf{X} \setminus X_n} f_m(Y_m | \mathbf{X}) \prod_{n' \neq n} \mu_{X_{n'} \rightarrow f_m}^{(i)}(X_{n'}). \quad (19)$$

In the opposite direction, the message passed from the cloning node to the channel transition node is given by

$$\mu_{X_n \rightarrow f_m}^{(i)}(X_n) = \mu_{\phi_n \rightarrow X_n}^{(i)}(X_n) \prod_{m' \neq m} \mu_{f_{m'} \rightarrow X_n}^{(i-1)}(X_n), \quad (20)$$

where $\mu_{\phi_n \rightarrow X_n}^{(i)}(X_n)$ is the message passed from the mapping node to the cloning node, which can be defined as

$$\mu_{\phi_n \rightarrow X_n}^{(i)}(X_n) = \prod_q \frac{\exp[c_n^q L^{(i)}(c_n^q)]}{1 + \exp[L^{(i)}(c_n^q)]} = \prod_q P^{(i)}(c_n^q), \quad (21)$$

where $P^{(i)}(c_n^q) = \exp[c_n^q L^{(i)}(c_n^q)] / (1 + \exp[L^{(i)}(c_n^q)])$ is the *a priori* probability.

Let us assume that X_n is continuous random variables, and approximating the message $\mu_{X_n \rightarrow f_m}^{(i)}(X_n)$ by the estimation of a complex Gaussian probability density function $\hat{\mu}_{X_n \rightarrow f_m}^{(i)}(X_n) = \mathcal{N}_{\mathbb{C}}(X_n; \hat{X}_{X_n \rightarrow f_m}^{(i)}, \hat{\tau}_{X_n \rightarrow f_m}^{(i)})$, the integration form of (19) is defined as

$$\begin{aligned} \mu_{f_m \rightarrow X_n}^{(i)}(X_n) &= \int_{\mathbf{X} \setminus X_n} f_m(Y_m | \mathbf{X}) \\ &\quad \cdot \prod_{n' \neq n} \mathcal{N}_{\mathbb{C}}(X_{n'}; \hat{X}_{X_{n'} \rightarrow f_m}^{(i)}, \hat{\tau}_{X_{n'} \rightarrow f_m}^{(i)}) \\ &= \mathcal{N}_{\mathbb{C}}(H_{m,n} X_n; z_{f_m \rightarrow X_n}^{(i)}, v_{f_m \rightarrow X_n}^{(i)}), \end{aligned} \quad (22)$$

where

$$\begin{aligned} z_{f_m \rightarrow X_n}^{(i)} &= Y_m - \sum_{n' \neq n} H_{m,n'} \hat{X}_{X_{n'} \rightarrow f_m}^{(i)}, \\ v_{f_m \rightarrow X_n}^{(i)} &= \sigma_W^2 + \sum_{n' \neq n} |H_{m,n'}|^2 \hat{\tau}_{X_{n'} \rightarrow f_m}^{(i)}. \end{aligned} \quad (23)$$

FIGURE 3: Factor graph model of multiuser MIMO communication system at one subcarrier.

And the parameters $\widehat{X}_{X_n \rightarrow f_m}^{(i)}$, $\widehat{\tau}_{X_n \rightarrow f_m}^{(i)}$ are obtained by the criterion of minimum KL divergence in [19] as

$$\begin{aligned} \widehat{X}_{X_n \rightarrow f_m}^{(i)} &= \sum_{\alpha_s \in \mathcal{A}} \alpha_s \mu_{X_n \rightarrow f_m}^{(i)}(X_n = \alpha_s), \\ \widehat{\tau}_{X_n \rightarrow f_m}^{(i)} &= \sum_{\alpha_s \in \mathcal{A}} |\alpha_s|^2 \mu_{X_n \rightarrow f_m}^{(i)}(X_n = \alpha_s) - \left| \widehat{X}_{X_n \rightarrow f_m}^{(i)} \right|^2. \end{aligned} \quad (24)$$

In practice, we can use symbol belief by the principle of expectation propagation in [20–22] at the i th iteration to get the approximate message $\widehat{\mu}_{X_n \rightarrow f_m}^{(i)}$, and from the factor graph theory we also find that the symbol belief is the *a posteriori* probability of the symbol X_n , so symbol belief can be approximately given by

$$\begin{aligned} \beta^{(i)}(X_n) &= \frac{\mu_{\phi_n \rightarrow X_n}^{(i-1)}(X_n) \prod_m \mu_{f_m \rightarrow X_n}^{(i-1)}(X_n)}{\sum_{X_n \in \mathcal{A}} \mu_{\phi_n \rightarrow X_n}^{(i-1)}(X_n) \prod_m \mu_{f_m \rightarrow X_n}^{(i-1)}(X_n)} \\ &= \frac{\mu_{\phi_n \rightarrow X_n}^{(i)}(X_n) \mu_{X_n \rightarrow \phi_n}^{(i-1)}(X_n)}{\sum_{X_n \in \mathcal{A}} \mu_{\phi_n \rightarrow X_n}^{(i)}(X_n) \mu_{X_n \rightarrow \phi_n}^{(i-1)}(X_n)} \\ &= \widetilde{P}^{(i)}(X_n^q) = \prod_q \widetilde{P}^{(i)}(c_n^q), \end{aligned} \quad (25)$$

where $\widetilde{P}^{(i)}(c_n^q)$ is the *a posteriori* probability of the coded bit fed back from the decoders. $\mu_{X_n \rightarrow \phi_n}^{(i)}(X_n)$ can be regarded as the multiplication of all the incoming messages by

$$\begin{aligned} \mu_{X_n \rightarrow \phi_n}^{(i)}(X_n) &= \prod_m \mu_{f_m \rightarrow X_n}^{(i)}(X_n) \\ &\propto \mathcal{N}_{\mathbb{C}}(X_n; \xi_{X_n}^{(i)}, \gamma_{X_n}^{(i)}), \end{aligned} \quad (26)$$

where

$$\begin{aligned} \xi_{X_n}^{(i)} &= \gamma_{X_n}^{(i)} \sum_m \frac{H_{m,n}^* z_{f_m \rightarrow X_n}^{(i)}}{v_{f_m \rightarrow X_n}^{(i)}}, \\ \gamma_{X_n}^{(i)} &= \left[\sum_m |H_{m,n}|^2 (v_{f_m \rightarrow X_n}^{(i)})^{-1} \right]^{-1}. \end{aligned} \quad (27)$$

By the expectation propagation (EP) principle and the canonical form of Gaussian PDF by [23], the approximate message $\widehat{\mu}_{X_n \rightarrow f_m}^{(i)}$ can be given as

$$\begin{aligned} \widehat{\mu}_{X_n \rightarrow f_m}^{(i)}(X_n) &= \mathcal{N}_{\mathbb{C}}(X_n; \widehat{X}_{X_n \rightarrow f_m}^{(i)}, \widehat{\tau}_{X_n \rightarrow f_m}^{(i)}) \\ &\propto \beta^{(i)}(X_n) [\mu_{f_m \rightarrow X_n}^{(i-1)}(X_n)]^{-1} \\ &\approx \widehat{\beta}^{(i)}(X_n) [\mu_{f_m \rightarrow X_n}^{(i-1)}(X_n)]^{-1}, \end{aligned} \quad (28)$$

where

$$\begin{aligned} \widehat{X}_{X_n \rightarrow f_m}^{(i)} &= \widehat{\tau}_{X_n \rightarrow f_m}^{(i)} \left[\widehat{X}_{X_n}^{(i)} (\widehat{\tau}_{X_n}^{(i)})^{-1} \right. \\ &\quad \left. - (v_{f_m \rightarrow X_n}^{(i-1)})^{-1} H_{m,n}^* z_{f_m \rightarrow X_n}^{(i-1)} \right], \\ \widehat{\tau}_{X_n \rightarrow f_m}^{(i)} &= \left[(\widehat{\tau}_{X_n}^{(i)})^{-1} - (v_{f_m \rightarrow X_n}^{(i-1)})^{-1} |H_{m,n}|^2 \right]^{-1}, \\ \widehat{\beta}^{(i)}(X_n) &= \mathcal{N}_{\mathbb{C}}(X_n; \widehat{X}_{X_n}^{(i)}, \widehat{\tau}_{X_n}^{(i)}), \end{aligned} \quad (29)$$

where $\widehat{X}_{X_n}^{(i)}$ denotes the mean and $\widehat{\tau}_{X_n}^{(i)}$ denotes the variance; they are defined as

$$\begin{aligned}\widehat{X}_{X_n}^{(i)} &= \sum_{\alpha_s \in \mathcal{A}} \alpha_s \widetilde{P}^{(i)}(X_n = \alpha_s) = E_{\widetilde{P}^{(i)}(X_n)}(X_n), \\ \widehat{\tau}_{X_n}^{(i)} &= \sum_{\alpha_s \in \mathcal{A}} |\alpha_s|^2 \widetilde{P}^{(i)}(X_n = \alpha_s) - \left| \widehat{X}_{X_n}^{(i)} \right|^2 \\ &= E_{\widetilde{P}^{(i)}(X_n)}[|X_n|^2] - \left| \widehat{X}_{X_n}^{(i)} \right|^2.\end{aligned}\quad (30)$$

In order to reduce the complexity, we use the linear approximation (LA) in [13] to simplify the parameters; if the number of users N_T is large, we can get

$$\widehat{\tau}_{X_n \rightarrow f_m}^{(i)} \approx \widehat{\tau}_{X_n}^{(i)}, \quad (31)$$

$$\nu_{f_m \rightarrow X_n}^{(i)} \approx \nu_{f_m}^{(i)} = \sigma_W^2 + \sum_n |H_{m,n}|^2 \widehat{\tau}_{X_n}^{(i)}, \quad (32)$$

$$\begin{aligned}z_{f_m \rightarrow X_n}^{(i)} &\approx z_{f_m}^{(i)} + H_{m,n} \widehat{X}_{X_n}^{(i)} \\ &\quad - \widehat{\tau}_{X_n}^{(i)} \left[\nu_{f_m}^{(i-1)} \right]^{-1} |H_{m,n}|^2 z_{f_m \rightarrow X_n}^{(i-1)}.\end{aligned}\quad (33)$$

Now, by using the first-order linear approximation and Taylor formula in (27), (33) and after some mathematical simplifications, we can get

$$\begin{aligned}z_{f_m}^{(i)} &\approx Y_m - \sum_n H_{m,n} \widehat{X}_{X_n}^{(i)} \\ &\quad + z_{f_m}^{(i-1)} \left(\nu_{f_m}^{(i)} \right)^{-1} \sum_{n'} \widehat{\tau}_{X_{n'}}^{(i)} |H_{m,n'}|^2,\end{aligned}\quad (34)$$

$$\xi_{X_n}^{(i)} \approx \widehat{X}_{X_n}^{(i)} + \gamma_{X_n}^{(i)} \sum_n H_{m,n}^* z_{f_m}^{(i)} \left(\nu_{f_m}^{(i)} \right)^{-1}, \quad (35)$$

$$\gamma_{X_n}^{(i)} = \left[\sum_m |H_{m,n}|^2 \left(\nu_{f_m}^{(i)} \right)^{-1} \right]^{-1}. \quad (36)$$

Finally, the LLRs of coded bits of signal X_n in frequency domain can be expressed in form as

$$\begin{aligned}L_e^{(i)}(c_n^q) &= \ln \frac{\sum_{\mathcal{A}_q^1} \mathcal{N}_{\mathbb{C}}(X_n; \xi_{X_n}^{(i)}, \gamma_{X_n}^{(i)}) \prod_{q' \neq q} \widetilde{P}^{(i)}(c_n^{q'})}{\sum_{\mathcal{A}_q^0} \mathcal{N}_{\mathbb{C}}(X_n; \xi_{X_n}^{(i)}, \gamma_{X_n}^{(i)}) \prod_{q' \neq q} \widetilde{P}^{(i)}(c_n^{q'})} \\ &= \ln \frac{\sum_{\mathcal{A}_q^1} \widetilde{P}_{eq}^{(i)}(X_n)}{\sum_{\mathcal{A}_q^0} \widetilde{P}_{eq}^{(i)}(X_n)} - L^{(i)}(c_n^q),\end{aligned}\quad (37)$$

where

$$\widetilde{P}_{eq}^{(i)}(X_n) = \mu_{\phi_n \rightarrow X_n}^{(i)}(X_n) \mathcal{N}_{\mathbb{C}}(X_n; \xi_{X_n}^{(i)}, \gamma_{X_n}^{(i)}). \quad (38)$$

The detailed processing procedure of AMP algorithm simplified by linear approximation (AMP-LA) is illustrated in Algorithm 1.

4. Proposed Iterative Detection Algorithm and Complexity Analysis

4.1. Iterating Soft Turbo Detection Using Approximate Message Passing. We choose the main detector and the branch detector both to apply the AMP-LA algorithm, so the processing procedure will be simple. The iteration between main detector and branch detector can be called Main-Branch self-iteration, while the turbo iteration between main detector and decoders can be seen as outer iteration. Total iteration numbers of the proposed algorithm equal the numbers of outer iterations. The Main-Branch self-iteration and the outer iteration are performed in parallel. At the specified number of outer iterations i , combining with the *a priori* LLR from the decoders and the branch detector for the main detector, we have $L^{(i)}(c_n^q) = L_{a,M}^{(i)}(c_n^q) + L_a^{(i)}(c_n^q)$. When the main detector generates the extrinsic LLRs $L_{e,M}^{(i)}(c_n^q)$ by Algorithm 1, $L_e^{(i)}(c_n^q) = L_{e,M}^{(i)}(c_n^q)$ for (16). We pass the extrinsic information and the correlation compensated extrinsic information of the main detector to the decoders and to the branch detector, respectively. Likewise, when the branch detector generates the extrinsic LLRs $L_{e,b}^{(i)}(c_n^q)$ by Algorithm 1, $L_e^{(i)}(c_n^q) = L_{e,b}^{(i)}(c_n^q)$ and $L^{(i)}(c_n^q) = L_{a,b}^{(i)}(c_n^q)$ for (16). The main steps of proposed iterative detection structure joint Approximate Message Passing are described as in Algorithm 2.

4.2. Complexity Analysis. Actually, the computational complexity of different algorithms is mainly measured by the numbers of floating-point operations (FLOPs) used for multiplications. We can find that the N point multiplications of a complex number and a real number need $2N$ point FLOPs, and the N point multiplications of two complex numbers require $6N$ point FLOPs operations. In this paper, the complexity of channel estimation and decoder is not considered because the different detection schemes will estimate the channel and decode using the same amount of computations.

We compare the average complexity of the proposed detection algorithm and other algorithms in Table 1; $O(\cdot)$ is the Big-O notation expressing the complexity of an algorithm as a function of a given input.

The traditional MMSE-SIC frequency domain algorithm in [7] has to estimate the complexity of the extrinsic mean and variances of X_i . We let N_T be the number of users, and N_R is the numbers of receiver antennas. The MMSE-SIC method has the largest computational complexity compared to AMP using Gaussian approximation (AMP-G), AMP-EP, and AMP-LA but it got better performance improvement than them also. All the message passing algorithms require $3N_T N_R$ FLOPs to compute $|\mathbf{H}(k)|^2$ in the preprocessing stage, but the proposed algorithm only needs once preprocessing for main and branch iterating detector. At the observation nodes, AMP-G and AMP-EP both require $13N_T N_R$ FLOPs, the AMP-LA algorithm needs $10N_T N_R + 3N_R$ FLOPs, and the AMP-LA algorithm needs $10N_T N_R + 3N_R$ FLOPs. The computational complexity of extrinsic information LLRs is $[(Q+7)|\mathcal{A}|+1]N_T$ FLOPs. We give the normalized computational complexity versus number of antennas N_T for different

TABLE 1: Computational complexity comparison for subcarrier per iteration.

Algorithm	Total calculation complexity (FLOPs)
AMP-G	$(17 \mathcal{A} + 32) N_T N_R + (2Q + 6) \mathcal{A} N_T - 2N_T$
AMP-EP	$43N_T N_R + (2Q + 12) \mathcal{A} N_T + 5N_T$
AMP-LA	$25N_T N_R + (2Q + 12) \mathcal{A} N_T + 4N_T + 3N_R$
BP based on MRF	$O(N_T^2 N_R)$
Turbo-MMSE-SIC	$8N_R^3 + 8N_T N_R^2 + 8N_T^2 N_R + 6N_R^2 + 24N_T N_R + (2Q + 12) \mathcal{A} N_T + 5N_T - N_R$
Proposed algorithm	$47N_T N_R + (4Q + 24) \mathcal{A} N_T + 8N_T + 6N_R$

- (1) **Initialization:** Input $z_{f_m \rightarrow X_n}^{(0)} = 0$, $v_{f_m \rightarrow X_n}^{(0)} = \infty$, $P^{(1)}(c_n^q) = \tilde{P}^{(1)}(c_n^q) = 1/2$.
- (2) **for** $n = 1 \rightarrow N_T$ **do** computation at the cloning nodes.
- (3) Calculate $\tilde{P}^{(i)}(X_n^q)$ by (25).
- (4) Calculate $\hat{X}_{X_n}^{(i)}$ and $\hat{\tau}_{X_n}^{(i)}$ by (30).
- (5) **end for**
- (6) **for** $m = 1 \rightarrow N_R$ **do** computation at the channel transition nodes.
- (7) Calculate $v_{f_m}^{(i)}$ by (32).
- (8) Calculate $z_{f_m}^{(i)}$ by (34).
- (9) **end for**
- (10) **for** $n = 1 \rightarrow N_T$ **do** computation at the cloning nodes.
- (11) Calculate $\xi_{X_n}^{(i)}$ and $\gamma_{X_n}^{(i)}$ by (35) and (36).
- (12) Calculate $\tilde{P}_{eq}^{(i)}(X_n)$ by (38).
- (13) **for** $q = 1 \rightarrow Q$ **do** computation of LLRs
- (14) Calculate $L_e^{(i)}(c_n^q)$ by (37).
- (15) **end for**
- (16) **end for**

ALGORITHM 1: Processing procedure of the AMP-LA algorithm.

detection algorithms in the $N_T \times N_T$ multiuser MIMO-OFDM systems with QPSK modulation as shown in Figure 4. For the antennas number $N_T = N_R$, both BP algorithm based on Markov random field (MRF) in [14] and MMSE-SIC algorithm have the same order complexity $O(N_T^2 N_R) = O(N_R^3)$, so we only consider the normalized computational complexity of MMSE-SIC algorithm. Figure 4 shows that the floating-point operations of proposed algorithm are fewer than other conventional MMSE-SIC algorithm and AMP-G algorithm, especially when the number of antennas is large. With the number of antennas N_T increasing, the normalized computational complexity of AMP-EP is very close to the proposed detection method. The AMP-LA has the minimum complexity. The similar normalized complexity comparison can be got from other types of modulation.

5. Simulation Results and Performance Analysis

We use convolution-coded style to test the performance of different detection methods in MIMO system, respectively. The channels are 16-tap Rayleigh fading with equal tap exponential power delay profile model and 9-tap ITU-EVA model which has unequal taps power delay profile, so we can simulate frequency selective fading channel satisfactorily.

TABLE 2: Simulation parameters of large-scale MIMO-OFDM system.

Parameters	Value
Channel coding	Recursive systematic convolutional (RSC) code
Decoding algorithm	BCJR algorithm
Doppler maximal frequency	5 Hz
Data modulation format	QPSK/16QAM
Transmitter FFT size	128
CP size	16
Channel estimation	Ideal estimation
Noise model	Independent AWGN
Information bits	8192

When the number of taps is large and paths have equal energy, the MIMO-ISI channel has severely delay spread so we can simulate the scene just like [13, 14]. When the channel is ITU-EVA we can simulate the real application scenario in 5G.

The simulation parameters of the large-scale MIMO-OFDM system are given in Table 2. Suppose that the synchronization is ideal in the iteration process. The BER- E_b/N_0 performance of our algorithm and other conventional

- (1) **Initialization:**
if $i = 0$, input the initial *a priori* LLRs of the main detector and the branch detector, set $L_{a,M}^{(0)}(c_n^q) = 0$, $L_a^{(0)}(c_n^q) = 0$, $L_{a,b}^{(0)}(c_n^q) = 0$.
- (2) **if** $i \geq 0$, combining with the *a priori* LLR from the decoders and the branch detector, let $L^{(i)}(c_n^q) = L_{a,M}^{(i)}(c_n^q) + L_a^{(i)}(c_n^q)$, **do** computation extrinsic LLRs $L_{e,M}^{(i)}(c_n^q)$ of the main detector by Algorithm 1.
- (3) Pass $L_{e,M}^{(i)}(c_n^q)$ to the decoders and to the branch detector in parallel, respectively.
- (4) The decoders Calculate $L_a^{(i)}(c_n^q)$ by decoding algorithm.
- (5) Calculate $\rho_M^{(i)}$ between $L_{e,M}^{(i)}(c_n^q)$ and $L_{a,M}^{(i)}(c_n^q)$ by (14).
- (6) Generate $L_{a,b}^{(i)}(c_n^q)$ by (12).
- (7) While given the *a priori* LLRs $L_{a,b}^{(i)}(c_n^q)$,
do computation extrinsic LLRs $L_{e,b}^{(i)}(c_n^q)$ of the branch detector by Algorithm 1.
- (8) Calculate $\rho_b^{(i)}$ between $L_{e,b}^{(i)}(c_n^q)$ and $L_{a,b}^{(i)}(c_n^q)$ by (13).
- (9) Set $i = i + 1$, **do** computation the *a priori* LLRs $L_{a,M}^{(i)}(c_n^q)$ by (11) for the next iteration.

ALGORITHM 2: Processing procedure of the proposed Main-Branch structure turbo iteration detection.

FIGURE 4: Comparison of normalized computational complexity versus number of antennas.

FIGURE 5: BER performance for different detection algorithms in a 64×64 MIMO system, with QPSK and 16QAM modulation, 16-tap Rayleigh channel with equal tap power. FIGURE 6: BER performance versus E_b/N_0 for the AMP-EP algorithm with different number of iterations in a 64×64 MIMO system with QPSK and 16QAM modulation. 16-tap Rayleigh channel with equal tap power.

algorithms is simulated by above given channel model shown as Figures 5–9, respectively.

We first consider a RSC-coded QPSK and 16QAM modulation 64×64 MIMO system, and the channel is 16-tap Rayleigh fading with equal tap exponential power delay profile model. It is seen from Figure 5 that the proposed method has better BER performance than conventional detector using soft interference cancellation (SIC) based on the Minimum Mean Square Error (MMSE) criterion when the iterative numbers are $i > 2$ in QPSK and $i > 5$ in 16QAM. Our method also gets better performance than Approximate Message Passing detection by expectation propagation (AMP-EP) scheme and AMP with linear approximation (AMP-LA) scheme in [13]. For instance, at the BER 10^{-3} , the proposed

FIGURE 7: BER performance versus E_b/N_0 for the proposed algorithm with different number of iterations in a 64×64 MIMO system, QPSK and 16QAM modulation, 16 taps Rayleigh channel with equal tap power.

FIGURE 8: BER performance for different detection algorithms in a 16×16 MIMO system, with QPSK and 16QAM modulation, 16-tap Rayleigh channel with equal tap power.

method gets about 0.3 dB gain compared with the MMSE-SIC when the iterative numbers are $i = 6$ with QPSK. At the BER level of 10^{-3} and the number of iterations being 12, the proposed method gets about 1.1 dB gains compared with MMSE-SIC in 16QAM. For this BER level and 6 iterations, the proposed iterative equalization scheme also achieves better BER performance compared with the MMSE-SIC, while the AMP-EP performance is worse than MMSE-SIC. Note that, in the low E_b/N_0 , the performance curve of proposed algorithm first reaches the matched filter bound (MFB). It should be stressed that the proposed method is very precise especially regarding the first iteration.

FIGURE 9: BER performance for different detection algorithms in a 4×4 MIMO system, with QPSK and 16QAM modulation, 16-tap Rayleigh channel with equal tap power.

From Figures 6 and 7 we can observe that the proposed method also has better BER performance than AMP-EP algorithm with different number of iterations in multiuser 64×64 MIMO systems. When the number of iterations is increased, the performance improvement of our algorithm is obviously more remarkable. Figure 6 presents that 6 and 12 iterations are enough for the AMP-EP to approach the MFB of QPSK and 16QAM modulation at $\text{BER} = 10^{-5}$, while the BER performance of the proposed algorithm can approach the MFB after just 5 and 10 iterations at $\text{BER} = 10^{-4}$ as shown in Figure 7. The performance gap between the AMP-EP algorithm and the MFB is obviously larger than that between the proposed method and the MFB for the same BER level in the low E_b/N_0 region. Because the AMP-LA is simplified by AMP-EP, the performance of AMP-LA is worse than AMP-EP. So we just consider comparison of the performance of AMP-EP.

It should be also noted that the performance of our scheme for large-scale MIMO systems can be seriously affected due to error propagation effects on the approximation made in the message updating by (34) and (35). If we want the performance curves of our scheme to approach the MFB, we need more number of iterations for the low-dimensional MIMO systems. As is shown in Figures 8 and 9, different methods of BER performance are present in 16×16 and 4×4 MIMO systems with QPSK and 16QAM modulation. Compared with the MFB, at $\text{BER} = 10^{-2}$, the performance loss for the proposed algorithm is about 1.3 dB and 1.5 dB when the number of iterations i is 6, respectively, in a 16×16 and 4×4 MIMO system with the 16QAM constellation, while the performance degraded for the proposed algorithm in a 64×64 MIMO system is about 0.7 dB by the same iterations compared with the MFB at the same BER level with 16QAM. This is very important for large-dimensional systems, because the error propagation effects will be lower. On the other hand, the performance

FIGURE 10: BER performance for different detection algorithms in a 64×64 MIMO system, with QPSK modulation and ITU-EVA channel.

gap between the proposed method and AMP-EP algorithm in the more antennas MIMO system is obviously larger than that in a small-scale 4×4 MIMO system for the same BER level. The proposed algorithm could also be applied in noncoded systems for the Main-Branch iterating soft detection structure not using interleaving. In fact, in noncoded systems our algorithm can be seen as complexity turbo detection implemented in the frequency domain which does not employ channel decoders in the feedback loop. So it has similar performance behavior with turbo detection just like the BER performance curves of our scheme approaching the MFB in larger E_b/N_0 region. It is seen from Figure 9 that the performance of the proposed algorithm is even lower than MMSE-SIC when the number of iterations is $i = 3$ for QPSK and $i = 6$ for 16QAM, and we also find that the performance of the proposed algorithm is also near other algorithms when the number of iterations is small in Figure 8. So we can conclude that our proposed algorithm is best used in the large-scale MIMO systems such as $N \geq 16$.

In Figures 10 and 11, the performance of the proposed algorithm is compared with other different methods in ITU-EVA channel for QPSK and 16QAM modulation, respectively. We can see that the performance gap between the proposed method and other detection algorithms over ITU-EVA channel is larger than that over 16-tap Rayleigh fading with equal tap power delay profile model.

6. Conclusions

In this paper, we proposed an iterative detection algorithm in order to defy intersymbol interference and improve the spectral efficiency for uplink large-scale MIMO communication systems. The iterative detection algorithm can achieve near-optimal performance by iteratively exchanging probabilistic information about the coded bits between a soft-input soft-output (SISO) Main-Branch structure detector and a SISO

FIGURE 11: BER performance comparison for different detection algorithms in a 64×64 MIMO system, with 16QAM modulation and ITU-EVA channel.

channel decoder. Our method calculates the correlation coefficients between iterative main detector and branch detector by computing gradient to minimize the MSE. By applying AMP-LA, the precision of detection is improved while the complexity can be acceptable. The simulation results show that the performance of the proposed method is better than traditional MMSE-SIC detection and AMP type detection in small number of iterations. Moreover, the complexity of the proposed algorithm is lower than MMSE-SIC algorithm and near the AMP-EP when the number of system antennas is large.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This study is funded by the National Natural Science Foundation of China (Grant nos. 61571162 and 61101122) and Science and Technology Project of Ministry of Public Security Foundation (Grant nos. 2015GABJC38 and 2015GABJC37), Major National Science and Technology Project (Grant nos. 2014ZX03004003 and 2011ZX03004-004), Municipal Exceptional Academic Leaders Foundation (Grant no. 2014RFXXJ002), and China Postdoctoral Science Foundation (Grant no. 2014M561347).

References

- [1] S. K. Mohammed, A. Zaki, A. Chockalingam, and B. S. Rajan, "High-rate space-time coded large-MIMO systems: low-complexity detection and channel estimation," *IEEE Journal on Selected Topics in Signal Processing*, vol. 3, no. 6, pp. 958–974, 2009.

- [2] C. Knievel, P. A. Hoeher, A. Tyrrell, and G. Auer, "Multi-dimensional graph-based soft iterative receiver for MIMO-OFDM," *IEEE Transactions on Communications*, vol. 60, no. 6, pp. 1599–1609, 2012.
- [3] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: how many antennas do we need?" *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, 2013.
- [4] S. Yang and L. Hanzo, "Fifty years of MIMO detection: the road to large-scale MIMOs," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 1941–1988, 2015.
- [5] R. Koetter, A. C. Singer, and M. Tüchler, "Turbo equalization," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 67–80, 2004.
- [6] E. G. Larsson and J. Jaldén, "Fixed-complexity soft MIMO detection via partial marginalization," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3397–3407, 2008.
- [7] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Transactions on Communications*, vol. 47, no. 7, pp. 1046–1061, 1999.
- [8] C. Studer, S. Fateh, and D. Seethaler, "ASIC implementation of soft-input soft-output MIMO detection using MMSE parallel interference cancellation," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 7, pp. 1754–1765, 2011.
- [9] R. J. Drost and A. C. Singer, "Factor-graph algorithms for equalization," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 2052–2065, 2007.
- [10] G. Colavolpe and G. Geremi, "On the application of factor graphs and the sum-product algorithm to ISI channels," *IEEE Transactions on Communications*, vol. 53, no. 5, pp. 818–825, 2005.
- [11] M. N. Kaynak, T. M. Duman, and E. M. Kurtas, "Belief propagation over SISO/MIMO frequency selective fading channels," *IEEE Transactions on Wireless Communications*, vol. 6, no. 6, pp. 2001–2005, 2007.
- [12] J. Goldberger and A. Leshem, "MIMO detection for high-order QAM based on a Gaussian tree approximation," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4973–4982, 2011.
- [13] S. Wu, L. Kuang, Z. Ni, J. Lu, D. D. Huang, and Q. Guo, "Low-complexity iterative detection for large-scale multiuser MIMO-OFDM systems using approximate message passing," *IEEE Journal on Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 902–915, 2014.
- [14] P. Som, T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "Low-complexity detection in large-dimension MIMO-ISI channels using graphical models," *IEEE Journal on Selected Topics in Signal Processing*, vol. 5, no. 8, pp. 1497–1511, 2011.
- [15] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727–1737, 2001.
- [16] S. Jeong and J. Moon, "Self-iterating soft equalizer," *IEEE Transactions on Communications*, vol. 61, no. 9, pp. 3697–3709, 2013.
- [17] S. Jeong and J. Moon, "Soft-in soft-out DFE and Bi-directional DFE," *IEEE Transactions on Communications*, vol. 59, no. 10, pp. 2729–2741, 2011.
- [18] S. Jeong and J. Moon, "Self-iterating soft equalizer," in *Proceedings of the IEEE Global Telecommunications Conference*, 2011.
- [19] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, NY, USA, 2006.
- [20] S. Wu, L. Kuang, Z. Ni, J. Lu, D. D. Huang, and Q. Guo, "Expectation propagation approach to joint channel estimation and decoding for OFDM systems," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '14)*, pp. 1941–1945, IEEE, Florence, Italy, May 2014.
- [21] T. P. Minka, "Divergence measures and message passing," Tech. Rep., Microsoft Research, Cambridge, UK, 2005.
- [22] T. P. Minka, *A family of algorithms for approximate Bayesian inference [Ph.D. dissertation]*, Massachusetts Institute of Technology, Cambridge, Mass, USA, 2001.
- [23] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*, MIT Press, Cambridge, Mass, USA, 2009.