

# Low-Complexity Iterative Detection for Large-Scale Multiuser MIMO-OFDM Systems Using Approximate Message Passing

Sheng Wu, *Student Member, IEEE*, Linling Kuang, *Member, IEEE*, Zuyao Ni, Jianhua Lu, *Senior Member, IEEE*, Defeng David Huang, *Senior Member, IEEE*, and Qinghua Guo, *Member, IEEE*

**Abstract**—One of the challenges in the design of large-scale multiuser MIMO-OFDM systems is developing low-complexity detection algorithms. To achieve this goal, we leverage message passing algorithms over the factor graph that represents the multiuser MIMO-OFDM systems and approximate the original discrete messages with continuous Gaussian messages through the use of the minimum Kullback-Leibler (KL) divergence criterion. Several signal processing techniques are then proposed to achieve near-optimal performance at low complexity. First, the principle of expectation propagation is employed to compute the approximate Gaussian messages, where the symbol belief is approximated by a Gaussian distribution and then the approximate message is calculated from the Gaussian approximate belief. In addition, the approximate symbol belief can be computed by the *a posteriori* probabilities fed back from channel decoders, which reduces the complexity dramatically. Second, the first-order approximation of the message is utilized to further simplify the message updating, leading to an algorithm that is equivalent to the AMP algorithm

Manuscript received September 10, 2013; revised January 15, 2014; accepted March 12, 2014. Date of publication March 25, 2014; date of current version September 11, 2014. This work was supported in part by the National Nature Science Foundation of China under Grants 91338101, 61021001, and 61132002, in part by the National Basic Research Program of China under Grant 2013CB329001, in part by the Co-Innovation Laboratory of Aerospace Broadband Network Technology, and in part by the Australian Research Council’s Discovery Project DP1093000 and DECRA Grant DE120101266. The guest editor coordinating the review of this manuscript and approving it for publication was Dr. Alexei Ashikhmin.

S. Wu is with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China, and also with the Key Laboratory of Wireless Broadband Communication and Signal Processing, Research Institute of Tsinghua University in Shenzhen, Shenzhen 518057, China (e-mail: peach.shengsheng@gmail.com).

L. Kuang is with the Tsinghua Space Center, Tsinghua University, Beijing 100084, China, and also with the Key Laboratory of Wireless Broadband Communication and Signal Processing, Research Institute of Tsinghua University in Shenzhen, Shenzhen 518057, China (e-mail: kll@tsinghua.edu.cn).

Z. Ni is with the Research Institute of Information Technology, Tsinghua University, Beijing 100084, China, and also with the Key Laboratory of Wireless Broadband Communication and Signal Processing, Research Institute of Tsinghua University in Shenzhen, Shenzhen 518057, China (e-mail: nzy@tsinghua.edu).

J. Lu is with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: lhh-dee@mail.tsinghua.edu.cn).

D. D. Huang is with the School of Electrical, Electronic and Computer Engineering, The University of Western Australia, Perth, WA 6009, Australia (e-mail: huangdf@ee.uwa.edu.au).

Q. Guo is with the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW 2522, Australia, and is also with the School of Electrical, Electronic and Computer Engineering, The University of Western Australia, Perth, WA 6009, Australia (e-mail: qinghua\_guo@uow.edu.au).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTSP.2014.2313766

proposed by Donoho *et al.* Finally, the message updating is further simplified using the central-limit theorem. Compared with the single tree search sphere decoder (STS-SD) and the iterative (turbo) minimum mean-square error based soft interference cancellation (MMSE-SIC) in the literature through extensive simulations, the proposed message passing algorithms can achieve a near-optimal performance while the complexity is decreased by tens of times for a  $64 \times 64$  MIMO system. In addition, it is shown that the proposed message passing algorithms exhibit desirable tradeoffs between performance and complexity for a low-dimensional MIMO system.

**Index Terms**—Approximate message passing, expectation propagation, factor graph, large-scale MIMO-OFDM.

## I. INTRODUCTION

MULTIPLE-INPUT multiple-output orthogonal frequency-division multiplexing (MIMO-OFDM) is a key technology for many wireless communication systems, due to its high spectral efficiency [1]–[3]. Recently, large-scale multiuser MIMO systems with tens to hundreds of antennas at the base-station (BS) have gained significant attention [4]–[12]. The motivation to consider such large-scale MIMO systems is their potential to meet the growing demands for higher throughput and improved quality-of-service of next-generation multiuser wireless communication systems [6]. As the signal of each user is needed to be extracted from the received interfered signal at the BS receiver [9], one key issue in the design of a practical receiver for large-scale MIMO-OFDM systems is how to reduce the complexity of detection without much compromise in performance [4], [7]–[9], [13].

Turbo detection can achieve near-optimal performance by iteratively exchanging probabilistic information about the coded bits between a soft-input soft-output (SISO) detector and a SISO channel decoder [14]. An optimal SISO detector employing the maximum *a posteriori* probability (MAP) algorithm has an exponential complexity  $\mathcal{O}(|\mathcal{A}|^N)$ , where  $|\mathcal{A}|$  is the modulation constellation size and  $N$  is the number of transmit antennas. To reduce the complexity, there has been considerable interest in the development of suboptimal SISO detectors. Various tree search based algorithms perform approximate MAP detection via different pruning strategies to generate soft-output [14]–[17], but their computational complexity  $\mathcal{O}(|\mathcal{A}|^{\alpha N})$ ,  $\alpha \in (0, 1)$  still grows exponentially [18]. The partial marginalization method in [19] marginalizes out part of the symbols by using the “max-log” approximation. It

has a complexity of  $\mathcal{O}(|\mathcal{A}|^r + |N - r|^3)$ ,  $1 \leq r \leq N$ , which is between that of the MAP detector and the linear detector. Partial Gaussian approximation based method treats part of the symbols as Gaussian variables to reduce the computational complexity, which has a complexity similar to the partial marginalization method [20]. To further reduce the complexity, soft interference cancellation (SIC) detections based on the linear minimum mean square error (LMMSE) filtering can be used [21]–[25], which has a complexity of  $\mathcal{O}(M^3)$ , where  $M$  is the number of receive antennas. Nevertheless, the complexity of these suboptimal detectors remains prohibitive for the large-scale systems.

The belief propagation (BP) algorithm and its variants have been applied in MIMO detection [8], [26]–[33]. When the exact BP algorithm is performed over the underlying factor graph of a MIMO system, the complexity is as high as that of the MAP detector due to the marginalization operations over discrete symbols. There have been many approximate algorithms to reduce the complexity. One class of the algorithms pass exact messages over a pruned factor graph that approximates the original system structure. It was suggested in [29] to prune some edges of the factor graph based on the strength of the channel coefficients, which finally reduces the computational complexity to  $\mathcal{O}(|\mathcal{A}|^d)$ , where  $d$  is the maximum edge degree after pruning. With a similar idea, an approximate BP algorithm based on Gaussian tree approximation has been proposed recently for the MIMO detection [32], which approximates the dense factor graph of the MIMO system into a tree and passes exact messages over the resultant tree. Another class of the algorithms pass approximate messages with tractable operations over the original factor graph. Damped Gaussian BP algorithm was considered in [8], [30] and [31], where the BPSK symbols and the messages are all assumed to be Gaussian leading to a simple message-updating rule with the complexity of  $\mathcal{O}(MN)$ . The Gaussian message passing approach to detection (in frequency selective channels) was also investigated in [34], where the symbols were assumed to be Gaussian. Recently, the approximate message passing (AMP) technique with low complexity is proposed for compressive sensing [35]–[37], and then is extended to generalized AMP (GAMP) for iterative estimation of the random vector whose elements are mixed by a matrix [38]. The GAMP has been applied to the iterative frequency domain equalization in [39], and joint sparse channel estimation and decoding for OFDM systems in [40]. Besides those algorithms derived from factor graph, iterative BP algorithm based on Markov random field (MRF) was also investigated in [8] and [33], where the algorithm complexity is  $\mathcal{O}(MN^2)$ .

In this paper, we develop several message passing algorithms to estimate coded symbols in multiuser MIMO-OFDM systems, which are discrete random variables with *a priori* distributions fed back from the decoders. We approximate the discrete messages as continuous Gaussian messages using the minimum Kullback-Leibler (KL) divergence criterion, and the computational complexity of the resulting algorithm is dramatically reduced to  $\mathcal{O}(MN)$  due to the marginalization operations achieved by simple linear processing. To further reduce the computational complexity, several signal processing techniques are proposed. First, the principle of expectation propagation

is used to find the approximate Gaussian messages, where the symbol belief, rather than the message itself, is approximated by a Gaussian distribution. The approximate message is then computed from the approximate symbol belief, resulting in the reduction of the number of messages that need to be tracked. In addition, we use the *a posteriori* probabilities fed back from the decoders to compute the approximate symbol belief at a lower complexity. Second, the first-order approximation of the messages is utilized to further simplify the message updating, which can reduce the number of messages from  $\mathcal{O}(MN)$  to  $\mathcal{O}(M + N)$ . By examining the derivation of the AMP algorithm in [35] and [37] and the GAMP algorithm in [38] and [39], it can be found that the algorithm based on the first-order approximation is equivalent to the sum-product GAMP, although our derivation is based on the framework of expectation propagation. Finally, the message updating is further simplified by the central-limit theorem. These proposed message passing algorithms offer tradeoffs between performance and complexity with desirable characteristics. Compared with the STS-SD [17] and the turbo MMSE-SIC [21], [41] in the literature through extensive simulations, the proposed message passing algorithms can achieve a near-optimal performance while the complexity is decreased by tens of times in a  $64 \times 64$  MIMO system. In addition, it is shown that these message passing algorithms exhibit desirable tradeoffs between performance and complexity in a low-dimensional MIMO system. We note that the scope of this work is different from that of [42]–[46]. The works in [42]–[46] deal with joint channel estimation and decoding in MIMO-OFDM system, while this work focuses on *low-complexity* turbo detection for the large-scale multiuser MIMO systems with known channel state information. Following the approach of minimizing the local information divergence [47], [48], our proposed algorithms can be extended to the joint channel estimation and decoding in multiuser MIMO-OFDM systems.

This paper is organized as follows: In Section II, the signal model is presented. In Section III, algorithms based on approximate message passing are proposed, and their complexities are analyzed. In Section IV, simulation results are presented. Finally, Section V concludes this paper.

Throughout the paper, we use the following notations. Lowercase letters (e.g.,  $x$ ) denote scalars, bold lowercase letters (e.g.,  $\mathbf{x}$ ) denote column vectors, and bold uppercase letters (e.g.,  $\mathbf{X}$ ) denote matrices. The superscripts  $(\cdot)^*$ ,  $(\cdot)^\top$  and  $(\cdot)^\text{H}$  denote the complex conjugate, transpose and conjugate transpose, respectively. Also, the symbol  $\mathbf{I}$  denotes an identity matrix;  $\ln(\cdot)$  denotes the natural logarithm; and  $\mathcal{N}_{\mathbb{C}}(x, \hat{x}, \hat{\tau}) \triangleq (\pi\hat{\tau})^{-1} \exp(-|x - \hat{x}|^2/\hat{\tau})$  denotes a complex Gaussian function. Furthermore,  $\mathbf{x} \setminus x_i$  denotes the symbols in  $\mathbf{x}$  with  $x_i$  excluded; and  $\propto$  denotes equality up to a scale. Finally,  $\mathbb{E}_{p(x)}[\cdot]$  denotes the statistical expectation operation with respect to the distribution  $p(x)$ .

## II. SYSTEM MODEL AND ITERATIVE RECEIVER

Consider the uplink of a multiuser system with  $N$  independent users, where each user is equipped with one transmit antenna, and the receiver is equipped with an array of  $M$  antennas. As shown in Fig. 1, the transmission scheme for each user is

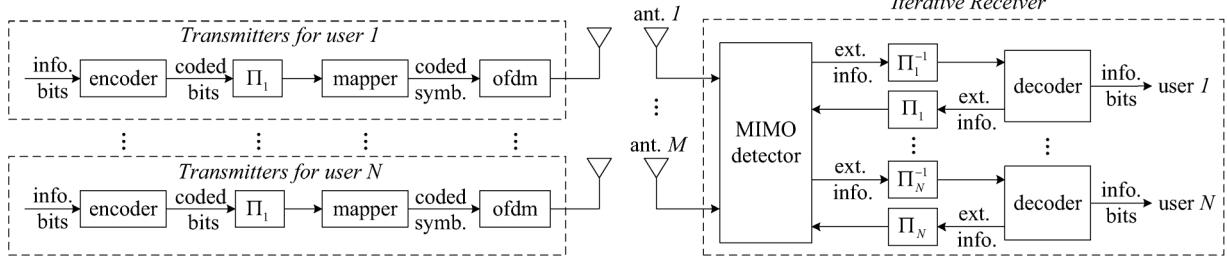


Fig. 1. Block diagram of a coded multiuser MIMO system, where  $\Pi_n$  and  $\Pi_n^{-1}$  denote the interleaver and the corresponding de-interleaver for the  $i$ th user, respectively.

based on OFDM to combat the inter-symbol interference, and an iterative signal processing structure is employed at the receiver. For a  $2^Q$ -ary modulation with the constellation set  $\mathcal{A}$ , every  $Q$  interleaved coded bits are mapped to one modulation symbol. Then the frequency domain symbols transmitted by the  $i$ th user are denoted by  $\mathbf{x}_i = [x_i(1), x_i(2), \dots, x_i(K)]^\top$ , where  $x_i(k) \in \mathcal{A}$  is the symbol transmitted at the  $k$ th subcarrier and  $K$  is the number of subcarriers. A  $K$ -point IFFT is applied to the symbol sequence  $\mathbf{x}_i$  and a cyclic prefix (CP) is added before transmission. Finally,  $N$  blocks of data from the  $N$  users are transmitted simultaneously over the MIMO channel. At the receiver, the CP is removed and the received signal from each receive antenna is converted into the frequency domain through a  $K$ -point FFT. It is assumed that the  $N$  transmitters and the receiver are synchronized and all channel coefficients are known at the receiver. The received signal at the  $k$ th subcarrier  $\mathbf{y}(k) \in \mathbb{C}^{M \times 1}$  can be written as

$$\mathbf{y}(k) = \mathbf{H}(k)\mathbf{x}(k) + \mathbf{n}(k), \quad (1)$$

where  $\mathbf{x}(k) \in \mathbb{C}^{N \times 1}$  denotes the symbols at all the  $k$ th subcarriers,  $\mathbf{n}(k)$  denotes a circularly symmetric complex noise vector with zero-mean and covariance matrix  $\sigma_n^2 \mathbf{I}$ , and  $\mathbf{H}(k)$  denotes the frequency domain MIMO channel, which is given by

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

where  $L$  is the number of channel taps, and  $\mathbf{H}_l \in \mathbb{C}^{M \times N}$  is the channel matrix of the  $l$ th channel taps in the time-domain, which are assumed to be independent for different  $l$ . For notational simplicity, we will henceforth omit the subcarrier index  $k$ . We note here that, although in this paper we focus on OFDM systems, the proposed algorithms can also be applied in single carrier systems.

Turbo MIMO detection exchanges extrinsic log likelihood ratios (LLRs) of the coded bits between a detector and a bank of channel decoders. Based on the received signal  $\mathbf{y}$  and the *a priori* LLRs  $\{L(c_i^q)\}$  fed back from the decoders, the task of the detector is to generate extrinsic LLR for each coded bit  $c_i^q$  [21]

$$L_e(c_i^q) = \ln \frac{p(c_i^q = 1 | \mathbf{y})}{p(c_i^q = 0 | \mathbf{y})} - L(c_i^q). \quad (3)$$

The optimal MAP detection has an intractable complexity for large  $N$ , which needs to enumerate the *a posteriori* distribution

$p(\mathbf{x} | \mathbf{y})$  for all  $x_i$  to calculate  $p(x_i | \mathbf{y}) = \sum_{\mathbf{x} \setminus x_i} p(\mathbf{x} | \mathbf{y})$ , and then compute the extrinsic LLR of  $c_i^q$

$$\begin{aligned} L_e(c_i^q) &= \ln \frac{\sum_{x_i \in \mathcal{A}_q^1} p(x_i | \mathbf{y})}{\sum_{x_i \in \mathcal{A}_q^0} p(x_i | \mathbf{y})} - L(c_i^q) \\ &= \ln \frac{\sum_{x_i \in \mathcal{A}_q^1} p(x_i) p(\mathbf{y} | x_i)}{\sum_{x_i \in \mathcal{A}_q^0} p(x_i) p(\mathbf{y} | x_i)} - L(c_i^q), \end{aligned} \quad (4)$$

where  $\mathcal{A}_q^1$  and  $\mathcal{A}_q^0$  denote the subset of all the symbols in which the  $q$ th bit has the value of 1 and 0, respectively. To reduce the complexity, linear filter based methods were proposed to extract the extrinsic information [21], [49]. These methods compute the symbol estimate  $\hat{\mathbf{x}}$  by performing matrix operations on the received signal  $\mathbf{y}$ , and model the estimate error vector  $\mathbf{e} = \hat{\mathbf{x}} - \mathbf{x}$  as a Gaussian random vector with zero mean and covariance matrix  $\mathbf{V}_e$ . The elements in  $\mathbf{e}$  are assumed to be independent of each other, and then we can get  $p(\mathbf{y} | x_i) = p(\hat{x}_i | x_i) = \mathcal{N}_{\mathbb{C}}(x_i; \hat{x}_i, \hat{\tau}_i)$ , where  $\hat{x}_i$  is the estimate of  $x_i$ , and  $\hat{\tau}_i$  is the  $i$ th diagonal element of  $\mathbf{V}_e$ , which are given in [50] as follows

$$\hat{x}_i = (\mathbf{h}_i^\top \mathbf{V}_y^{-1} \mathbf{h}_i)^{-1} \mathbf{h}_i^\top \mathbf{V}_y^{-1} (\mathbf{y} - \mathbf{Hm}_x) + m_{x_i}, \quad (5)$$

$$\hat{\tau}_i = (\mathbf{h}_i^\top \mathbf{V}_y^{-1} \mathbf{h}_i)^{-1} - v_{x_i}, \quad (6)$$

where  $\mathbf{h}_i$  is the  $i$ th column of  $\mathbf{H}$ ,  $\mathbf{V}_y = \mathbf{H}\mathbf{V}_x\mathbf{H}^\top + \sigma_n^2 \mathbf{I}$  is the *a priori* covariance matrix of  $\mathbf{y}$ ,  $\mathbf{m}_x$  and  $\mathbf{V}_x$  are the *a priori* mean and the *a priori* covariance matrix of  $\mathbf{x}$ , respectively,  $m_{x_i}$  is the  $i$ th element of  $\mathbf{m}_x$ , and  $v_{x_i}$  is the  $i$ th diagonal element of  $\mathbf{V}_x$ .

### III. TURBO DETECTION: GRAPHICAL MODEL AND APPROXIMATE MESSAGE PASSING

#### A. Factor Graph and Message Passing Algorithm

For the presentation of the factor graph and message passing algorithm, we use the same convention as in [51] and [52], to which we refer the reader for an in-depth review. The joint distribution  $p(\mathbf{x}, \mathbf{y})$  can be factorized into

$$\begin{aligned} p(\mathbf{x}, \mathbf{y}) &= p(\mathbf{x})p(\mathbf{y} | \mathbf{x}) \\ &= \prod_i p(x_i) \prod_j f_j(y_j | \mathbf{x}), \end{aligned} \quad (7)$$

where

$$f_j(y_j | \mathbf{x}) = \frac{1}{\pi \sigma_n^2} \exp \left\{ -\frac{|y_j - \sum_i h_{j,i} x_i|^2}{\sigma_n^2} \right\}, \quad (8)$$

and  $h_{j,i}$  denotes the component of  $\mathbf{H}$  in the  $j$ th row and  $i$ th column. This factorization is represented by the factor graph in Fig. 2. In the factor graph, node  $\phi_i$  represents the mapping

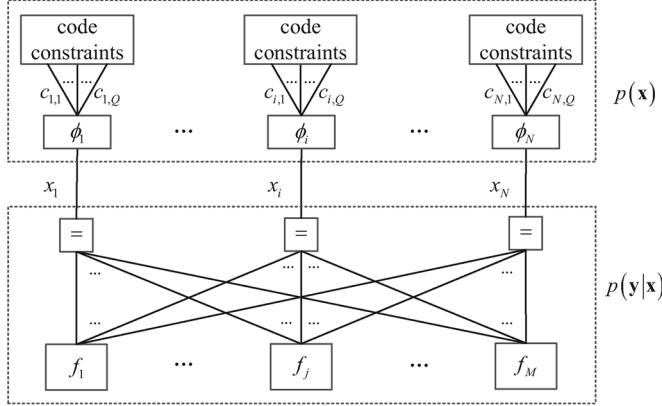


Fig. 2. Factor graph representation of the multiuser MIMO-OFDM system for one subcarrier. Note that the code constraints are across all the subcarriers.

constraint  $\delta(\varphi(\mathbf{c}_i) - x_i)$ , where  $\varphi(\mathbf{c}_i)$  is the mapping function and  $\delta(\cdot)$  is the Kronecker delta function, and the node with “=” represents the cloning of variable as in [52]. Let  $\mu_{x_i \rightarrow f_j}^t(x_i)$  denote the message sent from the cloning node of  $x_i$  to the channel transition node  $f_j$  in the  $t$ th turbo iteration, and let  $\mu_{f_j \rightarrow x_i}^t(x_i)$  denote the message in the opposite direction. The message-update rules are given by [51]

$$\mu_{x_i \rightarrow f_j}^t(x_i) = \mu_{\phi_i \rightarrow x_i}^t(x_i) \prod_{j' \neq j} \mu_{f_{j'} \rightarrow x_i}^{t-1}(x_i), \quad (9)$$

$$\mu_{f_j \rightarrow x_i}^t(x_i) = \sum_{\mathbf{x} \setminus x_i} f_j(y_j | \mathbf{x}) \prod_{i' \neq i} \mu_{x_{i'} \rightarrow f_j}^t(x_{i'}). \quad (10)$$

where  $\mu_{\phi_i \rightarrow x_i}^t(x_i) = \prod_q (\exp[c_i^q L^t(c_i^q)] / 1 + \exp[L^t(c_i^q)])$  denotes the message from the node  $\phi_i$  to the cloning node of  $x_i$ , and  $\{L^t(c_i^q)\}$  denotes the *a priori* LLRs fed back from the decoders in the  $t$ th turbo iteration.

Since the considered factor graph is loopy, a message passing schedule is required [51]. During a single turbo iteration, there may be multiple inner iterations within the dashed block at the bottom of the factor graph, which corresponds to  $p(\mathbf{y}|\mathbf{x})$  or the process of detection. In this work, to reduce complexity, we choose to pass messages from the top to the bottom of Fig. 2 and back again immediately, thereby removing the inner iteration of the detection. Once the decoders update extrinsic information and pass it to the cloning codes, a new iteration starts.

### B. Message Passing Algorithm Using Gaussian Approximation

As the symbols take on values in the discrete set  $\mathcal{A}$ , the computation of  $\mu_{f_j \rightarrow x_i}^t(x_i)$  in (10) requires exponential time to marginalize out the random vector  $\mathbf{x} \setminus x_i$ . If we consider  $x_i$  as a continuous random variable and approximate the message  $\mu_{x_i \rightarrow f_j}^t(x_i)$  into a complex Gaussian function  $\hat{\mu}_{x_i \rightarrow f_j}^t(x_i) = \mathcal{N}_{\mathbb{C}}(x_i; \hat{x}_{x_i \rightarrow f_j}^t, \hat{\tau}_{x_i \rightarrow f_j}^t)$ , where the parameters  $\hat{x}_{x_i \rightarrow f_j}^t$  and  $\hat{\tau}_{x_i \rightarrow f_j}^t$  will be defined later (see (19) and (20)),  $\mu_{f_j \rightarrow x_i}^t(x_i)$  can be calculated by integration, i.e.,

$$\begin{aligned} \mu_{f_j \rightarrow x_i}^t(x_i) &= \int_{\mathbf{x} \setminus x_i} f_j(y_j | \mathbf{x}) \prod_{i' \neq i} \mathcal{N}_{\mathbb{C}}(x_{i'}; \hat{x}_{x_{i'} \rightarrow f_j}^t, \hat{\tau}_{x_{i'} \rightarrow f_j}^t) \\ &= \mathcal{N}_{\mathbb{C}}(h_{j,i} x_i; z_{f_j \rightarrow x_i}^t, \nu_{f_j \rightarrow x_i}^t), \end{aligned} \quad (11)$$

where the parameters  $z_{f_j \rightarrow x_i}^t$  and  $\nu_{f_j \rightarrow x_i}^t$  are given by

$$z_{f_j \rightarrow x_i}^t = y_j - \sum_{i' \neq i} h_{j,i'} \hat{x}_{x_{i'} \rightarrow f_j}^t, \quad (12)$$

and

$$\nu_{f_j \rightarrow x_i}^t = \sigma_n^2 + \sum_{i' \neq i} |h_{j,i'}|^2 \hat{\tau}_{x_{i'} \rightarrow f_j}^t. \quad (13)$$

Then, by substituting  $\mu_{\phi_i \rightarrow x_i}^t(x_i) = \mathcal{N}_{\mathbb{C}}(h_{j,i} x_i; z_{f_j \rightarrow x_i}^t, \nu_{f_j \rightarrow x_i}^t)$  into (9),  $\mu_{x_i \rightarrow f_j}^t(x_i)$  can be normalized as follows

$$\mu_{x_i \rightarrow f_j}^t(x_i) = \frac{\mu_{\phi_i \rightarrow x_i}^t(x_i) \mathcal{N}_{\mathbb{C}}(x_i; \zeta_{x_i \rightarrow f_j}^{t-1}, \gamma_{x_i \rightarrow f_j}^{t-1})}{\sum_{x_i \in \mathcal{A}} \mu_{\phi_i \rightarrow x_i}^t(x_i) \mathcal{N}_{\mathbb{C}}(x_i; \zeta_{x_i \rightarrow f_j}^{t-1}, \gamma_{x_i \rightarrow f_j}^{t-1})}, \quad (14)$$

where  $\gamma_{x_i \rightarrow f_j}^{t-1}$  and  $\zeta_{x_i \rightarrow f_j}^{t-1}$  are given by

$$\gamma_{x_i \rightarrow f_j}^{t-1} = \left( \sum_{j' \neq j} \frac{|h_{j',i}|^2}{\nu_{f_{j'} \rightarrow x_i}^{t-1}} \right)^{-1}, \quad (15)$$

and

$$\zeta_{x_i \rightarrow f_j}^{t-1} = \gamma_{x_i \rightarrow f_j}^{t-1} \sum_{j' \neq j} \frac{h_{j',i}^* z_{f_{j'} \rightarrow x_i}^{t-1}}{\nu_{f_{j'} \rightarrow x_i}^{t-1}}. \quad (16)$$

We can find that  $\zeta_{x_i \rightarrow f_j}^t, j = 1, 2, \dots, M$  are distinct from each other in only two summation terms, and so do  $z_{f_j \rightarrow x_i}^t, i = 1, 2, \dots, N$ . Instead of directly computing these messages, we can retrieve them from a sum by means of an addition and a subtraction, respectively, as follows

$$\begin{aligned} z_{f_j \rightarrow x_i}^t &= y_j - \sum_{i' \neq i} h_{j,i'} \hat{x}_{x_{i'} \rightarrow f_j}^t \\ &= z_{f_j}^t + h_{j,i} \hat{x}_{x_i \rightarrow f_j}^t, \end{aligned} \quad (17)$$

$$\begin{aligned} \zeta_{x_i \rightarrow f_j}^t &= \gamma_{x_i \rightarrow f_j}^t \sum_{j' \neq j} \frac{h_{j',i}^* z_{f_{j'} \rightarrow x_i}^t}{\nu_{f_{j'} \rightarrow x_i}^t} \\ &= \gamma_{x_i \rightarrow f_j}^t \left( \xi_{x_i}^t - \frac{h_{j,i}^* z_{f_j \rightarrow x_i}^t}{\nu_{f_j \rightarrow x_i}^t} \right), \end{aligned} \quad (18)$$

where  $z_{f_j}^t = y_j - \sum_i h_{j,i} \hat{x}_{x_i \rightarrow f_j}^t$ , and  $\xi_{x_i}^t = \sum_j (h_{j,i}^* z_{f_j \rightarrow x_i}^t / \nu_{f_j \rightarrow x_i}^t)$ .

In order to find the Gaussian function  $\hat{\mu}_{x_i \rightarrow f_j}^t(x_i)$  to replace  $\mu_{x_i \rightarrow f_j}^t(x_i)$ , a natural approach is to minimize the inclusive KL divergence  $\text{KL}(\mu_{x_i \rightarrow f_j}^t(x_i) \| \hat{\mu}_{x_i \rightarrow f_j}^t(x_i))$  [53], [54], then we can get the following

$$\hat{x}_{x_i \rightarrow f_j}^t = \sum_{\boldsymbol{\alpha}_s \in \mathcal{A}} \boldsymbol{\alpha}_s \mu_{x_i \rightarrow f_j}^t(x_i = \boldsymbol{\alpha}_s), \quad (19)$$

$$\hat{\tau}_{x_i \rightarrow f_j}^t = \sum_{\boldsymbol{\alpha}_s \in \mathcal{A}} |\boldsymbol{\alpha}_s|^2 \mu_{x_i \rightarrow f_j}^t(x_i = \boldsymbol{\alpha}_s) - \left| \hat{x}_{x_i \rightarrow f_j}^t \right|^2. \quad (20)$$

According to the factor graph shown in Fig. 2, the message  $\mu_{x_i \rightarrow \phi_i}^t(x_i)$  from the cloning node of  $x_i$  to the mapping

TABLE I  
THE APPROXIMATE MESSAGE PASSING ALGORITHM USING  
GAUSSIAN APPROXIMATION FOR THE  $t$ th ITERATION  
BETWEEN THE DECODERS AND THE DETECTOR

---

```

1: Initialization:
2: Set  $\zeta_{x_i \rightarrow f_j}^0 = 0, \gamma_{x_i \rightarrow f_j}^0 = \infty, p^1(c_i^q) = \frac{1}{2}$ .
3: for  $i = 1 \rightarrow N$  do ▷ Computation in the cloning nodes
4:   for  $j = 1 \rightarrow M$  do
5:      $\mu_{x_i \rightarrow f_j}^t(x_i) = \frac{\mu_{\phi_i \rightarrow x_i}^t(x_i) \mathcal{N}_{\mathbb{C}}(x_i; \zeta_{x_i \rightarrow f_j}^{t-1}, \gamma_{x_i \rightarrow f_j}^{t-1})}{\sum_{x_i \in \mathcal{A}} \mu_{\phi_i \rightarrow x_i}^t(x_i) \mathcal{N}_{\mathbb{C}}(x_i; \zeta_{x_i \rightarrow f_j}^{t-1}, \gamma_{x_i \rightarrow f_j}^{t-1})}$ ,
6:      $\hat{x}_{x_i \rightarrow f_j}^t = \mathbb{E}_{\mu_{x_i \rightarrow f_j}^t(x_i)}[x_i]$ ,
7:      $\hat{\tau}_{x_i \rightarrow f_j}^t = \mathbb{E}_{\mu_{x_i \rightarrow f_j}^t(x_i)}[|x_i|^2] - |\hat{x}_{x_i \rightarrow f_j}^t|^2$ .
8:   end for
9: end for
10: for  $j = 1 \rightarrow M$  do ▷ Computation at the channel transition nodes
11:    $\nu_{f_j}^t = \sigma_n^2 + \sum_i |h_{j,i}|^2 \hat{\tau}_{x_i \rightarrow f_j}^t, z_{f_j}^t = y_j - \sum_i h_{j,i} \hat{x}_{x_i \rightarrow f_j}^t$ ,
12:   for  $i = 1 \rightarrow N$  do
13:      $\nu_{f_j \rightarrow x_i}^t = \nu_{f_j}^t - |h_{j,i}|^2 \hat{\tau}_{x_i \rightarrow f_j}^t, z_{f_j \rightarrow x_i}^t = z_{f_j}^t + h_{j,i} \hat{x}_{x_i \rightarrow f_j}^t$ .
14:   end for
15: end for
16: for  $i = 1 \rightarrow N$  do ▷ Computation at the cloning nodes
17:    $\gamma_{x_i}^t = \left( \sum_j \frac{|h_{j,i}|^2}{\nu_{f_j \rightarrow x_i}^t} \right)^{-1}, \xi_{x_i}^t = \sum_j \frac{h_{j,i}^* z_{f_j \rightarrow x_i}^t}{\nu_{f_j \rightarrow x_i}^t}, \zeta_{x_i}^t = \gamma_{x_i}^t \xi_{x_i}^t$ ,
18:   for  $j = 1 \rightarrow M$  do
19:      $\gamma_{x_i \rightarrow f_j}^t = \left( \frac{1}{\gamma_{x_i}^t} - \frac{|\nu_{f_j \rightarrow x_i}^t|^2}{\nu_{f_j \rightarrow x_i}^t} \right)^{-1}$ ,
20:      $\zeta_{x_i \rightarrow f_j}^t = \gamma_{x_i \rightarrow f_j}^t \left( \xi_{x_i}^t - \frac{h_{j,i}^* z_{f_j \rightarrow x_i}^t}{\nu_{f_j \rightarrow x_i}^t} \right)$ .
21:   end for
22:    $\tilde{p}_{eq}^t(x_i) = \mu_{\phi_i \rightarrow x_i}^t(x_i) \mathcal{N}_{\mathbb{C}}(x_i; \zeta_{x_i}^t, \gamma_{x_i}^t)$ ,
23:   for  $q = 1 \rightarrow Q$  do ▷ Computation of LLRs
24:      $L_e^t(c_i^q) = \ln \frac{\sum_{\mathcal{A}_q^1} \tilde{p}_{eq}^t(x_i)}{\sum_{\mathcal{A}_q^0} \tilde{p}_{eq}^t(x_i)} - L^t(c_i^q)$ .
25:   end for
26: end for

```

---

node  $\phi_i$  is the multiplication of all the incoming messages  $\{\mu_{f_j \rightarrow x_i}^t(x_i), \forall j\}$ , i.e.,

$$\begin{aligned} \mu_{x_i \rightarrow \phi_i}^t(x_i) &= \prod_j \mu_{f_j \rightarrow x_i}^t(x_i) \\ &\propto \mathcal{N}_{\mathbb{C}}(x_i; \zeta_{x_i}^t, \gamma_{x_i}^t), \end{aligned} \quad (21)$$

where

$$\gamma_{x_i}^t = \left( \sum_j \frac{|h_{j,i}|^2}{\nu_{f_j \rightarrow x_i}^t} \right)^{-1}, \quad (22)$$

$$\zeta_{x_i}^t = \gamma_{x_i}^t \sum_j \frac{h_{j,i}^* z_{f_j \rightarrow x_i}^t}{\nu_{f_j \rightarrow x_i}^t}. \quad (23)$$

Using the message  $\mu_{x_i \rightarrow \phi_i}^t(x_i)$  in (21), the LLRs of the coded bits corresponding to the symbol  $x_i$  are given as follows:

$$L_e^t(c_i^q) = \ln \frac{\sum_{\mathcal{A}_q^1} \mathcal{N}_{\mathbb{C}}(x_i; \zeta_{x_i}^t, \gamma_{x_i}^t) \prod_{q' \neq q} p^t(c_i^{q'})}{\sum_{\mathcal{A}_q^0} \mathcal{N}_{\mathbb{C}}(x_i; \zeta_{x_i}^t, \gamma_{x_i}^t) \prod_{q' \neq q} p^t(c_i^{q'})}, \quad (24)$$

for  $q = 1, 2, \dots, Q$ , where  $p^t(c_i^q) = (\exp[c_i^q L^t(c_i^q)]/1 + \exp[L^t(c_i^q)])$ .

We summarize the approximate message passing algorithm using Gaussian approximation in Table I, which will be referred to as AMP-G.

### C. Complexity Reduction Using Expectation Propagation

In the above AMP-G algorithm, the computation of  $\{\mu_{x_i \rightarrow f_j}^t(x_i)\}$  shown by (14) is still expensive. In addition, the number of messages  $\{\mu_{x_i \rightarrow f_j}^t(x_i)\}$  that need to be tracked is as large as  $MN$ . By using expectation propagation in [47] and [53], we can reduce the computational complexity of  $\{\mu_{x_i \rightarrow f_j}^t(x_i)\}$ .

For the Gaussian approximation of the message  $\hat{\mu}_{x_i \rightarrow f_j}^t(x_i)$ , besides using (19) and (20) based on conventional sum-product message passing rule, we can resort to the concept of symbol belief based on expectation propagation. The symbol belief

$$\begin{aligned} \beta_{x_i}^t(x_i) &\triangleq \frac{\mu_{\phi_i \rightarrow x_i}^t(x_i) \prod_j \mu_{f_j \rightarrow x_i}^{t-1}(x_i)}{\sum_{x_i \in \mathcal{A}} \mu_{\phi_i \rightarrow x_i}^t(x_i) \prod_j \mu_{f_j \rightarrow x_i}^{t-1}(x_i)} \\ &= \frac{\mu_{\phi_i \rightarrow x_i}^t(x_i) \mathcal{N}_{\mathbb{C}}(x_i; \zeta_{x_i}^{t-1}, \gamma_{x_i}^{t-1})}{\sum_{x_i \in \mathcal{A}} \mu_{\phi_i \rightarrow x_i}^t(x_i) \mathcal{N}_{\mathbb{C}}(x_i; \zeta_{x_i}^{t-1}, \gamma_{x_i}^{t-1})} \end{aligned} \quad (25)$$

rather than the message  $\mu_{x_i \rightarrow f_j}^t(x_i)$  itself, is replaced by a Gaussian PDF denoted by  $\hat{\beta}_{x_i}^t(x_i) = \mathcal{N}_{\mathbb{C}}(x_i; \hat{x}_{x_i}^t, \hat{\tau}_{x_i}^t)$ , and then the approximate message  $\hat{\mu}_{x_i \rightarrow f_j}^t(x_i)$  is computed from the approximate symbol belief  $\hat{\beta}_{x_i}^t(x_i)$ .  $\hat{\beta}_{x_i}^t(x_i)$  can also be obtained by moment matching as (19) and (20) for the minimum KL divergence, namely  $\hat{x}_{x_i}^t = \mathbb{E}_{\beta_{x_i}^t(x_i)}[x_i]$  and  $\hat{\tau}_{x_i}^t = \mathbb{E}_{\beta_{x_i}^t(x_i)}[|x_i|^2] - |\hat{x}_{x_i}^t|^2$ . Since  $\beta_{x_i}^t(x_i)$  is approximated by the Gaussian PDF  $\hat{\beta}_{x_i}^t(x_i) = \mathcal{N}_{\mathbb{C}}(x_i; \hat{x}_{x_i}^t, \hat{\tau}_{x_i}^t)$  and  $\mu_{f_j \rightarrow x_i}^{t-1}(x_i) = \mathcal{N}_{\mathbb{C}}(h_{j,i} x_i; z_{f_j \rightarrow x_i}^{t-1}, \nu_{f_j \rightarrow x_i}^{t-1})$  is a Gaussian PDF,  $\hat{\mu}_{x_i \rightarrow f_j}^t(x_i)$  is also a Gaussian PDF given by

$$\begin{aligned} \hat{\mu}_{x_i \rightarrow f_j}^t(x_i) &\stackrel{(a)}{\propto} \frac{\beta_{x_i}^t(x_i)}{\mu_{f_j \rightarrow x_i}^{t-1}(x_i)} \\ &\approx \frac{\hat{\beta}_{x_i}^t(x_i)}{\mu_{f_j \rightarrow x_i}^{t-1}(x_i)} \\ &= \mathcal{N}_{\mathbb{C}}(x_i; \hat{x}_{x_i \rightarrow f_j}^t, \hat{\tau}_{x_i \rightarrow f_j}^t), \end{aligned} \quad (26)$$

where (a) is according to the semantics of factor graph,  $\beta_{x_i}^t(x_i) \propto \mu_{x_i \rightarrow f_j}^t(x_i) \mu_{f_j \rightarrow x_i}^{t-1}(x_i)$ , and  $\hat{\tau}_{x_i \rightarrow f_j}^t$  and  $\hat{x}_{x_i \rightarrow f_j}^t$  can be derived using the canonical form of Gaussian PDF [55]

$$\hat{\tau}_{x_i \rightarrow f_j}^t = \left( \frac{1}{\hat{\tau}_{x_i}^t} - \frac{|h_{j,i}|^2}{\nu_{f_j \rightarrow x_i}^{t-1}} \right)^{-1}, \quad (27)$$

$$\hat{x}_{x_i \rightarrow f_j}^t = \hat{\tau}_{x_i \rightarrow f_j}^t \left( \frac{\hat{x}_{x_i}^t}{\hat{\tau}_{x_i}^t} - \frac{h_{j,i}^* z_{f_j \rightarrow x_i}^{t-1}}{\nu_{f_j \rightarrow x_i}^{t-1}} \right). \quad (28)$$

The computation of  $\hat{x}_{x_i}^t = \mathbb{E}_{\beta_{x_i}^t(x_i)}[x_i]$  and  $\hat{\tau}_{x_i}^t = \mathbb{E}_{\beta_{x_i}^t(x_i)}[|x_i|^2] - |\hat{x}_{x_i}^t|^2$  has slightly high complexity. Note that according to the semantics of the factor graph,  $\beta_{x_i}^t(x_i)$  is the *a posteriori* probability of  $x_i$ . To reduce the complexity of computing  $\hat{x}_{x_i}^t$  and  $\hat{\tau}_{x_i}^t$ , we use  $\tilde{p}^t(x_i) = \prod_q \tilde{p}^t(c_i^q)$  to replace  $\beta_{x_i}^t(x_i)$ , where  $\tilde{p}^t(c_i^q)$  is the *a posteriori* probability of the coded bit  $c_i^q$  fed back from the decoders. Then the mean  $\hat{x}_{x_i}^t$  and the variance  $\hat{\tau}_{x_i}^t$  are given by

$$\hat{x}_{x_i}^t = \sum_{\alpha_s \in \mathcal{A}} \alpha_s \tilde{p}^t(x_i = \alpha_s), \quad (29)$$

$$\hat{\tau}_{x_i}^t = \sum_{\alpha_s \in \mathcal{A}} |\alpha_s|^2 \tilde{p}^t(x_i = \alpha_s) - |\hat{x}_{x_i}^t|^2. \quad (30)$$

TABLE II  
THE APPROXIMATION MESSAGE PASSING ALGORITHM SIMPLIFIED  
BY EXPECTATION PROPAGATION FOR THE  $t$ th ITERATION  
BETWEEN THE DECODERS AND THE DETECTOR

1: <b>Initialization:</b>
2: Set $\{z_{f_j \rightarrow x_i}^0 = 0, \nu_{f_j \rightarrow x_i}^0 = \infty, p^1(c_i^q) = \frac{1}{2}, \tilde{p}^1(c_i^q) = \frac{1}{2}\}$ .
3: <b>for</b> $i = 1 \rightarrow N$ <b>do</b> <span style="float: right;">▷ Computation at the cloning nodes</span>
4: $\tilde{p}^t(x_i) = \prod_q p^t(c_i^q)$ ,
5: $\hat{x}_{x_i}^t = \mathbb{E}_{\tilde{p}^t(x_i)}[x_i], \hat{\tau}_{x_i}^t = \mathbb{E}_{\tilde{p}^t(x_i)}[ x_i ^2] -  \hat{x}_{x_i}^t ^2$ .
6: <b>for</b> $j = 1 \rightarrow M$ <b>do</b>
7: $\hat{\tau}_{x_i \rightarrow f_j}^t = \left( \frac{1}{\hat{\tau}_{x_i}^t} - \frac{ h_{j,i} ^2}{\nu_{f_j \rightarrow x_i}^{t-1}} \right)^{-1}$
8: $\hat{x}_{x_i \rightarrow f_j}^t = \hat{\tau}_{x_i \rightarrow f_j}^t \left( \frac{\hat{x}_{x_i}^t}{\hat{\tau}_{x_i}^t} - \frac{h_{j,i}^* z_{f_j \rightarrow x_i}^{t-1}}{\nu_{f_j \rightarrow x_i}^{t-1}} \right)$ ,
9: <b>end for</b>
10: <b>end for</b>
11: <b>for</b> $j = 1 \rightarrow M$ <b>do</b> <span style="float: right;">▷ Computation at the channel transition nodes</span>
12: $\nu_{f_j}^t = \sigma_n^2 + \sum_i  h_{j,i} ^2 \hat{\tau}_{x_i \rightarrow f_j}^t, z_{f_j}^t = y_j - \sum_i h_{j,i} \hat{x}_{x_i \rightarrow f_j}^t$ ,
13: <b>for</b> $i = 1 \rightarrow N$ <b>do</b>
14: $\nu_{f_j \rightarrow x_i}^t = \nu_{f_j}^t -  h_{j,i} ^2 \hat{\tau}_{x_i \rightarrow f_j}^t, z_{f_j \rightarrow x_i}^t = z_{f_j}^t + h_{j,i} \hat{x}_{x_i \rightarrow f_j}^t$ .
15: <b>end for</b>
16: <b>end for</b>
17: <b>for</b> $i = 1 \rightarrow N$ <b>do</b> <span style="float: right;">▷ Computation at the cloning nodes</span>
18: $\gamma_{x_i}^t = \left( \sum_j \frac{ h_{j,i} ^2}{\nu_{f_j \rightarrow x_i}^t} \right)^{-1}, \zeta_{x_i}^t = \gamma_{x_i}^t \sum_j \frac{h_{j,i}^* z_{f_j \rightarrow x_i}^t}{\nu_{f_j \rightarrow x_i}^t}$ ,
19: $\tilde{p}_{eq}^t(x_i) = \mu_{\phi_i \rightarrow x_i}^t(x_i) \mathcal{N}_C(x_i; \zeta_{x_i}^t, \gamma_{x_i}^t)$ ,
20: <b>for</b> $q = 1 \rightarrow Q$ <b>do</b> <span style="float: right;">▷ Computation of LLRs</span>
21: $L_e^t(c_i^q) = \ln \frac{\sum_{A_q^1} \tilde{p}_{eq}^t(x_i)}{\sum_{A_q^0} \tilde{p}_{eq}^t(x_i)} - L^t(c_i^q)$ .
22: <b>end for</b>
23: <b>end for</b>

We summarize the approximate message passing algorithm simplified by expectation propagation in Table II, which will be referred to as AMP-EP.

#### D. Complexity Reduction Using the First-Order Approximation of Messages

To further reduce the complexity, we rewrite the message-passing updating in a form of recursion and omit the negligible terms in the large system limit. As a result, for an  $M \times N$  MIMO system, the number of messages that need to be tracked is reduced to  $\mathcal{O}(M + N)$  per subcarrier.

After a few simple algebraic manipulations, (27) can be rewritten as follows

$$\hat{\tau}_{x_i \rightarrow f_j}^t = \hat{\tau}_{x_i}^t \frac{\nu_{f_j \rightarrow x_i}^{t-1}}{\nu_{f_j \rightarrow x_i}^{t-1} - \hat{\tau}_{x_i}^t |h_{j,i}|^2}, \quad (31)$$

If  $N$  is large, the term  $\hat{\tau}_{x_i}^t |h_{j,i}|^2$  in the denominator of (31) may be neglected, then we have

$$\hat{\tau}_{x_i \rightarrow f_j}^t \approx \hat{\tau}_{x_i}^t. \quad (32)$$

Likewise,  $\nu_{f_j \rightarrow x_i}^{t-1} = \sigma_n^2 + \sum_{i' \neq i} |h_{j,i'}|^2 \hat{\tau}_{x_{i'} \rightarrow f_j}^{t-1}$  can be approximated by  $\nu_{f_j}^{t-1} = \sigma_n^2 + \sum_i |h_{j,i}|^2 \hat{\tau}_{x_i}^{t-1}$ , then (22) and (28) are rewritten as follows

$$\gamma_{x_i}^t = \left( \sum_j \frac{|h_{j,i}|^2}{\nu_{f_j}^t} \right)^{-1}, \quad (33)$$

$$\hat{x}_{x_i \rightarrow f_j}^t \approx \hat{x}_{x_i}^t - \frac{\hat{\tau}_{x_i}^t h_{j,i}^* z_{f_j \rightarrow x_i}^{t-1}}{\nu_{f_j}^{t-1}}. \quad (34)$$

Define  $\lambda_j^t \triangleq \sum_i \hat{\tau}_{x_i}^t |h_{j,i}|^2 z_{f_j \rightarrow x_i}^{t-1}$ . Using (34),  $z_{f_j}^t = y_j - \sum_i h_{j,i} \hat{x}_{x_i \rightarrow f_j}^t$  and  $z_{f_j \rightarrow x_i}^t = y_j - \sum_{i' \neq i} h_{j,i'} \hat{x}_{x_{i'} \rightarrow f_j}^t$  can be rewritten as

$$z_{f_j}^t = y_j - \sum_i h_{j,i} \hat{x}_{x_i}^t + \frac{\lambda_j^t}{\nu_{f_j}^{t-1}}, \quad (35)$$

$$z_{f_j \rightarrow x_i}^t = z_{f_j}^t + h_{j,i} \hat{x}_i^t - \frac{\hat{\tau}_{x_i}^t |h_{j,i}|^2 z_{f_j \rightarrow x_i}^{t-1}}{\nu_{f_j}^{t-1}}. \quad (36)$$

With  $z_{f_j \rightarrow x_i}^t$  in (36),  $\lambda_j^t$  and  $\zeta_{x_i}^t$  can be written as

$$\begin{aligned} \lambda_j^t &= \sum_i \hat{\tau}_{x_i}^t |h_{j,i}|^2 (z_{f_j}^{t-1} + h_{j,i} \hat{x}_i^{t-1}) \\ &\quad - \sum_{i'} \frac{\hat{\tau}_{x_i}^t |h_{j,i'}|^2}{\nu_{f_j}^{t-2}} \hat{\tau}_{x_{i'}}^{t-1} |h_{j,i'}|^2 z_{f_j \rightarrow x_{i'}}^{t-2} \\ &\approx \sum_i \hat{\tau}_{x_i}^t |h_{j,i}|^2 (z_{f_j}^{t-1} + h_{j,i} \hat{x}_i^{t-1}) - \frac{\sum_{i'} \hat{\tau}_{x_{i'}}^t |h_{j,i'}|^2}{N \nu_{f_j}^{t-2}} \lambda_j^{t-1} \\ &= \sum_i \hat{\tau}_{x_i}^t |h_{j,i}|^2 \left( z_{f_j}^{t-1} + h_{j,i} \hat{x}_i^{t-1} - \frac{\lambda_j^{t-1}}{N \nu_{f_j}^{t-2}} \right), \end{aligned} \quad (37)$$

$$\begin{aligned} \zeta_{x_i}^t &= \hat{x}_{x_i}^t + \gamma_{x_i}^t \sum_j \frac{h_{j,i}^* z_{f_j}^t}{\nu_{f_j}^t} - \gamma_{x_i}^t \sum_{j'} \frac{\hat{\tau}_{x_i}^t |h_{j',i}|^2 h_{j',i}^* z_{f_{j'} \rightarrow x_i}^{t-1}}{\nu_{f_{j'} \rightarrow x_i}^t \nu_{f_{j'} \rightarrow x_i}^{t-1}} \\ &\approx \hat{x}_{x_i}^t + \gamma_{x_i}^t \sum_j \frac{h_{j,i}^* z_{f_j}^t}{\nu_{f_j}^t} - \frac{\gamma_{x_n}^t}{\nu_{x_n}^{t-1}} \frac{\sum_{j'} \frac{\hat{\tau}_{x_i}^t |h_{j',i}|^2}{\nu_{f_{j'} \rightarrow x_i}^t}}{M} \zeta_{x_i}^{t-1}. \end{aligned} \quad (38)$$

If  $N$  is large, the term  $((\sum_{i'} \hat{\tau}_{x_i}^t |h_{j,i'}|^2) / (N \nu_{f_j}^{t-2})) \lambda_j^{t-1} \approx \lambda_j^{t-1} / N$  in (37) and the term  $(\gamma_{x_n}^t / \nu_{x_n}^{t-1}) (\sum_{j'} ((\hat{\tau}_{x_i}^t |h_{j',i}|^2) / \nu_{f_{j'} \rightarrow x_i}^t) / M) \zeta_{x_i}^{t-1} \approx \zeta_{x_i}^{t-1} / N$  in (38) can be safely neglected. As a result, we have

$$\begin{aligned} z_{f_j}^t &= y_j - \sum_i h_{j,i} \hat{x}_{x_i}^t + \frac{\sum_{i'} \hat{\tau}_{x_i}^t |h_{j,i'}|^2 (z_{f_j}^{t-1} + h_{j,i'} \hat{x}_{i'}^{t-1})}{\nu_{f_j}^{t-1}} \\ &\approx y_j - \sum_i h_{j,i} \hat{x}_{x_i}^t + \frac{\sum_{i'} \hat{\tau}_{x_i}^t |h_{j,i'}|^2}{\nu_{f_j}^{t-1}} z_{f_j}^{t-1}, \end{aligned} \quad (39)$$

$$\zeta_{x_i}^t \approx \hat{x}_{x_i}^t + \gamma_{x_i}^t \sum_j \frac{h_{j,i}^* z_{f_j}^t}{\nu_{f_j}^t}. \quad (40)$$

From (33) and (40), it can be found that the number of messages to be tracked is reduced to  $\mathcal{O}(M + N)$ . Following the approach of AMP algorithm in [38] and [37], it is verified in the Appendix that  $\hat{x}_{x_i \rightarrow f_j}^t$  given by (34) is in the form of the first-order approximation (linear approximation) as in the AMP. We summarize the approximate message passing algorithm simplified by the first-order approximation in Table III, which will be referred to as AMP-LA.

Similar to the application of GAMP [38] in frequency domain equalization for single-antenna systems [39], we can employ GAMP algorithm to estimate the discrete random variables  $\{x_i\}$  in MIMO-OFDM systems. By examining the derivation of GAMP algorithm in [38] and the relationship between sum-product GAMP algorithm and AMP algorithm in [37], we can show that the sum-product GAMP algorithm applied to the detection for MIMO-OFDM systems is equivalent to the

TABLE III  
THE APPROXIMATION MESSAGE PASSING ALGORITHM SIMPLIFIED  
BY THE FIRST-ORDER APPROXIMATION FOR THE  $t$ th ITERATION  
BETWEEN THE DECODERS AND THE DETECTOR

---

```

1: Initialization:
2: Set  $\{z_{f_j \rightarrow x_i}^0 = 0, \nu_{f_j \rightarrow x_i}^0 = \infty, p^1(c_i^q) = \frac{1}{2}, \tilde{p}^1(c_i^q) = \frac{1}{2}\}$ .
3: for  $i = 1 \rightarrow N$  do ▷ Computation at the cloning nodes
4:    $\tilde{p}^t(x_i) = \prod_q \tilde{p}^t(c_i^q)$ ,
5:    $\hat{x}_{x_i}^t = \mathbb{E}_{\tilde{p}^t(x_i)}[x_i], \hat{\tau}_{x_i}^t = \mathbb{E}_{\tilde{p}^t(x_i)}[|x_i|^2] - |\hat{x}_{x_i}^t|^2$ .
6: end for
7: for  $j = 1 \rightarrow M$  do ▷ Computation at the channel transition nodes
8:    $\nu_{f_j}^t = \sigma_n^2 + \sum_i |h_{j,i}|^2 \hat{\tau}_{x_i}^t$ ,
9:    $z_{f_j}^t = y_j - \sum_i h_{j,i} \hat{x}_{x_i}^t + z_{f_j}^{t-1} \frac{\sum_{i'} \hat{\tau}_{x_i}^t |h_{j,i'}|^2}{\nu_{f_j}^{t-1}}$ .
10: end for
11: for  $i = 1 \rightarrow N$  do ▷ Computation at the cloning nodes
12:    $\gamma_{x_i}^t = \left( \sum_j \frac{|h_{j,i}|^2}{\nu_{f_j}^t} \right)^{-1}, \zeta_{x_i}^t = \hat{x}_{x_i}^t + \gamma_{x_i}^t \sum_j \frac{h_{j,i}^* z_{f_j}^t}{\nu_{f_j}^t}$ .
13:    $\tilde{p}_{eq}(x_i) = \mu_{\phi_i \rightarrow x_i}^t(x_i) \mathcal{N}_{\mathbb{C}}(x_i; \zeta_{x_i}^t, \gamma_{x_i}^t)$ ,
14:   for  $q = 1 \rightarrow Q$  do ▷ Computation of LLRs
15:      $L_e^t(c_i^q) = \ln \frac{\sum_{\mathcal{A}_q^1} \tilde{p}_{eq}^t(x_i)}{\sum_{\mathcal{A}_q^0} \tilde{p}_{eq}^t(x_i)} - L^t(c_i^q)$ .
16:   end for
17: end for

```

---

AMP-LA algorithm in our work. However, the main ideas behind the AMP-LA and GAMP are different. In our derivation the variables  $\{x_i\}$  are approximated into continuous Gaussian variables. As a result, the messages  $\mu_{f_j \rightarrow x_i}^t(x_i)$  can be analytically calculated by integration. In the GAMP, the random variable  $\sum_{i' \neq i} h_{j,i'} x_{i'}$  is approximated as Gaussian using the central limit theorem, and the discrete message  $\mu_{f_j \rightarrow x_i}^t(x_i)$  is approximated as a Gaussian function using a second-order Taylor series [40] (Note that for an AWGN channel  $\mu_{f_j \rightarrow x_i}^t(x_i)$  itself is a Gaussian function without Taylor approximation). The another difference is the way of calculating the parameters (mean and variance) of  $\mu_{x_i \rightarrow f_i}^t(x_i)$ . Our derivation employs the principle of expectation propagation, while the GAMP employs the first order Taylor series.

#### E. Complexity Reduction Using Central-Limit Theorem

For a large-scale system with lots of transmit antennas and receive antennas, according to the central-limit theorem, it is reasonable to approximate  $\sum_i |h_{j,i}|^2$  with the average value  $\bar{h}_r^2 = (1/M) \sum_j \sum_i |h_{j,i}|^2$ , and to approximate  $\sum_j |h_{j,i}|^2$  with the average value  $\bar{h}_c^2 = (M/N) \bar{h}_r^2$  [38]. Furthermore,  $\hat{\tau}_{x_i}^t$  can be approximated by the average variance  $\bar{\tau}_x^t = (1/N) \sum_i \hat{\tau}_{x_i}^t$ , which is often used in the frequency domain equalization [41]. Then,  $\tau_{f_j}^t$  and  $\gamma_{x_i}^t$  can be rewritten as

$$\nu_f^t = \sigma_n^2 + \bar{h}_r^2 \bar{\tau}_x^t, \quad (41)$$

$$\gamma_x^t = \frac{\nu_f^t}{\bar{h}_c^2}, \quad (42)$$

and  $\zeta_{x_i}^t$  and  $z_{f_j}^t$  can be rewritten as

$$\zeta_{x_i}^t = \hat{x}_{x_i}^t + \frac{\sum_j h_{j,i}^* z_{f_j}^t}{\bar{h}_c^2}, \quad (43)$$

$$z_{f_j}^t = y_j - \sum_i h_{j,i} \hat{x}_{x_i}^t + z_{f_j}^{t-1} \frac{\bar{h}_r^2 \bar{\tau}_x^t}{\nu_{f_j}^{t-1}}. \quad (44)$$

TABLE IV  
THE APPROXIMATION MESSAGE PASSING ALGORITHM SIMPLIFIED  
BY THE CENTRAL-LIMIT THEOREM FOR THE  $t$ th ITERATION  
BETWEEN THE DECODERS AND THE DETECTOR

---

```

1: Initialization: Set  $\bar{h}_r^2 = \frac{1}{M} \sum_j \sum_i |h_{j,i}|^2, \bar{h}_c^2 = \frac{M}{N} \bar{h}_r^2$ ,
2: and  $\{\zeta_{x_i \rightarrow f_j}^0 = 0, \gamma_{x_i \rightarrow f_j}^0 = \infty, p^1(c_i^q) = \frac{1}{2}, \tilde{p}^1(c_i^q) = \frac{1}{2}\}$ .
3: for  $i = 1 \rightarrow N$  do ▷ Computation at the cloning nodes
4:    $\tilde{p}^t(x_i) = \prod_q \tilde{p}^t(c_i^q)$ ,
5:    $\hat{x}_{x_i}^t = \mathbb{E}_{\tilde{p}^t(x_i)}[x_i], \hat{\tau}_{x_i}^t = \mathbb{E}_{\tilde{p}^t(x_i)}[|x_i|^2] - |\hat{x}_{x_i}^t|^2$ .
6: end for
7:  $\bar{\tau}_x^t = \frac{\sum_i \hat{\tau}_{x_i}^t}{N}, \nu_f^t = \sigma_n^2 + \bar{h}_r^2 \bar{\tau}_x^t, \gamma_x^t = \frac{\nu_f^t}{\bar{h}_c^2}$ .
8: for  $j = 1 \rightarrow M$  do ▷ Computation at the channel transition nodes
9:    $z_{f_j}^t = y_j - \sum_i h_{j,i} \hat{x}_{x_i}^t + z_{f_j}^{t-1} \frac{\bar{h}_r^2 \bar{\tau}_x^t}{\nu_{f_j}^{t-1}}$ .
10: end for
11: for  $i = 1 \rightarrow N$  do ▷ Computation at the cloning nodes
12:    $\zeta_{x_i}^t = \hat{x}_{x_i}^t + \frac{\sum_j h_{j,i}^* z_{f_j}^t}{\bar{h}_c^2}$ .
13:    $\tilde{p}_{eq}(x_i) = \mu_{\phi_i \rightarrow x_i}^t(x_i) \mathcal{N}_{\mathbb{C}}(x_i; \zeta_{x_i}^t, \gamma_x^t)$ ,
14:   for  $q = 1 \rightarrow Q$  do ▷ Computation of LLRs
15:      $L_e^t(c_i^q) = \ln \frac{\sum_{\mathcal{A}_q^1} \tilde{p}_{eq}^t(x_i)}{\sum_{\mathcal{A}_q^0} \tilde{p}_{eq}^t(x_i)} - L^t(c_i^q)$ .
16:   end for
17: end for

```

---

Using (41)–(44), the AMP-LA is reduced to the algorithm as shown in Table IV, which will be referred to as AMP-LS.

#### F. Complexity Analysis

As shown in Tables V and VI, the complexity is evaluated in terms of floating-point operations (FLOPs) counted per sub-carrier per iteration. As the complexity of STS-SD is random, only the complexity of the turbo MMSE-SIC detector is considered for comparison. It is assumed that the operation of  $\exp(\cdot)$  can be implemented by a look-up table,  $\{p^t(c_i^q)\}$  and  $\{\tilde{p}^t(c_i^q)\}$  are calculated by the decoders. Note that the multiplication of a complex number and a real number needs two FLOPs, and the multiplication of two complex numbers (excluding conjugate numbers) needs six FLOPs.

All the message passing algorithms requires  $3MN$  FLOPs to compute  $|\mathbf{H}|^2$  (the component-wise squared magnitude) in the preprocessing stage, and the AMP-LS algorithm requires  $MN + 1$  FLOPs for the calculation of  $\bar{h}_r^2$  and  $\bar{h}_c^2$ . For computing the downward messages at the cloning nodes,  $\{\tilde{p}^t(x_i), \forall i\}$  or  $\{\mu_{\phi_i \rightarrow x_i}^t(x_i), \forall i\}$  needs  $N|\mathcal{A}|(Q - 1)$  FLOPs,  $\{\mu_{x_i \rightarrow f_j}^t(x_i), \forall i, \forall j\}$  and  $\{\hat{x}_{x_i \rightarrow f_j}^t, \hat{\tau}_{x_i \rightarrow f_j}^t, \forall i, \forall j\}$  in the AMP-G algorithm need  $(11|\mathcal{A}| - 1)MN$  FLOPs and  $(6|\mathcal{A}| + 1)MN$  FLOPs, respectively.  $\{\hat{x}_{x_i}^t, \hat{\tau}_{x_i}^t, \hat{x}_{x_i \rightarrow f_j}^t, \hat{\tau}_{x_i \rightarrow f_j}^t, \forall i, \forall j\}$  in the AMP-EP algorithm needs  $(6|\mathcal{A}| + 4)N + 15MN$  FLOPs,  $\{\hat{x}_{x_i}^t, \hat{\tau}_{x_i}^t, \hat{x}_{x_i \rightarrow f_j}^t, \hat{\tau}_{x_i \rightarrow f_j}^t, \forall i, \forall j\}$  in the AMP-LA algorithm and the AMP-LS algorithm both needs  $(6|\mathcal{A}| + 1)N$  FLOPs. Additionally,  $\bar{\tau}_x^t$  in the AMP-LS algorithm needs  $N$  FLOPs.

For computing the upward messages at the channel transition nodes,  $\{\nu_{f_j}^t, z_{f_j}^t, \nu_{f_j \rightarrow x_i}^t, z_{f_j \rightarrow x_i}^t, \forall i, \forall j\}$  in the AMP-G algorithm and the AMP-EP algorithm both need  $13MN$  FLOPs;  $\{\nu_{f_j}^t, z_{f_j}^t, \forall j\}$  in the AMP-LA algorithm needs  $10MN + 3M$  FLOPs; and  $\{\nu_f^t, z_f^t, \forall j\}$  in the AMP-LS algorithm needs

TABLE V  
COMPLEXITY COMPARISON IN TERMS OF FLOPS PER SUBCARRIER PER ITERATION

	Preprocessing	Cloning nodes	Observation nodes	Cloning nodes again
AMP-G	$3MN$	$17 \mathcal{A} MN + (Q-1) \mathcal{A} N$	$13MN$	$16MN + [(Q+7) \mathcal{A} -2]N$
AMP-EP	$3MN$	$15MN + (Q+5) \mathcal{A} N + 4N$	$13MN$	$12MN + [(Q+7) \mathcal{A} +1]N$
AMP-LA	$3MN$	$(Q+5) \mathcal{A} N + N$	$10MN + 3M$	$12MN + [(Q+7) \mathcal{A} +3]N$
AMP-LS	$4MN + 1$	$(Q+5) \mathcal{A} N + 2N$	$8MN + 3M + 2$	$8MN + [(Q+7) \mathcal{A} +3]N + 1$
Turbo MMSE-SIC	–	$(Q+5) \mathcal{A} N + 2N$	$(8N-2)M^2 + 2MN + M$	$8M^3 + 8M^2N + 22MN + 8M^2 - 2M + [(Q+7) \mathcal{A} +3]N$

TABLE VI  
TOTAL COMPLEXITY PER SUBCARRIER PER ITERATION

AMP-G	$(17 \mathcal{A}  + 32)MN + (2Q + 6) \mathcal{A} N - 2N$
AMP-EP	$43MN + (2Q + 12) \mathcal{A} N + 5N$
AMP-LA	$25MN + (2Q + 12) \mathcal{A} N + 4N + 3M$
AMP-LS	$20MN + (2Q + 12) \mathcal{A} N + 5N + 3M + 4$
Turbo MMSE-SIC	$8M^3 + 8M^2N + 8NM^2 + 6M^2 + 24MN + (2Q + 12) \mathcal{A} N + 5N - M$

$8MN + 3M + 2$  FLOPs. Note that for computing  $z_{f_j}^t$ , we have used the intermediate result in the computation of  $\nu_{f_j}^t$ .

For computing the messages at the cloning nodes, the AMP-G algorithm needs  $16MN - 3N$  FLOPs to compute  $\{\zeta_{x_i}^t, \gamma_{x_i}^t, \zeta_{f_j \rightarrow x_i}^t, \gamma_{f_j \rightarrow x_i}^t \forall i, \forall j\}$ , the AMP-EP algorithm and the AMP-LA algorithm need  $12MN$  FLOPs and  $12MN + 2N$  FLOPs, respectively, to compute  $\{\gamma_{x_i}^t, \zeta_{x_i}^t, \forall i\}$ , and the AMP-LA algorithm needs  $8MN + 2N + 1$  FLOPs to compute  $\{\gamma_{x_i}^t, \zeta_{x_i}^t, \forall i\}$ . Finally, the complexity of computing  $\{L_e^t(c_e^q), \forall i, q\}$  is  $[(Q+7)|\mathcal{A}|+1]N$  FLOPs.

Next, we compare the complexity of turbo MMSE-SIC detection algorithm with that of the message passing algorithms. All these algorithms need to compute the mean and variance of each symbol as well as the extrinsic LLR of each code bit. The complexity of computing  $\mathbf{V}_y$  in the turbo MMSE-SIC algorithm is similar to computing  $\{\nu_{f_j}^t, \forall j\}$  or  $\{\nu_{f_j \rightarrow x_i}^t, \forall j, \forall i\}$  in the message passing algorithms, and the extrinsic mean and variance of  $x_i$  shown in (5) and (6) are similar to  $\gamma_{x_i}^t$  and  $\zeta_{x_i}^t$  in the message passing algorithms. The computations of  $\mathbf{V}_y$  needs  $(8N-2)M^2 + 2MN + M$  FLOPs per subcarrier per iteration and the extrinsic mean together with the extrinsic variance needs  $8M^3 + 22MN + 8M^2 - 2M + 8M^2N + 2N$  operations per subcarrier per iteration. Normalized over the complexity of the AMP-LS algorithm, the complexity of the considered algorithms is shown in Fig. 3.

#### IV. SIMULATION RESULTS

We consider the uplink of a multiuser system with  $N$  independent users, and each user is equipped with one transmit antenna. For each user, the transmission is based on OFDM with  $K = 128$  subcarriers, and 8192 information bits are encoded by a  $R = 1/2$  recursive systematic convolutional (RSC) code with generator polynomial  $[G_1, G_2] = [117, 155]_{\text{oct}}$ , and then interleaved by an S-random interleaver with  $S = 32$  [56]. The conventional BCJR algorithm is used to decode the convolutional codes. The channel model in the simulations is a 16-tap Rayleigh fading MIMO channel with equal tap power. We assume that the transmit antennas from different users are spatially uncorrelated, and the receive antennas are spaced sufficiently

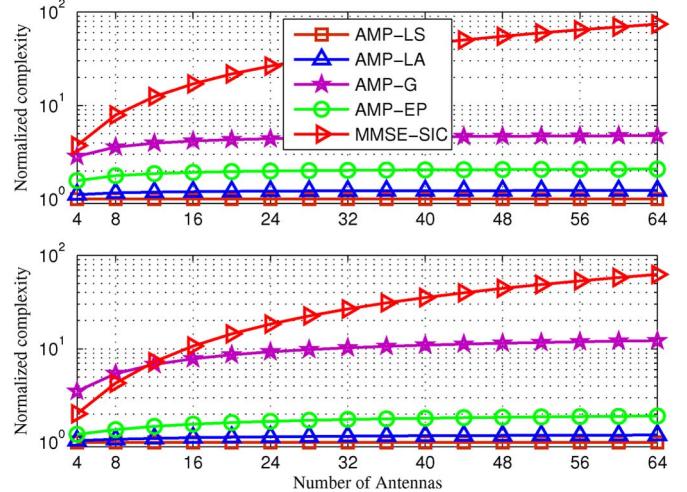


Fig. 3. Complexity versus number of antennas  $N$  for the algorithms in the  $N \times N$  MIMO-OFDM systems with QPSK (top) and 16QAM (bottom). The complexity is normalized over the complexity of AMP-LS algorithm.

so that they are also spatially uncorrelated. For each simulated point, a minimum of 100 frame errors were counted. The energy per bit to noise power spectral density ratio  $E_b/N_0$  is defined as [14]

$$\frac{E_b}{N_0} = \frac{E_s}{N_0} + 10 \log_{10} \frac{M}{RNQ}, \quad (45)$$

where  $E_s/N$  is the average energy per transmitted symbol.

First, we consider a large-scale MIMO system with  $N = 64$  users and  $M = 64$  antennas at the receiver. We make comparisons between our proposed message passing algorithms and the turbo MMSE-SIC algorithm that has often been used as a benchmark. For the turbo MMSE-SIC algorithm, the extrinsic information fed back from the decoders is used to estimate the interference. The matched filter bound (MFB) obtained by exactly removing all the interferences is used as a lower bound on the bit error rate (BER). In the following figures, “QPSK-#” (“16QAM-#”) indicates that QPSK (16QAM) modulation is

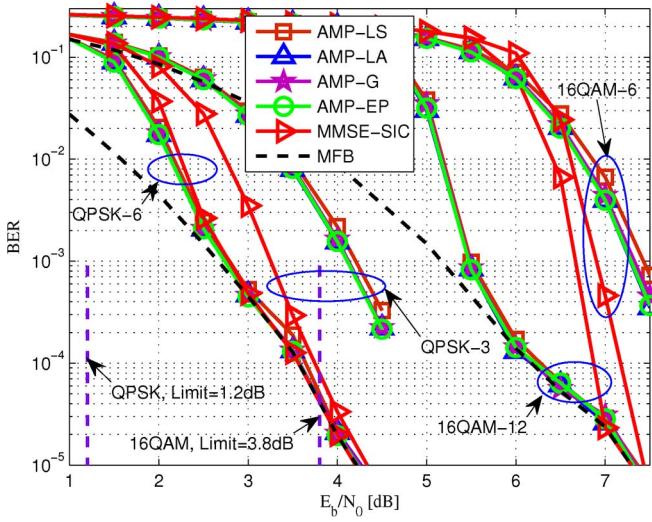


Fig. 4. BER performance for the algorithms in a  $64 \times 64$  MIMO system with QPSK and 16QAM.

used and the number of iterations is  $\#$ . The BER performance of the considered algorithms is shown in Fig. 4. For QPSK transmission, all the approximate message passing algorithms achieve the same performance as that of the turbo MMSE-SIC algorithm with 6 iterations, although the latter outperforms them with 3 iterations; for 16QAM transmission, the turbo MMSE-SIC algorithm outperforms the proposed approximate message passing algorithms with 6 iterations, but with more iterations, e.g., 12 iterations, all the approximate message passing algorithms reach the MFB and outperform the turbo MMSE-SIC algorithm about 0.7 dB at  $\text{BER} = 10^{-4}$ . The performance improvement may be due to the different way of forming the interference estimate. The turbo MMSE-SIC algorithm uses the extrinsic information fed back from the decoders, i.e.,  $\mu_{\phi_i \rightarrow x_i}^t(x_i)$ , to calculate the soft interference estimates. In contrast, in the proposed approximate message passing algorithms, the interference estimates are calculated more precisely using the “a-posteriori-like” information  $\mu_{x_i \rightarrow f_j}^t(x_i)$  at the equalizer output, as shown in (14), or the *a posteriori* information fed back from the decoders,  $\{\hat{p}^t(x_i)\}$ . In Fig. 4, vertical lines indicate the respective capacity limits (Gaussian input), i.e., the minimum  $E_b/N_0$  required for the target rate. It is seen that the  $64 \times 64$  system with QPSK achieving 64 bits per channel use (bpcu) is roughly 3.7 dB from capacity and the case with 16QAM achieving 128 bpcu is roughly 3.5 dB from capacity.

Fig. 5 presents the BER performance of the AMP-EP algorithm and the turbo MMSE-SIC algorithm with the number of iterations. It can be see that 5 iterations and 9 iterations are enough for the AMP-EP algorithm to achieve the MFBs of QPSK and 16QAM at  $\text{BER} = 10^{-4}$ , respectively. While the AMP-EP algorithm can uniformly improve the performance with iterations and achieve the MFB in the low  $E_b/N_0$  region, the turbo MMSE-SIC algorithm cannot achieve the MFB in the low  $E_b/N_0$  region with iterations.

Fig. 6 shows the BER performance versus the number of iterations for these algorithms in a  $64 \times 64$  MIMO system.

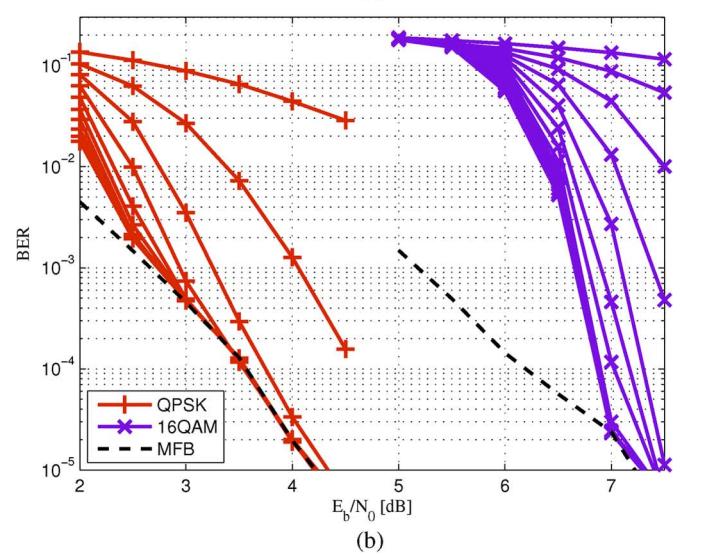
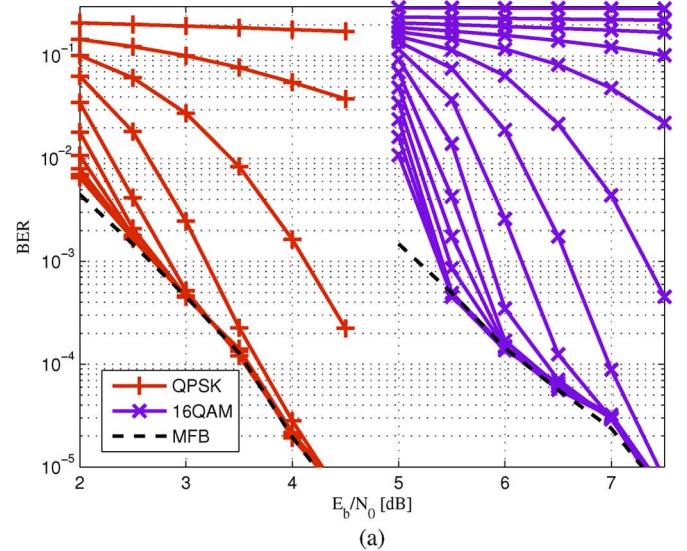


Fig. 5. BER performance versus  $E_b/N_0$  with multiple iterations in a  $64 \times 64$  MIMO system with QPSK and 16QAM. (a) The AMP-EP algorithm; (b) The turbo MMSE-SIC algorithm.

For QPSK transmission, compared with MMSE-SIC, the proposed approximate message passing algorithms need only one more iteration to converge. For 16QAM transmission, all the message passing algorithms outperform the turbo MMSE-SIC algorithm at  $E_b/N_0 = 5.5$  dB; the proposed message passing algorithms and the turbo MMSE-SIC algorithm need the same number of iterations to converge to the same performance at  $E_b/N_0 = 7.0$  dB. On the other hand, as shown in Fig. 3, the complexity per iteration of the turbo MMSE-SIC algorithm is about 70 and 60 times of that of the AMP-LS algorithm for the system with QPSK and 16QAM, respectively.

In order to investigate the robustness of the proposed approximate message passing algorithms, we consider a low-dimensional MIMO system with  $N = 4$  users and  $M = 4$  antennas at the receiver. Additionally, we use the STS-SD algorithm (for QPSK and 16QAM cases) and the optimal MAP algorithm (for QPSK case) by using (4) directly as benchmarks. For the STS-SD algorithm, we adopt the exact “clipping” approach to mitigating LLR estimation errors, and the estimated

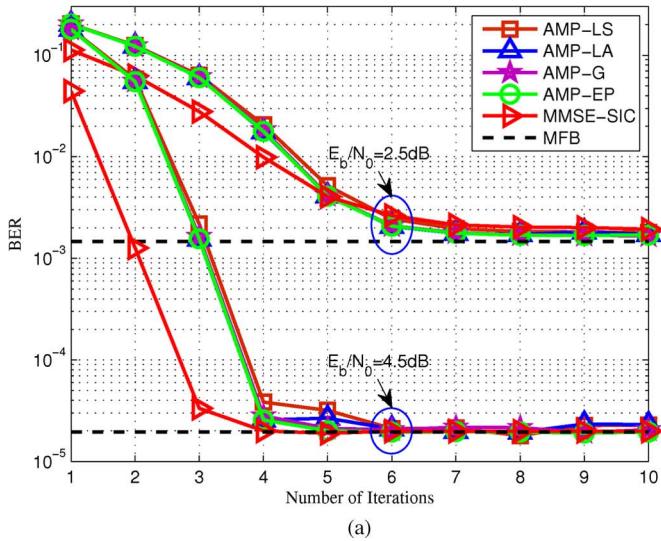


Fig. 6. BER performance versus number of iterations for the algorithms in a  $64 \times 64$  MIMO system with QPSK and 16QAM. (a) QPSK; (b) 16QAM.

LLRs are clipped to the interval  $[-5, +5]$ . The BER performance is shown in Fig. 7. For the system with QPSK, it is apparent that the AMP-G algorithm and the AMP-EP algorithm perform as well as the MAP algorithm, the STS-SD algorithm, and the turbo MMSE-SIC algorithm after 6 iterations, and the performance of the AMP-LA algorithm and the AMP-LS algorithm is degraded about 0.5 dB and 1.0 dB at  $\text{BER} = 10^{-5}$ , respectively. Similarly, for the system with 16QAM, the AMP-G algorithm and the AMP-EP algorithm reach the MFB at  $\text{BER} = 10^{-5}$  as the STS-SD algorithm and the turbo MMSE-SIC algorithm do after 12 iterations, and the AMP-LA algorithm and AMP-LS algorithm are deteriorated by 0.5 dB and 1.0 dB at  $\text{BER} = 10^{-4}$ .

Fig. 8 presents the BER performance with the number of iterations for the AMP-EP algorithm, the STS-SD algorithm and the turbo MMSE-SIC algorithm, respectively. It is also shown that 4 iterations and 9 iterations are enough for the AMP-EP algorithm to achieve the MFB of system with QPSK and 16QAM at  $\text{BER} = 10^{-3}$ , respectively, and more iterations could only improve the performance above  $\text{BER} = 10^{-3}$ .

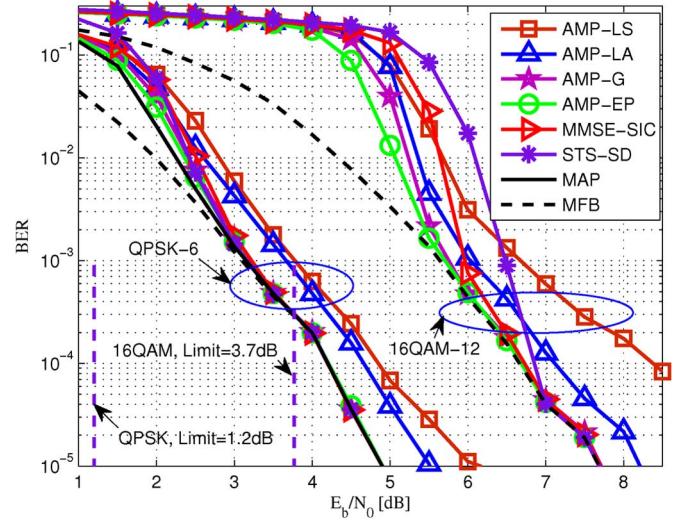


Fig. 7. BER performance of the algorithms in a  $4 \times 4$  MIMO system with QPSK and 16QAM.

While the AMP-EP algorithm can uniformly improve the performance with iterations and achieve the MFB even in the low  $E_b/N_0$  region, both the STS-SD algorithm and the turbo MMSE-SIC algorithm cannot achieve the MFB in the low region of  $E_b/N_0$  with iterations.

Fig. 9 shows the BER performance of system with QPSK versus number of iterations. For QPSK transmission, the AMP-G algorithm and the AMP-EP algorithm need only one more iteration to converge than the MAP algorithm, the STS-SD algorithm and the turbo MMSE-SIC algorithm. For 16QAM transmission, all the message passing algorithms outperform the turbo MMSE-SIC algorithm and the STS-SD algorithm at  $E_b/N_0 = 5.5\text{ dB}$ ; when at  $E_b/N_0 = 7.5\text{ dB}$ , the AMP-G algorithm and the AMP-EP algorithm need the same number of iterations to converge to the same performance as the turbo MMSE-SIC algorithm and the STS-SD algorithm, but the performance of the AMP-LA algorithm and the AMP-LS algorithm is degraded, which can also be observed in Fig. 7.

Compared with the performance in the large-scale MIMO system, the performance of the AMP-LA and AMP-LS is degraded slightly in the low-dimensional systems, such as  $N = M = 4$ . This may be due to the fact that, when  $N$  is not large enough, the approximation made in the message updating (37)–(39) is not valid and the central limit theorem may not apply in (41)–(44). Fig. 10 presents the performance for a  $16 \times 16$  MIMO system and a  $32 \times 32$  MIMO system. In the  $16 \times 16$  case, it is shown that the performance gap between the AMP-LS algorithm and the AMP-LA algorithm disappears, which indicates that  $N = 16$  is enough for applying the central limit theorem to the AMP-LA algorithm (resulting in the AMP-LS algorithm). On the other hand, the performance gap between the AMP-LA algorithm and the MFB becomes small as the number of transmit antennas  $N$  increases, although the performance of AMP-LA algorithm is still degraded slightly from the MFB in the  $32 \times 32$  case. From all the numerical results ( $N = M = 4, 16, 32, 64$ ), we may conclude that: in the low-dimensional  $N \times N$  MIMO systems such as  $N \leq 16$ , the AMP-EP

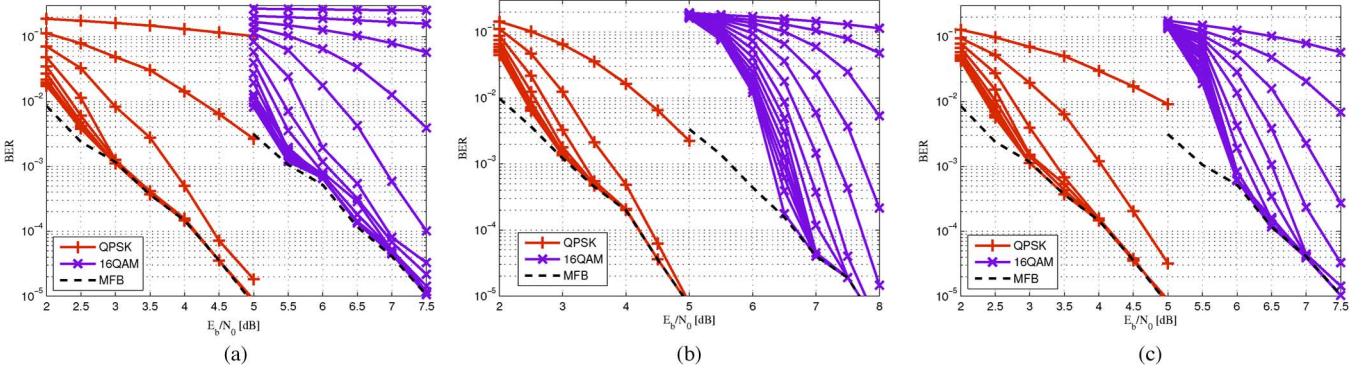


Fig. 8. BER performance versus  $E_b/N_0$  with multiple iterations in a  $4 \times 4$  MIMO system with QPSK and 16QAM. (a) The AMP-EP algorithm; (b) The STS-SD algorithm; (c) The turbo MMSE-SIC algorithm.

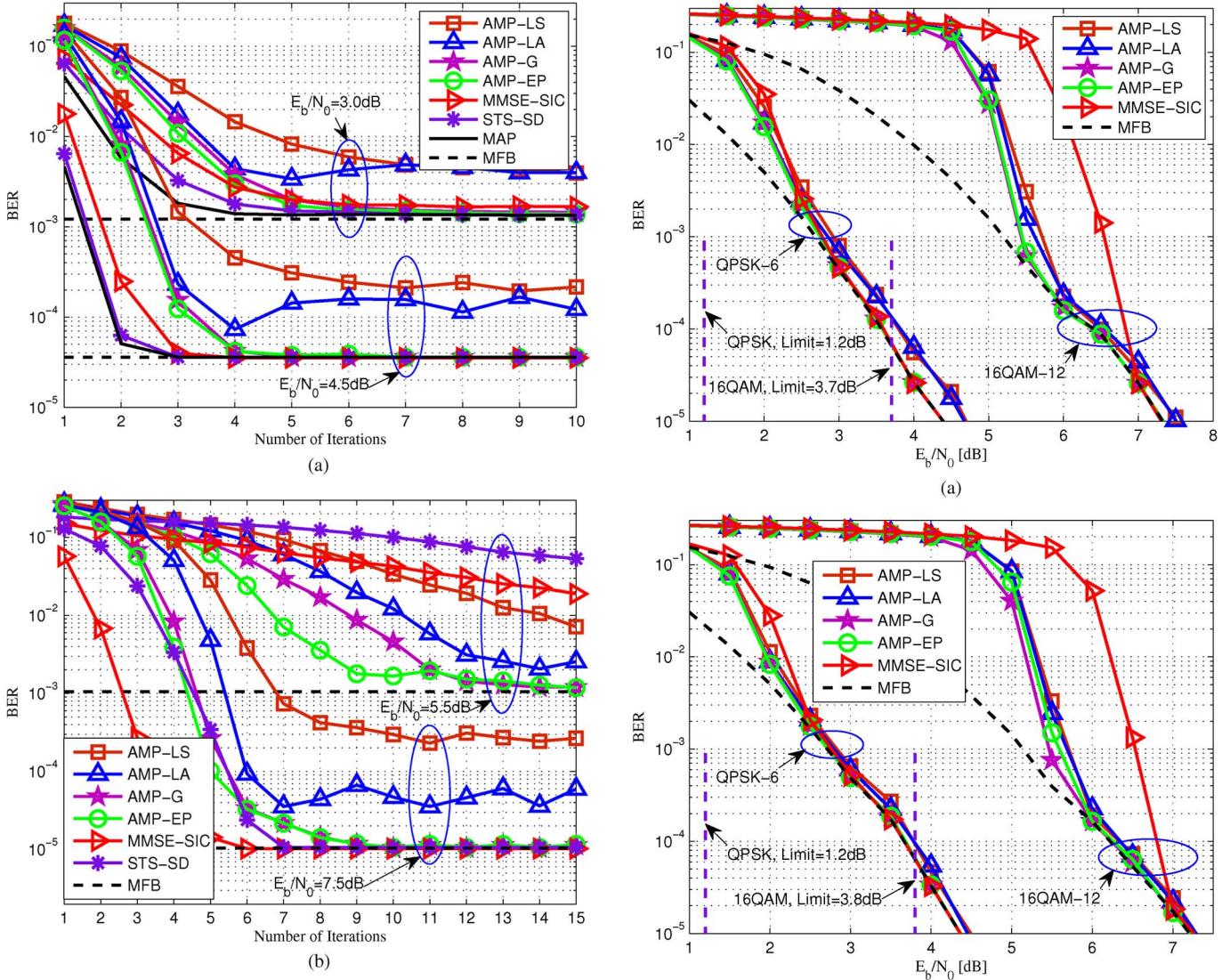


Fig. 9. BER performance versus number of iterations for the algorithms in a  $4 \times 4$  MIMO system with QPSK and 16QAM. (a) QPSK; (b) 16QAM.

algorithm is more suitable, as its performance is as good as that of the turbo MMSE-SIC and AMP-G, but with lower complexity. In the high-dimensional  $N \times N$  MIMO systems such as  $N \geq 32$ , the AMP-LS algorithm turns out to be a good choice, as it can perform as well as the AMP-LA algorithm that can achieve near optimal performance when  $N \geq 32$ .

Fig. 10. BER performance of the algorithms in a MIMO system with QPSK and 16QAM. (a) a  $16 \times 16$  MIMO system; (b) a  $32 \times 32$  MIMO system.

## V. CONCLUSION

For the detection of large-scale multiuser MIMO-OFDM systems, we have proposed a range of low-complexity approximate message passing algorithms that can offer desirable tradeoff

between performance and complexity. It is verified through extensive simulations that our proposed approximate message passing algorithms can achieve near optimal performance with low complexity. Compared with existing turbo detection algorithms, the proposed schemes can achieve or even outperform the performance of some complex algorithms, such as the iterative decoding based on STS-SD and MMSE-SIC. In addition, the number of iterations required to achieve near-optimal performance is small and does not increase with the system dimension.

## APPENDIX

Following the approach in [37] and [38],  $\gamma_{x_i \rightarrow f_j}^{t-1} = (\sum_{j' \neq j} (|h_{j',i}|^2 / \nu_{f_{j'} \rightarrow x_i}^{t-1}))^{-1}$  and  $\zeta_{x_i \rightarrow f_j}^{t-1}$  in (14) are replaced, respectively, by  $\gamma_{x_i}^{t-1} = (\sum_j (|h_{j,i}|^2 / \nu_{f_j}^{t-1}))^{-1}$  and  $\zeta_{x_i \rightarrow f_j}^{t-1} = \gamma_{x_i}^{t-1} \sum_{j' \neq j} h_{j',i}^* (z_{f_{j'} \rightarrow x_i}^{t-1} / \nu_{f_{j'}}^{t-1})$ . Then  $\mu_{x_i \rightarrow f_j}^t(x_i)$  is normalized as follows

$$\mu_{x_i \rightarrow f_j}^t(x_i) \approx \frac{\mu_{\phi_i \rightarrow x_i}^t(x_i) \mathcal{N}_{\mathbb{C}}(x_i; \zeta_{x_i \rightarrow f_j}^{t-1}, \gamma_{x_i}^{t-1})}{\sum_{x_i \in \mathcal{A}} \mu_{\phi_i \rightarrow x_i}^t(x_i) \mathcal{N}_{\mathbb{C}}(x_i; \zeta_{x_i \rightarrow f_j}^{t-1}, \gamma_{x_i}^{t-1})}. \quad (46)$$

As a result,  $\hat{x}_{x_i \rightarrow f_j}^t$  given by (19) and (46) can be viewed as a function of  $\zeta_{x_i \rightarrow f_j}^{t-1}$  and  $\gamma_{x_i}^{t-1}$  as follows

$$\hat{x}_{x_i \rightarrow f_j}^t = \mathcal{F}^t(\zeta_{x_i \rightarrow f_j}^{t-1}, \gamma_{x_i}^{t-1}). \quad (47)$$

Let  $\Delta \zeta_{x_i \rightarrow f_j}^{t-1} = \zeta_{x_i \rightarrow f_j}^{t-1} - \zeta_{x_i}^{t-1}$  denote the difference between  $\zeta_{x_i \rightarrow f_j}^{t-1}$  and  $\zeta_{x_i}^{t-1}$

$$\Delta \zeta_{x_i \rightarrow f_j}^{t-1} = -\frac{\gamma_{x_i}^{t-1}}{\nu_{f_j}^{t-1}} h_{j,i}^* z_{f_j \rightarrow x_i}^{t-1}. \quad (48)$$

Then we can expand  $\hat{x}_{x_i \rightarrow f_j}^t$  at  $\zeta_{x_i}^{t-1}$  using the Taylor formula

$$\hat{x}_{x_i \rightarrow f_j}^t \approx \mathcal{F}^t(\zeta_{x_i}^{t-1}, \gamma_{x_i}^{t-1}) + \frac{\partial \mathcal{F}^t(\zeta_{x_i}^{t-1}, \gamma_{x_i}^{t-1})}{\partial \zeta_{x_i}^{t-1}} \Delta \zeta_{x_i \rightarrow f_j}^{t-1}, \quad (49)$$

where

$$\mathcal{F}^t(\zeta_{x_i}^{t-1}, \gamma_{x_i}^{t-1}) = \mathbb{E}_{\beta_{x_i}^t(x_i)}[x_i] \quad (50)$$

and

$$\frac{\partial \mathcal{F}^t(\zeta_{x_i}^{t-1}, \gamma_{x_i}^{t-1})}{\partial \zeta_{x_i}^{t-1}} = \frac{\mathbb{E}_{\beta_{x_i}^t(x_i)}[|x_i|^2] - |\mathbb{E}_{\beta_{x_i}^t(x_i)}[x_i]|^2}{\gamma_{x_i}^{t-1}}. \quad (51)$$

When  $\beta_{x_i}^t(x_i)$  in (50) and (51) is also replaced by  $\tilde{p}^t(x_i)$ , the first-order approximation of  $\hat{x}_{x_i \rightarrow f_j}^t$  shown in (49) is identical to (34).

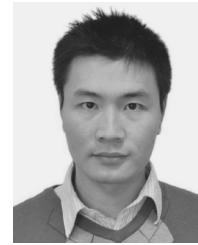
## REFERENCES

- [1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, 1998.
- [2] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecomm.*, vol. 10, no. 6, pp. 585–595, 1999.
- [3] G. L. Stüber, J. R. Barry, S. W. McLaughlin, Y. G. Li, M. A. Ingram, and T. Pratt, "Broadband MIMO-OFDM wireless communications," *Proc. IEEE*, vol. 92, no. 2, pp. 271–294, Feb. 2004.
- [4] S. Mohammed, A. Zaki, A. Chockalingam, and B. Rajan, "High-rate space-time coded large-MIMO systems: Low-complexity detection and channel estimation," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 6, pp. 958–974, Dec. 2009.
- [5] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [6] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [7] N. Srinidhi, T. Datta, A. Chockalingam, and B. Rajan, "Layered tabu search algorithm for large-MIMO detection and a lower bound on ML performance," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 2955–2963, Nov. 2011.
- [8] P. Som, T. Datta, N. Srinidhi, A. Chockalingam, and B. Rajan, "Low-complexity detection in large-dimension MIMO-ISI channels using graphical models," *IEEE J. Sel. Top. Signal Process.*, vol. 5, no. 8, pp. 1497–1511, Dec. 2011.
- [9] T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "A hybrid RTS-BP algorithm for improved detection of large-MIMO M-QAM signals," in *Proc. Nat. Conf. Commun. (NCC)*, 2011, pp. 1–5.
- [10] W. Feng, Y. Wang, N. Ge, J. Lu, and J. Zhang, "Virtual MIMO in multi-cell distributed antenna systems: Coordinated transmissions with large-scale CSIT," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 10, pp. 2067–2081, Oct. 2013.
- [11] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [12] C. Studer and E. G. Larsson, "PAR-aware large-scale multi-user MIMO-OFDM downlink," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 303–313, Feb. 2013.
- [13] S. Wu, L. Kuang, Z. Ni, J. Lu, D. Huang, and Q. Guo, "Expectation propagation based iterative groupwise detection for Large-Scale multiuser MIMO-OFDM systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Istanbul, Turkey, Apr. 2014.
- [14] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.
- [15] S. Bäro, J. Hagenauer, and M. Witzke, "Iterative detection of MIMO transmission using a list-sequential (LISS) detector," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2003, pp. 2653–2657.
- [16] Y. L. C. de Jong and T. J. Willink, "Iterative tree search detection for MIMO wireless systems," *IEEE Trans. Commun.*, vol. 53, no. 6, pp. 930–935, Jun. 2005.
- [17] C. Studer and H. Böleskei, "Soft-input soft-output single tree-search sphere decoding," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4827–4842, Oct. 2010.
- [18] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.
- [19] E. Larsson and J. Jaldén, "Fixed-complexity soft MIMO detection via partial marginalization," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3397–3407, Aug. 2008.
- [20] Q. Guo, L. Fang, D. D. Huang, and S. Nordholm, "A soft-in soft-out detection approach using partial Gaussian approximation," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, 2012.
- [21] X. Wang and H. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046–1061, Jul. 1999.
- [22] W.-J. Choi, K.-W. Cheong, and J. M. Cioffi, "Iterative soft interference cancellation for multiple antenna systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2000, vol. 1, pp. 304–309.
- [23] B. Lu, G. Yue, and X. Wang, "Performance analysis and design optimization of LDPC-coded MIMO-OFDM systems," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 348–361, Feb. 2004.
- [24] S. Ahmed, T. Ratnarajah, M. Sellathurai, and C. Cowan, "Iterative receivers for MIMO-OFDM and their convergence behavior," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 461–468, Jan. 2009.
- [25] C. Studer, S. Fateh, and D. Seethaler, "ASIC implementation of soft-input soft-output MIMO detection using MMSE parallel interference cancellation," *IEEE J. Solid-State Circuits*, vol. 46, no. 7, pp. 1754–1765, Jul. 2011.
- [26] M. Kaynak, T. Duman, and E. Kurtas, "Belief propagation over MIMO frequency selective fading channels," in *Proc. Joint Int. Conf. Auton. Autonom. Syst. and Int. Conf. Netw. Services (ICAS-ICNS)*, Oct. 2005, p. 45.
- [27] M. N. Kaynak, T. M. Duman, and E. M. Kurtas, "Belief propagation over SISO/MIMO frequency selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2001–2005, Jun. 2007.
- [28] X. Yang, Y. Xiong, and F. Wang, "An adaptive MIMO system based on unified belief propagation detection," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2007, pp. 4156–4161.

- [29] J. Hu and T. M. Duman, "Graph-based detection algorithms for layered space-time architectures," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 269–280, 2008.
- [30] P. Som, T. Datta, A. Chockalingam, and B. S. Rajan, "Improved large-MIMO detection based on damped belief propagation," in *Proc. IEEE Inf. Theory Workshop (ITW)*, 2010, pp. 1–5.
- [31] T. L. Narasimhan, A. Chockalingam, and B. S. Rajan, "Factor graph based joint detection/decoding for LDPC coded large-MIMO systems," in *Proc. IEEE Veh. Tech. Conf. (VTC-Spring)*, 2012, pp. 1–5.
- [32] J. Goldberger and A. Leshem, "MIMO detection for high-order QAM based on a Gaussian tree approximation," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4973–4982, Aug. 2011.
- [33] G. Colavolpe, D. Fertonani, and A. Piemontese, "SISO detection over linear channels with linear complexity in the number of interferers," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 8, pp. 1475–1485, Dec. 2011.
- [34] Q. Guo and D. D. Huang, "EM-based joint channel estimation and detection for frequency selective channels using gaussian message passing," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 4030–4035, Aug. 2011.
- [35] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci.*, vol. 106, no. 45, pp. 18 914–18 919, 2009.
- [36] D. L. Donoho, A. Maleki, and A. Montanari, IEEE, "Message passing algorithms for compressed sensing: I. motivation and construction," in *Proc. IEEE Inf. Theory Workshop (ITW)*, 2010, pp. 1–5.
- [37] A. Montanari, "Graphical models concepts in compressed sensing," *Compressed Sensing: Theory Applicat.*, pp. 394–438, 2012.
- [38] S. Rangan, Generalized Approximate Message Passing for Estimation With Random Linear Mixing, arXiv:1010.5141v2 [cs.IT], 2010.
- [39] Q. Guo, D. D. Huang, S. Nordholm, J. Xi, and Y. Yu, "Iterative frequency domain equalization with generalized approximate message passing," *IEEE Signal Process. Lett.*, vol. 20, no. 6, pp. 559–562, Jun. 2013.
- [40] P. Schniter, "A message-passing receiver for BICM-OFDM over unknown clustered-sparse channels," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 8, pp. 1462–1474, Dec. 2011.
- [41] M. Tüchler and A. Singer, "Turbo equalization: An overview," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 920–952, Feb. 2011.
- [42] P. S. Rossi and R. R. Müller, "Joint twofold-iterative channel estimation and multiuser detection for MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4719–4729, Nov. 2008.
- [43] G. E. Kirkelund, C. N. Manchón, L. P. B. Christensen, E. Riegler, and B. H. Fleury, "Variational message-passing for joint channel estimation and decoding in MIMO-OFDM," in *Proc. IEEE Global Telecomm. Conf. (GLOBECOM)*, 2010, pp. 1–6.
- [44] C. N. Manchón, G. E. Kirkelund, E. Riegler, L. P. B. Christensen, and B. H. Fleury, "Receiver architectures for MIMO-OFDM based on a combined VMP-SP algorithm," arXiv: 1111.5848, 2011.
- [45] C. Knievel, P. A. Hoher, A. Tyrrell, and G. Auer, "Multi-dimensional graph-based soft iterative receiver for MIMO-OFDM," *IEEE Trans. Commun.*, vol. 60, no. 6, pp. 1599–1609, Jun. 2012.
- [46] C. Novak, G. Matz, and F. Hlawatsch, "IDMA for the multiuser MIMO-OFDM uplink: A factor graph framework for joint data detection and channel estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4051–4066, Aug. 2013.
- [47] T. P. Minka, "Divergence measures and message passing," Microsoft Research Cambridge, 2005, Tech. Rep..
- [48] S. Wu, L. Kuang, Z. Ni, J. Lu, D. Huang, and Q. Guo, "Expectation propagation approach to joint channel estimation and decoding for OFDM systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Florence, Italy, May 2014.
- [49] M. Tüchler, A. Singer, and R. Koetter, "Minimum mean squared error equalization using *a priori* information," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 673–683, Mar. 2002.
- [50] Q. Guo and D. D. Huang, "A concise representation for the soft-in soft-out LMMSE detector," *IEEE Commun. Lett.*, vol. 15, no. 5, pp. 566–568, May 2011.
- [51] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.
- [52] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004.
- [53] T. P. Minka, "A family of algorithms for approximate Bayesian inference," Ph.D. dissertation, Mass. Inst. of Technol., Cambridge, MA, USA, 2001.
- [54] C. M. Bishop *et al.*, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.
- [55] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, MA, USA: MIT Press, 2009.
- [56] C. Heegard and S. B. Wicker, *Turbo Coding*. Boston, MA, USA: Kluwer, 1999.

**Sheng Wu** (S'13) received the B.E. and M.E. degrees from Beijing University of Post and Telecommunications, Beijing, China, in 2004 and 2007, respectively.

He is currently working toward the Ph.D. degree in electrical engineering at Tsinghua University, Beijing, China. His research interests are mainly in iterative detection and decoding, channel estimation, MIMO, and receiver design.



**Linling Kuang** (S'01–M'06) received the B.S. and M.S. degrees from the National University of Defense Technology, Changsha, China, in 1995 and 1998, respectively, and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2004.

Since 2007, she has been with Tsinghua University, where she is currently an Associate Professor in the Tsinghua Space Center. Her research interests include wireless broadband communications, signal processing, and satellite communication. Dr. Kuang

is a member of the IEEE Communications Society.



**Zuyao Ni** received the B.E. and M.E. degrees from Zhejiang University, Hangzhou, China, in 1998 and 2001, respectively, and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2006.

Since 2007, he has been with Tsinghua University, where he is currently an Associate Professor in the Research Institute of Information Technology. His research interests include wireless broadband communications, signal processing, and satellite communication.



**Jianhua Lu** (M'98–SM'07) received the B.S.E.E. and M.S.E.E. degrees from Tsinghua University, Beijing, China, in 1986 and 1989, respectively, and his Ph.D. in Electrical & Electronic Engineering from the Hong Kong University of Science & Technology.

Since 1989, he has been with the Dept. of Electronic Engineering, Tsinghua Univ. Beijing, China, where he now serves as a Professor. His current research interests include broadband wireless communication, multimedia signal processing, and wireless networking.

Dr. Lu has been an active member of professional societies and has published more than 180 technical papers in international journals and conference proceedings. He was one of the recipients of best paper awards at the IEEE International Conference on Communications, Circuits and Systems 2002, ChinaCom 2006, and IEEE Embedded-Com 2012, and was awarded the National Distinguished Young Scholar Fund by the NSF committee of China in 2005. He has served in numerous IEEE conferences as a member of Technical Program Committees and served as Lead Chair of the General Symposium of IEEE ICC 2008, as well as a Program Committee Co-Chair of the 9th IEEE International Conference on Cognitive Informatics in 2010. He is now a chief scientist of the National Basic Research Program (973), China. Dr. Lu is a Senior Member of the IEEE Communication Society and the IEEE Signal Processing Society.



**Defeng (David) Huang** (M'01–S'02–M'05–SM'07) received the B. E. E. E. and M. E. E. E. degree in electronic engineering from Tsinghua University, Beijing, China, in 1996 and 1999, respectively, and the Ph.D. degree in electrical and electronic engineering from the Hong Kong University of Science and Technology (HKUST), Kowloon, Hong Kong, in 2004.

Currently, he is a professor with the School of Electrical, Electronic and Computer Engineering at the University of Western Australia. Dr. Huang serves as an Editor for the IEEE WIRELESS COMMUNICATIONS LETTERS, he also served as an Editor (2005–2011) for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.



**Qinghua Guo** (S'07–M'08) received the B.E. degree in electronic engineering and the M.E. degree in signal and information processing from Xidian University, Xi'an China, in 2001 and 2004, respectively, and the Ph.D. degree in electronic engineering from City University of Hong Kong, Kowloon, Hong Kong SAR, in 2008.

He is currently with the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW, Australia, and also with the School of Electrical, Electronic, and Computer Engineering, University of Western Australia, Perth, WA, Australia. His research interests include statistical signal processing and wireless communications. Dr Guo is a recipient of the Australian Research Council's inaugural DECRA.