2014 IEEE 17th International Conference on
Intelligent Transportation Systems (ITSC)
October 8-11, 2014. Qingdao, China

Non-Gaussian Approximation Iterative Receiver for MIMO-OFDM Vehicular Networks

Shuangshuang Han[1], Liuqing Yang[2], Xiang Cheng[3], Shichao Chen[2],
Fenghua Zhu[2], Zhongdong Yu[4] and Haitao Xie[4]

Abstract— **Multiple-Input Multiple-Output - Orthogonal Frequency Division Multiplexing (MIMO-OFDM) is adopted to vehicular networks to increase the capacity, reliability and speed. In this paper, iterative demodulation and decoding algorithms are studied to approach the capacity of MIMO-OFDM vehicular networks. By analysing the drawbacks of the Gaussian approximation on the interference cancellation, Non-Gaussian approximation is proposed to enhance the performance of interference cancellation based detectors with large constellations. Simulation results demonstrate that the proposed non-Gaussian algorithm can achieve a significant performance gain over existing ones with high order constellations.**

I. INTRODUCTION

To make transportation safer, more efficient and less harmful, vehicle-to-infrastructure and vehicle-to-vehicle communications are currently being intensely investigated and developed [1]–[8]. Key characteristics of vehicular channels are shadowing, high Doppler shifts, and inherent nonstationarity. An international standard, IEEE 802.11p [9], which has gained considerable importance. For the improvements in coverage, reliability, scalability and delay, multiple antennas techniques are recommended to exploit spatial diversity for increased diversity and reliability.

To meet ambitious target data rates as 1 Gb/s in local areas and 100 Mb/s in wide areas, MIMO technique [10], [11]is expected to be deployed which promise a linear increase of the wireless link capacity. Furthermore, OFDM has been selected for the downlink due to its robustness to the multipath as well as its flexibility in the resource allocation. To exploit the potentials of MIMO-OFDM, one of the challenges is the complexity of decoding algorithm at the receiver end.

For uncoded MIMO systems, a low complexity detection scheme is proposed in [12], which is a nulling and cancelling scheme. In [13], a technique referred to as "sphere decoding"(SD) is proposed for lattice code decoding. This SD approach is extended to coded MIMO systems in [14], which iteratively detect and decode any linear space-time mapper combined with an outer channel code. Computing the exact log likelihood ratio (LLR) in [14] requires a complexity exponential in the number of antennas and in the size of the constellation. A list SD and max-log approximation are used to approach the optimal performance with low complexity. However, the complexity of SD based algorithms depends on signal-to-noise (SNR). Different approach [15]–[20] to adopt soft information in iterative detection and decoding is using non-linear interference cancellation followed by a single input and single output (SISO) iterative detector and decoder in [21], which has a low complexity.

In this paper, we efficient iterative detection and decoding for MIMO-OFDM systems. We follow the iterative architecture in [14], where the detector and the decoder exchange information. Different algorithms such as those in [14]–[20] differ in the way how the soft information is generated from the detector. We first review the algorithm using Gaussian approximation on the interference in the LLR value computation. However, the Gaussian approximation does not work well for high order modulations such as 64 quadrature amplitude modulation (QAM). We propose a class of non-Gaussian approximation for any constellation by writing the probability mass function (pmf) of the constellation points into a unified form and relaxing the variable to take real numbers rather than a discrete set of numbers. By replacing the Gaussian distribution with the non-Gaussian one and integrating over the resulting continuous probability density function (pdf), we obtain closed-form of LLR value. In the proposed algorithm, the *a priori* probability is used to update the list at each stage, where the *a priori* probability is approximated using Gaussian and non-Gaussian approximations. Simulation results show that the proposed algorithms achieve a superior performance over existing ones particularly for high order modulation.

The rest of the paper is organized as follows. In Section II, the MIMO-OFDM system model is presented. In Section III, the iterative receiver structure is outlined. A non-Gaussian approximation based algorithm is given in Section IV. Simulation results are given in V and Section VI contains the conclusions.

II. SYSTEM MODEL

We consider a MIMO OFDM system with $\tilde{M}$ transmit and N receive antennas. The system has N_s subcarriers in an

[1]Shuangshuang Han is with Qingdao Academy of Intelligent Industries, Qingdao, Shandong Province, China. shuangshuang.han@qaii.ac.cn

[2]Liuqing Yang, Shichao Chen and Fenghua Zhu are with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation Chinese Academy of Sciences, Beijing, China. {liuqing.yang, shichao.chen, fenghua.zhu}@ia.ac.cn

[3]Xiang Cheng is with School of Electronics Engineering and Computer Science, Peking University, Beijing, China. xiangcheng@pku.edu.ca

[4]Zhongdong Yu and Haitao Xie are with Guang Zhou Communication Information Construction Investment and Operation Co. Ltd. 510405, Guangzhou, China. {yuzhd,xiehaitao}@96900.com.cn

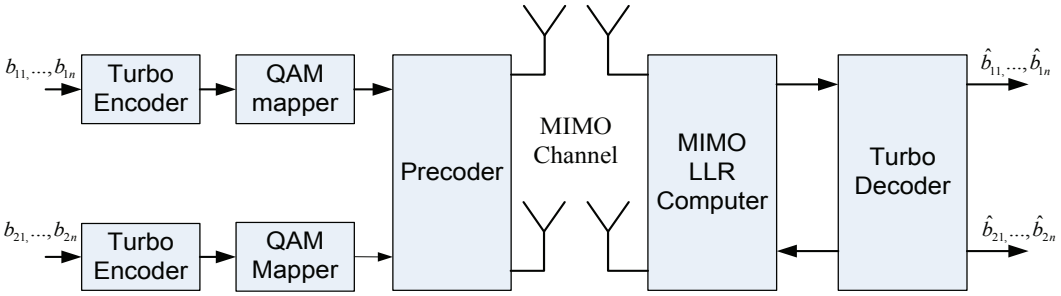

Fig. 1. The diagram of a MIMO-OFDM system.

OFDM block. There are M data streams to be transmitted[1]. The constellation $\mathcal{Q}_m$ is applied on stream m, where C_m is the number of bits per constellation symbol. The incoming bits of each stream m of length $N_s C_m R_m, m = 1, \ldots, M$, is encoded using a channel code (typically a convolutional or turbo code) of rate R_m, resulting in a bit vector $\mathbf{b}_m$. The encoded bits are converted into symbols using a mapping function $x_{i,m} = \mathcal{M}_m(\mathbf{b}_m((i-1)C_m + 1 : iC_m))$ (e.g., Gray mapping and set partitioning mapping), $i = 0, ..., N_s - 1$, where $x_{i,m}$ is the symbol to be transmitted over subcarrier i and antenna m and we have used Matlab notation. The Inverse Discrete Fourier Transform (IDFT) of the data block $x_{0,m}, \ldots, x_{N_s-1,m}$ yields the time domain sequence, i.e.,

$$X_{j,m} = \frac{1}{\sqrt{N_s}} \sum_{i=0}^{N_s-1} x_{i,m} e^{j2\pi ij/N_s}, j = 0, \ldots, N_s - 1. \quad (1)$$

The time domain symbol X_j, m is assumed to obey the component-wise energy constraint $E\|X_{j,m}\|^2 = \mathcal{E}_s/M$. A cyclic prefix (CP) is added to mitigate for the residual ISI due to previous OFDM symbol. After parallel-to-serial (P/S) conversion, signal is transmitted from the corresponding antenna. The channel between each transmitter/receiver pair is modelled as multipath channel. The channel between transmit antenna m and receive antenna n is expressed as

$$h_{n,m}(t) = \sum_{l=0}^{\Gamma_{n,m}-1} \alpha_{n,m,l} \delta(t - \tau_{n,m,l}), \quad (2)$$

where $\Gamma_{n,m}$ is the number of taps, $\alpha_{n,m,l}$ is the l-th complex path gain, and $\tau_{n,m,l}$ is the corresponding path delay. We assume block fading model in this paper, where the channel is assumed to be constant in each OFDM data block.

At the receiver side, serial-to-parallel (S/P) conversion is first performed and the CP is removed. After DFT operation, the received signal in frequency domain can be expressed as

$$y_{i,n} = \sum_{m=1}^{M} H_{i,n,m} x_{i,m} + w_{i,n}, \quad (3)$$

where $i = 0, \ldots, N_s - 1, n = 1, \ldots, N$, n denotes the receiver antenna indexing, $w_{i,n}$ is the additive white Gaussian noise (AWGN) with zero mean and variance σ^2, and

$$H_{i,n,m} = \frac{1}{\sqrt{N_s}} \sum_{l=0}^{\Gamma_{n,m}-1} \alpha_{n,m,l} e^{-j2\pi \lceil \tau_{n,m,l/T_s} \rfloor i/N_s}, \quad (4)$$

[1]We have $M \leq \tilde{M}$ due to possible beamforming at the transmitter. In this case, we consider an equivalent channel with M transmit antennas.

where T_s is the symbol duration. We can write Eq. (3) in vector form as

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{x}_i + \mathbf{w}_i, i = 0, \ldots, N_s - 1. \quad (5)$$

The diagram of a MIMO-OFDM system is given in Fig. 1. We can consider Eq. (5) as a MIMO system on each subcarrier. In the following parts of this paper, we will neglect the subscript i in Eq. (5).

III. ITERATIVE DETECTION AND DECODING ALGORITHMS

The channel code and the MIMO channel can be considered as a serially concatenated scheme [14] [21]with an outer channel encoder and inner constellation mapping with block encoding matrix $\mathbf{H}_i$ at each subcarrier. To decode $\mathbf{b}_1, \ldots, \mathbf{b}_M$, the optimal joint detector and decoder should compute the likelihood of each bit given all the received signals $\mathbf{y}_1, \ldots, \mathbf{y}_{N_s-1}$ on all subcarriers. However, this is computationally infeasible in practice. Several algorithms [14], [17]–[19], [21] solve this problem approximately using the "turbo principle", where information is exchanged between the detector and decoder in an iterative fashion. In this section, we focus on how to generate extrinsic information at each subcarrier using the received signals on this subcarrier using the a priori information on each bits from the channel decoder. The generated extrinsic information on all subcarrier is then put into the soft in and soft out channel decoder (e.g., Bahl-Cocke-Jelinek-Raviv (BJCR) algorithm) for the next iteration decoding and detection. Different joint detection and decoding algorithms share the same outer channel decoder. Their difference lies in how the extrinsic information from the inner mapping. The diagram of iterative decoding and demodulation for MIMO-OFDM is given in Fig. 2.

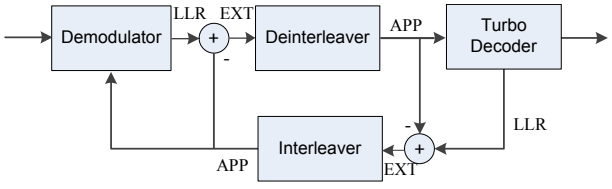

Fig. 2. The diagram of iterative decoding and demodulation for MIMO-OFDM.

A. iterative MAP Detection and Decoding

The a priori probability (APP) is usually expressed as a log-likelihood ratio (LLR) value. In the rest of this paper, the logical zero for a bit is represented by amplitude level

$b_i = -1$ and logical one by $b_i = +1$, respectively. After obtaining the APP from the channel decoder (initially the APP is set to be zero), the a posteriori LLR value of the bit $b_i, i = 0, ..., \sum_{m=1}^{M} C_m - 1$ conditioned on the received vector $\mathbf{y}$ is

$$L(b_i|\mathbf{y}) = \log \frac{\Pr(b_i = +1|\mathbf{y})}{\Pr(b_i = -1|\mathbf{y})}. \quad (6)$$

If we assume that the interleaver at the encoder is ideal such that the bits in each modulation symbol are approximately statistically independent of one another, we can rewrite Eq. (6) using Bayes' theorem as Eq. (7) on next page, where $\mathcal{X}_{i,+1}$ and $\mathcal{X}_{i,-1}$ are the set of $2^{\sum_{m=1}^{M} C_m - 1}$ symbols vectors such that the i-th bit is $+1$ or -1, respectively, i.e., $\mathcal{X}_{i,\pm 1} = \{\mathbf{x}|\mathcal{M}(\mathbf{b}) = \mathbf{x}, b_i = \pm 1\}, \mathbf{b} = \mathcal{B}(\mathbf{x})$ is the inverse mapping of $\mathbf{x} = \mathcal{M}(\mathbf{b})$ and $\mathcal{B}_j(\mathbf{x})$ is the j-th bit of $\mathcal{B}(\mathbf{x})$.

In the case of Gaussian channel as in Eq. (5), we can further write $L(b_i|\mathbf{y})$ as Eq. (8) on next page. In [14], instead of using Eq. (8) directly, max-log approximation is adopted to compute $L_E(b_i|\mathbf{y})$ as Eq. (9). However, the simplification Eq. (9) still has a complexity exponential in the total number of bits or $\sum_{m=1}^{M} C_m$. In [14], a list sphere decoder (LSD) is used to resolve this issue by searching only over a list $\mathcal{L}$ containing N_{cand} elements, i.e., Eq. (10). The list is generated by checking only points within the hypersphere of radius r, i.e., $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \leq r^2$, where r is chosen according to the noise variance such that the number of points within the hypersphere is not far away from $\mathbf{y}$. The performance of the LSD base algorithm depends on the size of the list. Clearly, when the list size is equal to the number of all possible constellation points, i.e., $2^{\sum_{m=1}^{M} C_m}$, Eq. (10) reduces to (9).

B. Iterative MAP Detection and Decoding with Gaussian Approximation

The complexity of computing the LLR value from Eq. (7) is high. Note that we can rewrite (7) as Eq. (11), where x_m denotes the symbol that b_i belongs to, i.e., $\sum_{m'=1}^{m-1} C_{m'} \leq i < \sum_{m'=1}^{m} C_{m'}$, $\mathbf{x}_{-m}$ denotes the vectors contains all entries of $\mathbf{x}$ except the m-th entry, and and $\mathcal{X}_{i,+1}^{m}$ and $\mathcal{X}_{i,-1}^{m}$ are the set of $2^{C_m - 1}$ symbols such that b_i is $+1$ or -1, respectively. From Eq. (11), we need to compute $\sum_{\mathbf{x}_{-m}} \Pr(\mathbf{x}_{-m}, x_m) \Pr(\mathbf{x}_{-m})$ for any given x_m. A suboptimal approach is to replace the summation over $\mathbf{x}_{-m}$ with an integration over a continuous distribution. One typical assumption is to use the Gaussian distribution. We assume the entries of x_{-m} are independent Gaussian random variables with mean

$$\mu_{m'} = E\{x_{m'}\} = \sum_{x_{m'}} \Pr(x_{m'}) x_{m'} \quad (12)$$

and variance

$$\begin{aligned} \nu_{m'}^2 &= E\{|x_{m'}|^2\} - E^2\{x_{m'}\} \\ &= \sum_{x_{m'} \in \mathcal{C}_{m'}} \Pr(x_{m'}) |x_{m'}|^2 - |\mu_{m'}|^2, \end{aligned} \quad (13)$$

$m' = 1, \ldots, M, m' \neq m$. When Gaussian channel model Eq. (5) is used, we have

$$\begin{aligned} &\Pr(\mathbf{y}|x_m) = \sum_{\mathbf{x}_{-m}} \Pr(\mathbf{y}|\mathbf{x}_{-m}, x_m) \Pr(\mathbf{x}_{-m}) \\ &\approx \int_{-\infty}^{+\infty} \Pr(\mathbf{y}|\mathbf{x}_{-m}, x_m) f(\mathbf{x}_{-m}) d\mathbf{x}_{-m} \\ &= \int_{-\infty}^{+\infty} \frac{1}{(\pi\sigma^2)^N} \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}_{-m}\mathbf{x}_{-m} - \mathbf{h}_m x_m\|^2}{\sigma^2} \right) \times \\ &\qquad \frac{1}{\pi^N \prod_{m'=1, m' \neq m}^{M} \nu_{m'}^2} \times \\ &\qquad \exp\left(- \sum_{m'=1, m' \neq m}^{M} \frac{|x_{m'} - \mu_{m'}|^2}{\nu_{m'}^2} \right) d\mathbf{x}_{-m} \\ &\propto \left(-(\mathbf{y} - \mathbf{H}_{-m}\mu_{-m} - \mathbf{h}_m x_m)^H \mathbf{R}_m^{-1} \right. \\ &\qquad \left. (\mathbf{y} - \mathbf{H}_{-m}\mu_{-m} - \mathbf{h}_m x_m) \right), \end{aligned} \quad (14)$$

where the integral is from $-\infty$ to ∞ in each dimension, $\mathbf{H}_{-m}$ contains the columns of $\mathbf{H}$ except the m-th column, $\mathbf{h}_m$ is the m-th column of $\mathbf{H}$, $\mu_{-m} = [\mu_1, \ldots, \mu_{m-1}, \mu_{m+1}, \ldots, \mu_M]$,

$$\begin{aligned} \mathbf{R}_m =& \mathbf{H}_{-m} \text{diag}\{\nu_1^2, \ldots, \nu_{m-1}^2, \nu_{m+1}^2, \ldots, \nu_M^2\} \mathbf{H}_{-m}^H \\ &+ \sigma^2 \mathbf{I}_N, \end{aligned} \quad (15)$$

and $\mathbf{I}_N$ is an N by N identity matrix. Substituting (14) into (11), we obtain the LLR value under Gaussian approximation. The complexity of computing LLR reduces from $2^{\sum_{m=1}^{M} C_m}$ to 2^{C_m}.

IV. PROPOSED ITERATIVE DETECTION AND DECODING ALGOTITHMS

A. Motivation

There are several issues with the existing algorithms.

- Many practical wireless communications standards now adopt high order constellations such as 64QAM. The max-log approximation in (9) may not work well with high order constellations as the number of terms in the summation in (8) is large in this case. Moreover, the LSD may be hard to implement in hardware directly due to its sequential nature.
- The Gaussian approximation based algorithms avoid the max-log approximation but the Gaussian assumption incurs some performance loss. It is commented in [19] that the performance of Gaussian approximation algorithms is not good for higher order modulations.

We wonder whether we could combine these two strategies and take the advantage of both. The main contribution in this section is summarized as follows: We propose an non-Gaussian approximations for LLR computation. As practical constellations have a finite alphabet structure, the non-Gaussian distribution is integrated over a bounded set instead of from $-\infty$ to $+\infty$.

In this section, we assume squared-QAM is used at all transmit antennas, which is the case in many wireless

$$L(b_i|\mathbf{y}) = \log \frac{\sum_{\mathbf{x}\in\mathcal{X}_{i,+1}} \Pr(\mathbf{x}|\mathbf{y})}{\sum_{\mathbf{x}\in\mathcal{X}_{i,-1}} \Pr(\mathbf{x}|\mathbf{y})} = \underbrace{\log \frac{\sum_{\mathbf{x}\in\mathcal{X}_{i,+1}} \Pr(\mathbf{y}|\mathbf{x}) \prod_{j=0,j\neq i}^{\sum_{m=1}^{M} C_m - 1} \Pr(b_i = \mathcal{B}_j(\mathbf{x}))}{\sum_{\mathbf{x}\in\mathcal{X}_{i,-1}} \Pr(\mathbf{y}|\mathbf{x}) \prod_{j=0,j\neq i}^{\sum_{m=1}^{M} C_m - 1} \Pr(b_i = \mathcal{B}_j(\mathbf{x}))}}_{L_E(b_i|\mathbf{y})} + \underbrace{\log \frac{\Pr(b_i=+1)}{\Pr(b_i=-1)}}_{L_A(b_i)} \tag{7}$$

$$\begin{aligned} L(b_i|\mathbf{y}) &= \log \frac{\sum_{\mathbf{x}\in\mathcal{X}_{i,+1}} \exp\left(-\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{\sigma^2}\right) \prod_{j=0,j\neq i}^{\sum_{m=1}^{M} C_m - 1} \Pr(b_i = \mathcal{B}_j(\mathbf{x}))}{\sum_{\mathbf{x}\in\mathcal{X}_{i,-1}} \exp\left(-\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{\sigma^2}\right) \prod_{j=0,j\neq i}^{\sum_{m=1}^{M} C_m - 1} \Pr(b_i = \mathcal{B}_j(\mathbf{x}))} + L_A(b_i) \\ &= \log \frac{\sum_{\mathbf{x}\in\mathcal{X}_{i,+1}} \exp\left(-\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{\sigma^2} + \sum_{j=0,j\neq i,b_i=1}^{\sum_{m=1}^{M} C_m - 1} L_A(b_j)\right)}{\sum_{\mathbf{x}\in\mathcal{X}_{i,-1}} \exp\left(-\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{\sigma^2} + \sum_{j=0,j\neq i,b_i=-1}^{\sum_{m=1}^{M} C_m - 1} L_A(b_j)\right)} + L_A(b_i) \end{aligned} \tag{8}$$

$$L_E(b_i|\mathbf{y}) \approx \max_{\mathbf{x}\in\mathcal{X}_{i,+1}} \left\{ -\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{\sigma^2} + \sum_{j=0,j\neq i,b_i=1}^{\sum_{m=1}^{M} C_m - 1} L_A(b_j) \right\} - \max_{\mathbf{x}\in\mathcal{X}_{i,-1}} \left\{ -\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{\sigma^2} + \sum_{j=0,j\neq i,b_i=1}^{\sum_{m=1}^{M} C_m - 1} L_A(b_j) \right\} \tag{9}$$

$$L_E(b_i|\mathbf{y}) \approx \max_{\mathbf{x}\in\mathcal{L}\cap\mathcal{X}_{i,+1}} \left\{ -\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{\sigma^2} + \sum_{j=0,j\neq i,b_i=1}^{\sum_{m=1}^{M} C_m - 1} L_A(b_j) \right\} - \max_{\mathbf{x}\in\mathcal{L}\cap\mathcal{X}_{i,-1}} \left\{ -\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{\sigma^2} + \sum_{j=0,j\neq i,b_i=1}^{\sum_{m=1}^{M} C_m - 1} L_A(b_j) \right\} \tag{10}$$

$$L_E(b_i|\mathbf{y}) = \log \frac{\sum_{\mathbf{x}\in\mathcal{X}_{i,+1}} \Pr(\mathbf{y}|\mathbf{x}) \Pr(\mathbf{x})}{\sum_{\mathbf{x}\in\mathcal{X}_{i,-1}} \Pr(\mathbf{y}|\mathbf{x}) \Pr(\mathbf{x})} = \log \frac{\sum_{x_m\in\mathcal{X}_{i,+1}^m} \Pr(x_m) \sum_{\mathbf{x}_{-m}} \Pr(\mathbf{y}|\mathbf{x}_{-m}, x_m) \Pr(\mathbf{x}_{-m})}{\sum_{x_m\in\mathcal{X}_{i,-1}^m} \Pr(x_m) \sum_{\mathbf{x}_{-m}} \Pr(\mathbf{y}|\mathbf{x}_{-m}, x_m) \Pr(\mathbf{x}_{-m})}, \tag{11}$$

communications standards. But the proposed algorithm can be readily extended to other general constellations. With squared-QAM, we can write (5) as a real system, i.e.,

$$\begin{aligned} \underbrace{\begin{bmatrix} \Re(\mathbf{y}_i) \\ \Im(\mathbf{y}_i) \end{bmatrix}}_{\tilde{\mathbf{y}}_i} &= \underbrace{\begin{bmatrix} \Re(\mathbf{H}_i) & -\Im(\mathbf{H}_i) \\ \Im(\mathbf{H}_i) & \Re(\mathbf{H}_i) \end{bmatrix}}_{\tilde{\mathbf{H}}_i} \underbrace{\begin{bmatrix} \Re(\mathbf{x}_i) \\ \Im(\mathbf{x}_i) \end{bmatrix}}_{\tilde{\mathbf{x}}_i} \\ &+ \underbrace{\begin{bmatrix} \Re(\mathbf{w}_i) \\ \Im(\mathbf{w}_i) \end{bmatrix}}_{\tilde{\mathbf{w}}_i}, i = 0, , \ldots, N_s - 1, \end{aligned} \tag{16}$$

where $\Re(x)$ and $\Im(x)$ denote the real part and imaginary part of x, respectively and the entries of $\tilde{\mathbf{x}}_i$ are from PAM constellations. With a slight abuse of notations, we still use (5) to represent the real system (16) in this section with the entries of $\mathbf{x}_i$ from PAM.

B. Iterative MAP Detection and Decoding with Non-Gaussian Approximation

To motivate our non-Gaussian approximation, we start with the BPSK, i.e., $X \in \{+1, -1\}$. Let $\Pr(X = +1) = p$ and $\Pr(X = -1) = 1 - p$. We can write this probability mass function (pmf) into a single equation as

$$\Pr(X = x) = p^{\left(\frac{x+1}{2}\right)^2} (1-p)^{\left(\frac{x-1}{2}\right)^2}, \quad x = \pm 1. \tag{17}$$

A nature continuous approximation to this pmf is by relaxing x to be a real number with a scaling factor to keep $\int \Pr(X = x)dx = 1$. Note that there are several choices of the pmf (17). For example, we can choose $\Pr(X = x) = p^{\frac{x+1}{2}}(1-p)^{\frac{1-x}{2}}$. But this function will go to ∞ when x goes to ∞, which is undesired. We can also choose $\Pr(X = x) = p^{\frac{|x+1|}{2}}(1-p)^{\frac{|x-1|}{2}}$. But this function is hard to obtain closed form of integration.

The idea can be extended to higher modulation. For a given modulation $\mathcal{Q}$ with $\Pr(X = x_i) = p_i$ and $\sum p_i = 1$, we can write the pmf into a single equation as

$$\begin{aligned} \Pr(X = x) &= \prod_{x_i\in\mathcal{Q}} p_i^{\frac{\prod_{x_j\in\mathcal{Q}, x_j\neq x_i} (x-x_j)^2}{\prod_{x_j\in\mathcal{Q}, x_j\neq x_i} (x_i-x_j)^2}}, \quad x \in \mathcal{Q} \\ &= \exp\left(\sum_{l=0}^{2(|Q|-1)} a_l x^l\right) \end{aligned} \tag{18}$$

The pdf can be obtained by relaxing x to be a real number. When $|\mathcal{Q}| > 2$, if we use (18) directly in (14), the integral involves a polynomial greater than second order in the exponential function, whose closed form may be hard to obtain. Therefore, we approximate the pmf (18) with a second order polynomial in the exponential function for any $\mathcal{Q}$, i.e.,

$$\Pr(X = x) = \exp(-(c + 2rx + ax^2)). \tag{19}$$

Note that the Gaussian distribution is a special case of (19), which contains only two variables. The coefficients a, r, c are found by solving

$$\min_{a,r,c} \sum_i \omega_i (\exp(-(c + 2rx_i + ax_i^2)) - p_i)^2, \tag{20}$$

or

$$\min_{a,r,c} \sum_i \omega_i (c + 2rx_i + ax_i^2 + \log(p_i))^2, \qquad (21)$$

where $\omega_i \geq 0$ is a weight for symbol x_i. In practical systems, we may only care about the symbols with the largest probability. In this case, we can choose $\omega_i = 1$ for the three largest probability symbols and $\omega_i = 0$ otherwise. The solution of (21) can be readily obtained by least squares.

Except that the Gaussian approximation is not good for some pmf, the integration in (14) is from $-\infty$ to $+\infty$, which may distort the LLR value. Note that practical constellations typically are usually finite alphabets, e.g., 2D-PAM is $\{-2D+1, -2D+3, \ldots, 2D-3, 2D-1\}$. We can integrate from $-U$ to U instead. Some possible choices of U are $2D$ or $2D-1+\sigma$. When $U = 2D$, we approximate $\Pr(X = d)$ by the integral between $d-1$ and $d+1$. When $U = 2D-1+\sigma$, $\Pr(X = d)$ is approximated similarly as when $U = 2D$ but it takes into account of the noise variance at the two boundary points. With (19) and the finite integration, we can write (14) as

$$\begin{aligned}\Pr(\mathbf{y}|x_m) &\propto \int_{-U}^{+U} \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}_{-m}\mathbf{x}_{-m} - \mathbf{h}_m x_m\|^2}{\sigma^2}\right. \\ &\quad \left. -2\mathbf{r}_{-m}^T\mathbf{x}_{-m} - \mathbf{x}_{-m}^T\mathbf{A}_{-m}\mathbf{x}_{-m}\right) d\mathbf{x}_{-m} \\ &\propto \exp\left(-\frac{\|\mathbf{y} - \mathbf{h}_m x_m\|^2}{\sigma^2}\right) \\ &\int_{-U}^{+U} \exp\left(-2\underbrace{\left(\mathbf{r}_{-m}^T - \frac{1}{\sigma^2}(\mathbf{y} - \mathbf{h}_m x_m)^T\mathbf{H}_{-m}\right)}_{\mathbf{b}_{-m}^T}\mathbf{x}_{-m}\right. \\ &\quad \left. -\mathbf{x}_{-m}^T\left(\mathbf{A}_{-m} + \frac{\mathbf{H}_{-m}^T\mathbf{H}_{-m}}{\sigma^2}\right)\mathbf{x}_{-m}\right) d\mathbf{x}_{-m}\end{aligned} \qquad (22)$$

where $\mathbf{r}_{-m} = [r_1, \ldots, r_{m-1}, r_{m+1}, \ldots, r_M]^T$ and $\mathbf{A}_{-m} = \text{diag}\{a_1, \ldots, a_{m-1}, a_{m+1}, \ldots, a_M\}$, $r_{m'}$ and $a_{m'}$ are obtained from (20) and (21). Comparing with (14), we can see that there are two main differences. First, $\mathbf{r}_{-m}$ and $\mathbf{A}_{-m}$ are not from the matched mean and variance but from matching the pmf directly. Second, the integral is from $-U$ to U.

To compute the integral in (22), let the singular value decomposition of $\mathbf{R}_m$ be $\mathbf{V}^T\Lambda\mathbf{V}$ and $\mathbf{g}(x_m) = \mathbf{V}\mathbf{b}_{-m}$, where $\Lambda = \text{diag}\{\lambda_1, \ldots, \lambda_{M-1}\}$. We make a change of variables by defining $\mathbf{z} = \mathbf{V}\mathbf{x}_{-m}$. However, the integration region of $\mathbf{z}$ is $M-1$ dimensional, this makes the integral hard to compute. For simplicity, we enlarge the integration region by setting a bound $Z_i = U\sum_{j=1}^{M-1}|V_{i,j}|$ for dimension i. We can then upperbound (22) as

$$\begin{aligned}\Pr(\mathbf{y}|x_m) &\propto \exp\left(-\frac{\|\mathbf{y} - \mathbf{h}_m x_m\|^2}{\sigma^2}\right) \\ &\prod_{i=1}^{M-1}\int_{-Z_i}^{+Z_i} \exp\left(-2g_i(x_m)z_i - \lambda_i Z_i^2\right) dz_i\end{aligned} \qquad (23)$$

Note that the second product in (23) also depends on x_m. In some cases, λ_i may be negative, so we cannot write the integral into Q-function.

V. SIMULATION RESULTS

In this section, we consider a 2 by 2 MIMO-OFDM system with 1024 subcarriers and 960 subcarriers are used for data transmission. Perfect knowledge of channel state information is assumed. Each transmit antenna is assigned power P. The SNR is defined as P/N_0, where N_0 is the noise power. We consider Extended Vehicular A model (EVA) [22] in this section with delay profile [0 30 150 310 370 710 1090 1730 2510]ns and power profile $[0\ -1.5\ -1.4\ -3.6\ -0.6\ -9.1\ -7\ -12\ -16.9]$ dB. The channel power profile is normalized to unity. The scheme of turbo encoder is a Parallel Concatenated Convolutional Code (PCCC) with two 8-state constituent encoders and one turbo code internal interleaver. The coding rate of turbo encoder is 1/3. The transfer function of the PCCC is: $G(D) = \left[1, \frac{g_0(D)}{g_1(D)}\right]$ [23], where $g_0(D) = 1 + D^2 + D^3, g_1(D) = 1 + D + D^3$. Eight iterations are performed within the turbo decoder. 64 QAM and Gray mapping are considered in this section.

The algorithms using (8), (9), (10), (14) are denoted as MAP, MLM, LSD and MMSE-Soft SIC, respectively. MMSE-No SIC denotes using (14) without iterations. In LSD, we choose the list size $L = 512$ to be consistent with [14]. The non-Gaussian approximation algorithm in Section IV-B is denoted as Non-Gaussian.

A. BER Comparison of Different Algorithms

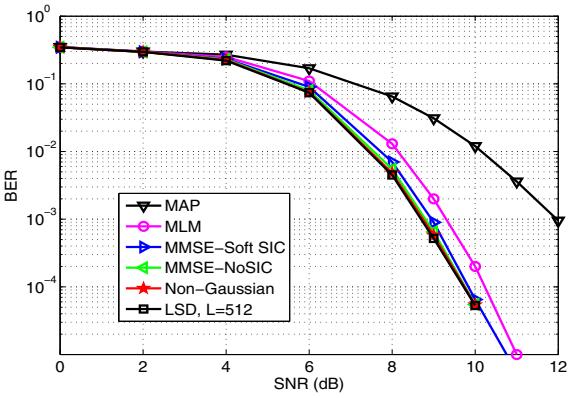

Fig. 3. BER comparison of different algorithms in a 2×2 MIMO-OFDM system over the EVA channel.

We first consider fixed scheduling, where both data streams transmit using transport block size (TBS) 1916. The bit error rates of different algorithms after simulating 20000 subframes are shown in Fig. 3. The channel varies independently from subframe to subframe. All algorithms except MMSE-No SIC use 6 iterations. It is clear that all the iterative algorithms benefit from the information exchange as compared with MMSE-No SIC. We can see that by using only the max term in MAP MLM incurs a 0.5 dB loss over MAP at BER $= 10^{-3}$. MMSE-Soft SIC only has a 0.1 dB loss over MLM at BER $= 10^{-3}$. But the former only needs to sum over 64 terms while the latter needs to compute $32 \times 64 = 2048$ terms in the numerator and denominator

in (9), respectively. LSD with L=512 incurs a 1 dB loss over MAP at BER $= 10^{-3}$ but with a higher complexity than MMSE-Soft SIC. The Non-Gaussian approximation achieves a 0.22 dB gain over MMSE-Soft SIC.

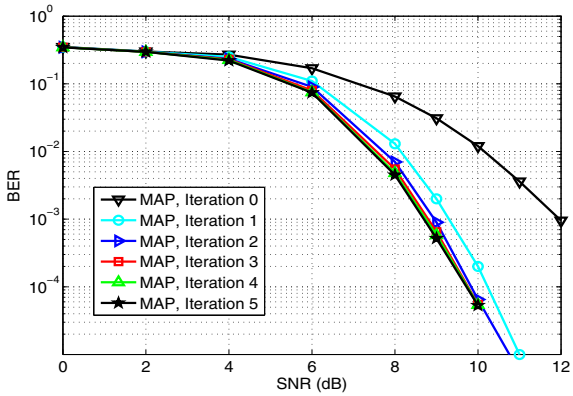

Fig. 4. BER comparison of MAP algorithm with different number of iterations in a 2×2 MIMO-OFDM system over the EVA channel.

In practice, it is important that the decoding time satisfies the timeline constraint. It is important to understand the performance of the iterative decoders in terms of the number of iterations. Fig. 4 shows the BER performance of the MAP decoder with different number of iterations. We can see that marginal performance gain diminishes as the number of iterations increase. Three to four iterations is good enough to achieve close to optimal performance.

VI. CONCLUSIONS

There is currently a significant interest in the design of receiver to meet the increasing requirement on high data rate of networks. In this paper, we have developed iterative detection and decoding algorithms for MIMO-OFDM with Non-Gaussian approximation, which was proposed to enhance the performance of interference cancellation based detectors with large constellations. Simulation results demonstrate that the proposed algorithms can achieve a significant performance gain over existing ones with high order constellations.

VII. ACKNOWLEDGMENT

This work is partly supported by the National Natural Science Foundation of China (Grant no. 61172105, 61322107, 91124001, 61101079, and 61101220), the Chinese Academy of Sciences (Grant no. 2F11C01), Ministry of Transport (Grant no. 2012-364-X03-104), the Science Foundation for the Youth Scholar of Ministry of Education of China (Grant no. 20110001120129) and the Early Career Development Award of SKLMCCS (Grant no. Y3S9021F34).

REFERENCES

[1] C. Mecklenbrauker, A. Molisch, J. Karedal, F. Tufvesson, A. Paier, L. Bernado, T. Zemen, O. Klemp, and N. Czink, "Vehicular channel characterization and its implications for wireless system design and performance," *Proc. of the IEEE*, vol. 99, no. 7, pp. 1189–1212, 2011.

[2] L. Yang and F.-Y. Wang, "Driving into intelligent spaces with pervasive communications," *IEEE Intell. Syst.*, vol. 22, no. 1, pp. 12–15, Jan. 2007.

[3] X. Cheng, X. Hu, L. Yang, I. Husain, K. Inoue, P. Krein, R. Lefevre, Y. Li, H. Nishi, J. Taiber, F. Wang, Y. Zha, W. Gao, and Z. Li, "Electrified vehicles and the smart grid: the ITS perspective," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 4, pp. 1388–1404, Aug. 2014.

[4] X. Cheng, C.-X. Wang, B. Ai, and H. Aggoune, "Envelope level crossing rate and average fade duration of nonisotropic vehicle-to-vehicle Ricean fading channels," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 62–72, Feb 2014.

[5] X. Cheng, Q. Yao, M. Wen, C.-X. Wang, L.-Y. Song, and B.-L. Jiao, "Wideband channel modeling and intercarrier interference cancellation for vehicle-to-vehicle communication systems," *IEEE J. Sel. Areas Commun*, vol. 31, no. 9, pp. 434–448, Sep. 2013.

[6] X. Cheng, Q. Yao, C.-X. Wang, B. Ai, G. Stuber, D. Yuan, and B.-L. Jiao, "An improved parameter computation method for a MIMO V2V Rayleigh fading channel simulator under Non-Isotropic scattering environments," *IEEE Commun. Lett.*, vol. 17, no. 2, pp. 265–268, Feb. 2013.

[7] X. Cheng, C.-X. Wang, D. Laurenson, S. Salous, and A. Vasilakos, "An adaptive geometry-based stochastic model for non-isotropic MIMO mobile-to-mobile channels," *IEEE Trans. Wireless Comm.*, vol. 8, no. 9, pp. 4824–4835, Sep. 2009.

[8] K. Zhong, Y.-C. Wu, and S. Li, "Signal detection for OFDM-based virtual MIMO systems under unknown doubly selective channels, multiple interferences and phase noises," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5309–5321, 2013.

[9] "IEEE Standard for Information technology– Local and metropolitan area networks– Specific requirements– Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments," *IEEE Std 802.11p-2010 (Amendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008, IEEE Std 802.11r-2008, IEEE Std 802.11y-2008, IEEE Std 802.11n-2009, and IEEE Std 802.11w-2009)*, pp. 1–51, July 2010.

[10] X. Cheng, B. Yu, L. Yang, J. Zhang, G. Liu, Y. Wu, and L. Wan, "Communicating in the real world: 3D MIMO," *IEEE Wireless Commun.*, vol. 21, no. 4, pp.136–144, Aug. 2014.

[11] X. Cheng, C.-X. Wang, H. Wang, X. Gao, X.-H. You, D. Yuan, B. Ai, Q. Huo, L.-Y. Song, and B.-L. Jiao, "Cooperative MIMO channel modeling and multi-link spatial correlation properties," *IEEE J. Sel. Areas Commun*, vol. 30, no. 2, pp. 388–396, Feb. 2012.

[12] G. Golden, C. J. Foschini, R. Valenzuela, and P. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 35, no. 1, pp. 14–16, Jan 1999.

[13] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1639–1642, Jul 1999.

[14] B. 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. Han, T. Cui, and C. Tellambura, "Improved K-best sphere detection for uncoded and coded MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 1, no. 5, pp. 472 –475, Oct. 2012.

[16] Z. Guo and P. Nilsson, "Algorithm and implementation of the K-best sphere decoding for MIMO detection," *IEEE J. Sel. Areas Commun*, vol. 24, no. 3, pp. 491–503, Mar. 2006.

[17] 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.

[18] M. Tuchler, 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.

[19] D. Pham, K. Pattipati, P. Willett, and J. Luo, "A generalized probabilistic data association detector for multiple antenna systems," in *IEEE International Conference on Communications*, vol. 6, June 2004, pp. 3519–3522 Vol.6.

[20] S. Liu and Z. Tian, "Near-optimum soft decision equalization for frequency selective MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 721–733, March 2004.

[21] S. Ten Brink, J. Speidel, and R.-H. Yan, "Iterative demapping and decoding for multilevel modulation," in *IEEE Global Telecommunications Conference*, vol. 1, 1998, pp. 579–584 vol.1.

[22] "Evolved Universal Terrestrial Radio Access (E-UTRA); user equipment (UE) radio transmission and reception," The 3rd Generation Partnership Project, TS 36.101, Mar. 2014.

[23] "Evolved Universal Terrestrial Radio Access (E-UTRA); multiplexing and channel coding," The 3rd Generation Partnership Project, TS 36.212, Dec. 2013.