

## Research Article

# Low-Complexity Gaussian Detection for MIMO Systems

**Tianbin Wo and Peter Adam Hoeher**

*The Information and Coding Theory Lab, University of Kiel, Kaiserstrasse 2, 24143 Kiel, Germany*

Correspondence should be addressed to Tianbin Wo, [wtb@tf.uni-kiel.de](mailto:wtb@tf.uni-kiel.de)

Received 13 March 2010; Accepted 30 August 2010

Academic Editor: Christian Schlegel

Copyright © 2010 T. Wo and P. A. Hoeher. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For single-carrier transmission over delay-spread multi-input multi-output (MIMO) channels, the computational complexity of the receiver is often considered as a bottleneck with respect to (w.r.t.) practical implementations. Multi-antenna interference (MAI) together with intersymbol interference (ISI) provides fundamental challenges for efficient and reliable data detection. In this paper, we carry out a systematic study on the interference structure of MIMO-ISI channels, and sequentially deduce three different Gaussian approximations to simplify the calculation of the global likelihood function. Using factor graphs as a general framework and applying the Gaussian approximation, three low-complexity iterative detection algorithms are derived, and their performances are compared by means of Monte Carlo simulations. After a careful inspection of their merits and demerits, we propose a graph-based iterative Gaussian detector (GIGD) for severely delay-spread MIMO channels. The GIGD is characterized by a strictly linear computational complexity w.r.t. the effective channel memory length, the number of transmit antennas, and the number of receive antennas. When the channel has a sparse ISI structure, the complexity of the GIGD is strictly proportional to the number of nonzero channel taps. Finally, the GIGD provides a near-optimum performance in terms of the bit error rate (BER) for repetition encoded MIMO systems.

## 1. Introduction

In single-carrier mobile transmission systems not exploiting a guard interval, there are two sources of intersymbol interference (ISI): static ISI due to pulse shaping and receive filtering, and dynamic ISI due to the time-varying delay spread of the physical channel. Static ISI degrades the receiver performance, but can be avoided or limited by proper signal design. Dynamic ISI is particularly severe if the delay spread exceeds the symbol period, which is likely the case for high-rate data transmission. Dynamic ISI, however, provides a diversity gain in the time domain (fast fading) and the frequency domain (multipath fading). In addition to ISI, MIMO-ISI channels are characterized by another type of interference, namely, multi-antenna interference (MAI), which is caused by the simultaneous transmission of data streams via multiple antennas. MAI together with ISI manifests a fundamental challenge for efficient and reliable data detection. On the other hand, MAI provides a diversity gain in the spatial domain, from an information theoretical point of view.

There are two obvious facts that impede a practical implementation of high-rate single-carrier transmission over MIMO-ISI channels. First, with increasing signal bandwidth the effective channel memory length increases, which degrades the system performance in case of linear or decision-feedback equalization. Second, state-space-based detectors, such as the Viterbi algorithm [1, 2] and the BCJR algorithm [3], provide an excellent performance since they benefit from the diversity gain of dynamic ISI and MAI, but their computational complexity is typically prohibitive. Therefore, multi-carrier transmission schemes, particularly orthogonal frequency-division multiplexing (OFDM) [4], are often applied to circumvent the problem of ISI. An important question is if it is truly impossible to implement a single-carrier transmission system with reasonable performance and complexity for MIMO-ISI channels. We will try to answer this question by proposing a new detection algorithm, called graph-based iterative Gaussian detector (GIGD).

As the detection complexity of MIMO-ISI channels is mainly caused by multi-antenna interference and intersymbol interference, we will first carry out a systematic study on

the interference structure and try to find the opportunities of easy treatment. Based on the knowledge obtained from this study, we deduce three different Gaussian approximations, namely, joint Gaussian approximation (JGA), grouped joint Gaussian approximation (GJGA), and independent Gaussian approximation (IGA), to simplify the calculation of the global likelihood function and sequentially reduce the data detection complexity. The JGA is already well known [5–10], while the GJGA and the IGA are new approaches proposed by the authors. Corresponding to these three Gaussian approximations, three low-complexity iterative parallel soft interference cancellation [5, 11] algorithms, namely, joint Gaussian detector (JGD), grouped joint Gaussian detector (GJGD), and graph-based iterative Gaussian detector (GIGD), will be described by utilizing factor graphs [12, 13] as a general framework. From the JGD to the GJGD, and from the GJGD to the GIGD, the detection complexity is reduced dramatically in each step.

For severely delay-spread MIMO-ISI channels, we propose the GIGD as a promising solution. Applying the independent Gaussian approximation, the GIGD has a computational complexity strictly linear w.r.t. the number of nonzero channel taps, the number of transmit antennas, and the number of receive antennas. Meanwhile, the performance loss incurred by the independent Gaussian approximation can be well compensated by using a repetition code. More importantly, the GIGD shows a satisfying capability in exploiting the frequency/time/space diversity provided by the MIMO-ISI fading channels.

The remainder of this paper is organized as follows. Section 2 introduces a conventional output-oriented channel model as well as a symbol-oriented channel model. Section 3 provides a deep insight into the interference structure of MIMO-ISI channels, and Section 4 gives a brief introduction on factor graphs and message passing algorithms. Section 5 revises the known joint Gaussian detector, Section 6 derives a grouped joint Gaussian detector, and Section 7 proposes a graph-based iterative Gaussian detector. Numerical results by means of Monte Carlo simulations are provided in Section 8 and Section 9 to assess and compare the performance of the three Gaussian detectors. Finally, conclusions are drawn in Section 10.

## 2. Channel Model

In this section, we will first introduce a conventional MIMO-ISI channel model, and then convert it into a symbol-oriented channel model to facilitate the mathematical derivation of the new algorithms.

*2.1. Output-Oriented Channel Model.* The equivalent discrete-time model of a MIMO-ISI channel (including transmit and receive filters, physical channel and symbol-rate sampling) can be written in complex baseband notation as

$$y_n[k] = \sum_{m=1}^{N_T} \sum_{l=0}^L h_{n,m}^l \cdot x_m[k-l] + w_n[k], \quad 1 \leq n \leq N_R, \quad (1)$$

where  $N_R$  denotes the number of receive (Rx) antennas,  $N_T$  the number of transmit (Tx) antennas,  $L$  the effective memory length of all subchannels, and  $k \in \{0, 1, \dots, K-1\}$  the discrete time index with  $K$  denoting the block length.  $y_n[k] \in \mathbb{C}$  is the channel output sample at the  $n$ th Rx antenna at time index  $k$ , and  $x_m[k]$  is the channel input symbol at the  $m$ th Tx antenna at time index  $k$ .  $h_{n,m}^l \in \mathbb{C}$  marks the  $l$ th tap of the subchannel connecting the  $n$ th Rx antenna and the  $m$ th Tx antenna.  $w_n[k]$  represents a complex additive white Gaussian noise (AWGN) sample at the  $n$ th Rx antenna at time index  $k$  with zero mean and variance  $\sigma_w^2$ . By convention, the single-sided noise spectral density in the passband is denoted by  $N_0$ . Noting that  $w_n[k]$  is a complex noise sample, we have  $\sigma_w^2 = N_0/2 + N_0/2 = N_0$ . Throughout this paper, the signal-to-noise ratio per info bit will be defined as  $E_b/N_0$ , where  $E_b$  stands for the energy used for transmitting one info bit. In case of coded transmission, we have  $E_b = E_s/R$  with  $E_s \doteq E\{|x_m[k]|^2\}$  denoting the energy used for transmitting one symbol and  $R$  denoting the coding rate.

We assume that all channel taps are constant within each data burst while varying independently from burst to burst. Moreover, we assume that the fading processes of channel taps all have the same average power and are mutually independent. This equal delay power profile is often used for the purpose of equalizer test, for example, in the 3GPP GSM standard, since it is the most challenging case for linear equalization. Nevertheless, we will show that low-complexity high-performance data detection is in fact possible for this type of MIMO-ISI channel, by means of the receiver algorithm proposed in this paper.

If we take a second look at (1), we may recognize that it is actually an output-oriented channel model, that is, this channel model explains how a channel output sample is formed given multiple channel inputs. Such kind of channel model is convenient to derive state-space-based detection algorithms, but inconvenient for the derivation of factor-graph-based detection algorithms, which requires a channel model that explicitly states the information spread of a data symbol over multiple channel outputs.

*2.2. Symbol-Oriented Channel Model.* Let us consider an arbitrary data symbol  $x_m[k]$ . Due to multiple Rx antennas and delay spread, there will be in total  $N_R(L+1)$  channel outputs containing the information of  $x_m[k]$ . From now on, we call these channel outputs the observations of symbol  $x_m[k]$ . To facilitate the following mathematical elaboration, we collect the observations of  $x_m[k]$  into a matrix

$$\mathbf{Y}[k] = \begin{bmatrix} y_1[k] & y_1[k+1] & \dots & y_1[k+L] \\ y_2[k] & y_2[k+1] & \dots & y_2[k+L] \\ \vdots & \vdots & \ddots & \vdots \\ y_{N_R}[k] & y_{N_R}[k+1] & \dots & y_{N_R}[k+L] \end{bmatrix} \quad (2)$$

which may be termed the observation matrix of  $x_m[k]$ . Note that  $\mathbf{Y}[k]$  is shared by all  $x_m[k]$  for  $m = 1, 2, \dots, N_T$ . Hence, there is no necessity for  $\mathbf{Y}[k]$  to have a subscript  $m$  making this distinction.

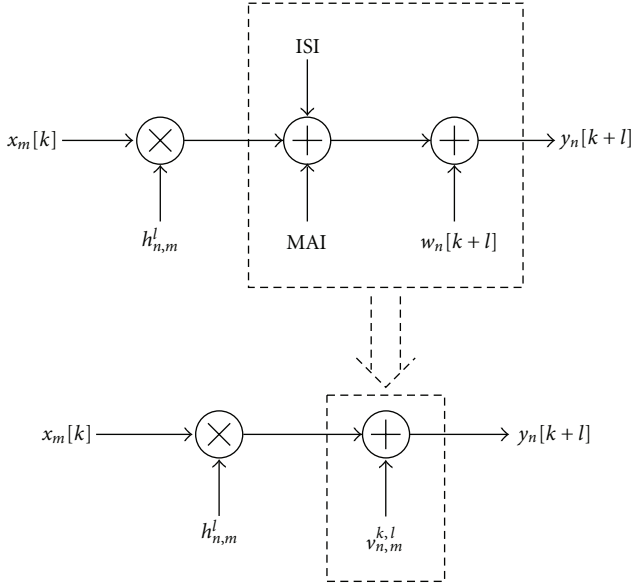


FIGURE 1: Relationship between a data symbol and one of its observations.

Revisiting (1), we find that the relationship between  $x_m[k]$  and one of its observations  $y_n[k+l]$  ( $1 \leq n \leq N_R$ ,  $0 \leq l \leq L$ ) can be written as

$$\begin{aligned}
 y_n[k+l] &= \sum_{i=1}^{N_T} \sum_{j=0}^L h_{n,i}^j x_i[k+l-j] + w_n[k+l] \\
 &= h_{n,m}^l x_m[k] + \underbrace{\sum_{j=0, j \neq l}^L h_{n,m}^j x_m[k+l-j]}_{\text{ISI}} \\
 &\quad + \underbrace{\sum_{i=1, i \neq m}^{N_T} \sum_{j=0}^L h_{n,i}^j x_i[k+l-j]}_{\text{MAI}} + \underbrace{w_n[k+l]}_{\text{AWGN}}.
 \end{aligned} \tag{3}$$

Defining the summation of ISI, MAI, and AWGN as an effective noise term  $v_{n,m}^{k,l}$ , the relationship between  $x_m[k]$  and  $y_n[k+l]$  can be simplified as

$$y_n[k+l] = h_{n,m}^l x_m[k] + v_{n,m}^{k,l}, \tag{4}$$

c.f. Figure 1. Combining (2) and (4), we obtain the following symbol-oriented channel model:

$$\mathbf{Y}[k] = \mathbf{H}_m x_m[k] + \mathbf{V}_m^k, \tag{5}$$

with

$$\begin{aligned}
 \mathbf{H}_m &= \begin{bmatrix} h_{1,m}^0 & h_{1,m}^1 & \cdots & h_{1,m}^L \\ h_{2,m}^0 & h_{2,m}^1 & \cdots & h_{2,m}^L \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_R,m}^0 & h_{N_R,m}^1 & \cdots & h_{N_R,m}^L \end{bmatrix} \\
 \mathbf{V}_m^k &= \begin{bmatrix} v_{1,m}^{k,0} & v_{1,m}^{k,1} & \cdots & v_{1,m}^{k,L} \\ v_{2,m}^{k,0} & v_{2,m}^{k,1} & \cdots & v_{2,m}^{k,L} \\ \vdots & \vdots & \ddots & \vdots \\ v_{N_R,m}^{k,0} & v_{N_R,m}^{k,1} & \cdots & v_{N_R,m}^{k,L} \end{bmatrix}
 \end{aligned} \tag{6}$$

being the channel matrix of the  $m$ th Tx antenna and the effective noise matrix in  $\mathbf{Y}[k]$  w.r.t.  $x_m[k]$ , respectively.

With this new channel model, it is clear that all information about  $x_m[k]$  that we can extract from the channel outputs is fully represented by the following global likelihood function:

$$p(\mathbf{Y}[k] | x_m[k]) = p(\mathbf{V}_m^k = \mathbf{Y}[k] - \mathbf{H}_m x_m[k]). \tag{7}$$

Now, the question is how to calculate this likelihood function in an efficient manner. According to (7), the key for this task is the probability density function (PDF) of the effective noise matrix, that is,  $p(\mathbf{V}_m^k)$ . As a matter of fact, the main differences between the three Gaussian detectors to be described are in their way of dealing with  $p(\mathbf{V}_m^k)$ .

### 3. Statistical Properties of the Effective Noise Matrix

From (3), (4), and (5), we see that the effective noise matrix  $\mathbf{V}_m^k$  consists of multi-antenna interference, intersymbol interference, and additive noise samples. Due to the large amount of variables involved in  $\mathbf{V}_m^k$ , an exact calculation of  $p(\mathbf{V}_m^k)$  typically incurs a prohibitive complexity. Therefore, reasonable approximations are necessary to make things easier. In this section, we will carefully study the statistical properties of the effective noise matrix and try to find a way towards complexity reduction.

**3.1. Distribution of Effective Noise Samples.** Noting that each effective noise sample  $v_{n,m}^{k,l}$  is a sum of  $N_T(L+1)$  independent random variables, its probability density function may be approximated by a complex Gaussian distribution:

$$p(v_{n,m}^{k,l}) \approx \frac{1}{\pi \sigma_v^2} \exp\left(-\frac{|v_{n,m}^{k,l} - \mu_v|^2}{\sigma_v^2}\right), \tag{8}$$

where  $\mu_v$  and  $\sigma_v^2$  are defined as

$$\mu_v \doteq E\{v_{n,m}^{k,l}\}, \quad \sigma_v^2 \doteq E\{|v_{n,m}^{k,l} - \mu_v|^2\}. \tag{9}$$

(Here, we neglect the correlation between the real part and the imaginary part. Concerning this issue, interested readers may refer to [7].) According to the rule of thumb, as long as

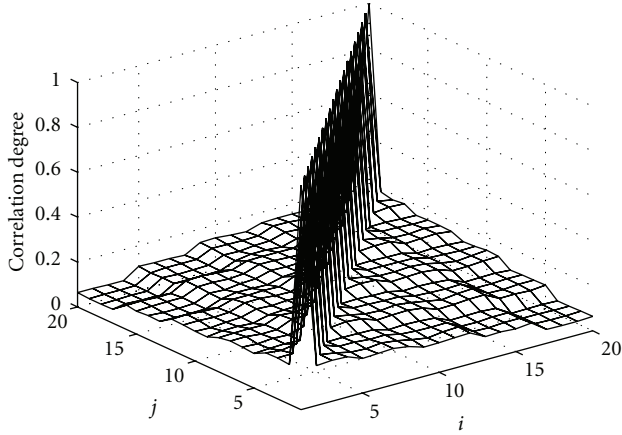


FIGURE 2: Average magnitude of correlation coefficient between effective noise samples,  $N_T = N_R = 4$ ,  $L = 4$ , BPSK mapping, and  $E_b/N_0 = 4$  dB.

$N_T(L + 1) \geq 12$  holds, the accuracy of (8) is satisfying. This approximation is often called Gaussian approximation, and its feasibility in the scenario of MIMO-ISI channels has been proven in the available literature [6, 8, 9].

**3.2. Dependence between Effective Noise Samples.** Due to more or less common sources of randomness, the elements of  $\mathbf{V}_m^k$  are in general statistically dependent on each other. However, it is so far unclear whether this dependence is strong or weak. In the following, we will carry out some numerical measurements to obtain a deeper insight into this issue. Many previous works [6–10] show that  $p(\mathbf{V}_m^k)$  can be well approximated by a joint Gaussian distribution, as long as the product  $N_T(L + 1)$  is large enough. Besides, it is well known that two jointly Gaussian distributed variables are independent if they are uncorrelated, and their dependence structure is completely defined by the correlation coefficient. Therefore, by measuring the correlation between the elements of  $\mathbf{V}_m^k$ , we will be able to get a rough impression on the dependence between the elements of  $\mathbf{V}_m^k$ .

First, we define that

$$\mathbf{v} = [v_1, v_2, \dots, v_Q]^T \doteq \text{vec}\{\mathbf{V}_m^k\}, \quad (10)$$

with  $Q \doteq N_R(L + 1)$ .  $\text{vec}\{\cdot\}$  denotes the column stacking operator and  $(\cdot)^T$  denotes the matrix/vector transpose operator. Since for a block-fading channel the statistics of  $\mathbf{V}_m^k$  do not change with  $m$  and  $k$ , the subscript  $m$  and the superscript  $k$  are omitted in  $\mathbf{v}$ . Next, we define the magnitude of the correlation coefficient between two effective noise samples as

$$\varphi_{i,j} \doteq \frac{\left| \mathbf{E}\left\{ (v_i - \mu_{v_i})(v_j - \mu_{v_j})^* \right\} \right|}{\sigma_{v_i} \sigma_{v_j}}, \quad (11)$$

where  $(\cdot)^*$  denotes complex conjugate. Since  $\varphi_{i,j}$  is in fact a function of the random channel taps, we further define

$$\phi_{i,j} \doteq \mathbf{E}\{\varphi_{i,j}\}, \quad (12)$$

where the expectation is taken over random realizations of channel taps. Last, we collect  $\phi_{i,j}$  into a matrix

$$\Phi \doteq \begin{bmatrix} \phi_{1,1} & \phi_{1,2} & \dots & \phi_{1,Q} \\ \phi_{2,1} & \phi_{2,2} & \dots & \phi_{2,Q} \\ \dots & \dots & \ddots & \dots \\ \phi_{Q,1} & \phi_{Q,2} & \dots & \phi_{Q,Q} \end{bmatrix}. \quad (13)$$

Clearly, the entries on the main diagonal of  $\Phi$  will always be 1, because these entries are the magnitudes of autocorrelation coefficients. For entries not on the main diagonal of  $\Phi$ , their values reflect the strongness of correlation between effective noise samples and sequentially the strongness of dependence between effective noise samples.

Figure 2 demonstrates the measured values of  $\Phi$  in a BPSK system with independent Rayleigh fading channel taps and an equal delay power profile. Observing Figure 2, we see that the values of  $\phi_{i,j}$  ( $i \neq j$ ) are small, which means that the correlation between the elements of  $\mathbf{V}_m^k$  is actually very weak. As a matter of fact, the correlation between effective noise samples drops steadily as the product  $N_T(L + 1)$  increases [14]. This observation delivers a good message: it may be feasible to partially or even fully neglect the mutual dependence between the effective noise samples, for the sake of complexity reduction. Certainly, the detailed dependence structure of effective noise samples will be different from Figure 2 if one uses another type of channel delay power profile. However, the contour of Figure 2 holds in general.

## 4. Factor Graph and Message Passing

Before specific algorithm derivation, we briefly revisit the concept of factor graphs and message passing.

**4.1. Factor Graphs and Factorization.** A factor graph is a type of bipartite graph which visualizes the factorization of certain global functions object to maximization or minimization. To easily understand it, let us consider a simple example. Suppose that we have a BPSK symbol  $x$  with three observations:

$$y_1 = x + n_1, \quad y_2 = x + n_2, \quad y_3 = x + n_3, \quad (14)$$

where  $n_1$ ,  $n_2$ , and  $n_3$  are additive noise terms. Assuming that no a priori information is available for  $x$ , an optimal detector tries to maximize the global likelihood function according to

$$\hat{x} = \arg \max_{\tilde{x} \in \{\pm 1\}} \{p(y_1, y_2, y_3 | \tilde{x})\}. \quad (15)$$

If  $n_1$ ,  $n_2$ , and  $n_3$  are mutually independent, we may factorize the above global likelihood function into a product of local likelihood functions:

$$p(y_1, y_2, y_3 | x) = p(y_1 | x)p(y_2 | x)p(y_3 | x), \quad (16)$$

which can be visualized by the factor graph given in Figure 3, where a circle represents a symbol node and a square box represents an observation node.

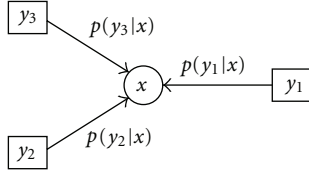


FIGURE 3: A symbol node connected with three observation nodes.

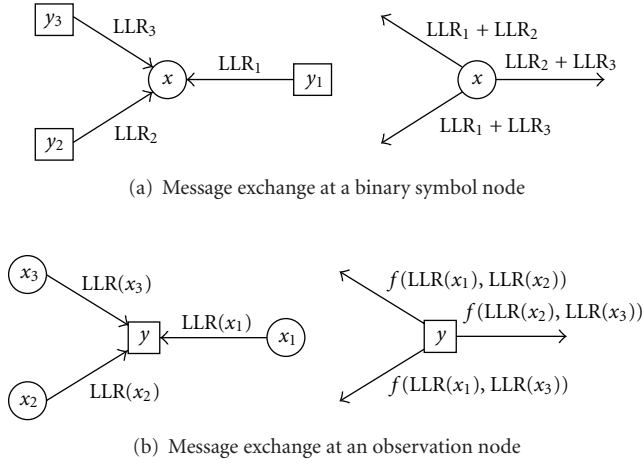


FIGURE 4: Message exchange at different nodes.

**4.2. Iterative Message Passing Algorithms.** Given a factor graph, the task of variable estimation can be accomplished by combining and exchanging the messages (knowledge) from various sources over this probabilistic network. Such an algorithm is often called an iterative message passing algorithm. For message passing over factor graphs, only extrinsic information should be exchanged and propagated. Although different type of nodes often apply different type of message processing operations, this rule must be carefully followed.

**4.3. Message Exchange at Symbol Nodes.** For binary variables, it is often convenient to use log-likelihood ratios (LLRs). Define that

$$\text{LLR}_i \doteq \ln \frac{p(y_i | x = +1)}{p(y_i | x = -1)}, \quad (17)$$

the message exchanging at a BPSK symbol node proceeds as Figure 4(a). The underlying principle is that LLR messages from independent observations are additive. In practice,  $\text{LLR}_{\text{sum}} \doteq \sum_i \text{LLR}_i$  is first calculated, then each new message is obtained as  $(\text{LLR}_{\text{sum}} - \text{LLR}_i)$ . Consequently, the complexity of this operation is always proportional to the amount of edges diverging from this symbol node.

**4.4. Message Exchange at Observation Nodes.** Considering an observation node connected with three BPSK symbols, the message exchange proceeds as illustrated in Figure 4(b), where  $f(\cdot)$  denotes a certain message combining function,

often called a message update rule. Different from the situation at symbol nodes, here message combining can no longer be accomplished by a simple linear addition. As a matter of fact,  $f(\cdot)$  is the major source of complexity in a graph-based detection algorithm, and hence will be the object of simplification in the remaining part of this paper.

## 5. Joint Gaussian Detector

According to Section 3, the elements of  $\mathbf{V}_m^k$  are roughly Gaussian distributed, and they are in general dependent on each other, although weakly. Hence, a straightforward way to calculate  $p(\mathbf{V}_m^k)$  is to approximate the elements of  $\mathbf{V}_m^k$  as jointly Gaussian distributed. This approach is usually termed joint Gaussian approximation (JGA), and the algorithm based on this approach is called joint Gaussian detector (JGD), which has been known for years [6–10]. In this section, we will give a clean mathematical derivation of the JGD (For the sake of simple mathematical expression, BPSK mapping is assumed in the rest of the paper.).

**5.1. Joint Gaussian Approximation.** Using the symbol-oriented channel model (5), the joint Gaussian approximation can be written as

$$p(\mathbf{V}_m^k) \approx \frac{1}{\pi^Q |\boldsymbol{\Sigma}|^Q} \exp(-(\mathbf{v} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{v} - \boldsymbol{\mu})) \quad (18)$$

with

$$\begin{aligned} \mathbf{v} &\doteq \text{vec}\{\mathbf{V}_m^k\}, & \boldsymbol{\mu} &\doteq \text{E}\{\mathbf{v}\}, \\ \boldsymbol{\Sigma} &\doteq \text{E}\{(\mathbf{v} - \boldsymbol{\mu})(\mathbf{v} - \boldsymbol{\mu})^H\}, & Q &\doteq N_R(L + 1). \end{aligned} \quad (19)$$

Note that  $\mathbf{v}$  is a  $Q \times 1$  column vector. Therefore, the order of the covariance matrix  $\boldsymbol{\Sigma}$  is  $Q = N_R(L + 1)$ . In the literature, however, this covariance matrix usually has an order  $N_R K$ , where  $K$  is the burst length, due to using an output-oriented channel model. The concept of sliding windows is introduced in [6] in order to reduce this order from  $N_R K$  to  $N_R(L + 1)$ . Nevertheless, with the symbol-oriented channel model, it is clarified that there is in fact no reason for the order of  $\boldsymbol{\Sigma}$  to be related to the burst length.

**5.2. Factor Graph with Joint Gaussian Approximation.** Applying the joint Gaussian approximation, we admit the mutual dependence between the elements of  $\mathbf{V}_m^k$ , and hence the PDF  $p(\mathbf{V}_m^k)$  as well as the global likelihood function  $p(\mathbf{Y}[k] | x_m[k])$  will not be factorizable at all. We also notice that the observation matrices for neighboring data symbols, namely,  $\mathbf{Y}[k], \mathbf{Y}[k + 1], \dots, \mathbf{Y}[k + L]$ , partially overlap with each other. For these two reasons, the factor graph of a MIMO-ISI channel will look like Figure 5, where  $\mathbf{Y}$  denotes the matrix which collects all channel outputs within the current data burst. No factorization exists and also no cycles are present.



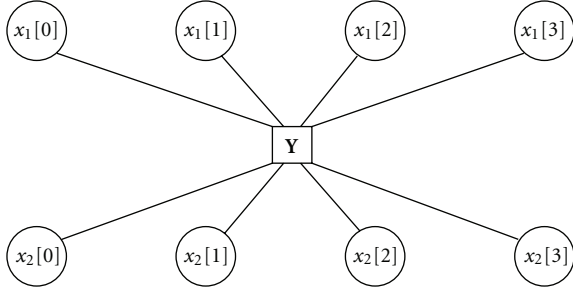


FIGURE 5: Factor graph of a MIMO-ISI channel with joint Gaussian approximation,  $N_T = 2$ .

5.3. *Message Update Rule at Observation Node.* Revisiting (7) and applying (18), the message from an observation node to a symbol node can be calculated as

$$\begin{aligned} \text{LLR}(x_m[k]) &= \ln \frac{p(\mathbf{Y}[k] | x_m[k] = +1)}{p(\mathbf{Y}[k] | x_m[k] = -1)} \\ &= -(\mathbf{v}_1 - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{v}_1 - \boldsymbol{\mu}) \\ &\quad + (\mathbf{v}_2 - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{v}_2 - \boldsymbol{\mu}) \end{aligned} \quad (20)$$

with

$$\mathbf{v}_1 = \text{vec}\{\mathbf{Y}[k] - \mathbf{H}_m\}, \quad \mathbf{v}_2 = \text{vec}\{\mathbf{Y}[k] + \mathbf{H}_m\}. \quad (21)$$

$\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$ , covering the statistical properties of the effective noise matrix  $\mathbf{V}_m^k$ , are calculated according to (19), utilizing the incoming LLR messages from all relevant symbol nodes. Due to limited space, we would like to refer interested readers to [6] for a detailed description of this calculation.

5.4. *Computational Complexity.* The computational complexity of (20) mainly comes from the inversion of the covariance matrix  $\boldsymbol{\Sigma}$ . Noting that (20) needs to be calculated for  $N_T$  data symbols per time index and matrix inversion is an operation with complexity cubic in the matrix order, we have

$$\mathcal{O}(\text{JGD}) \propto N_T N_R^3 (L + 1)^3. \quad (22)$$

This complexity is much lower than that of the BCJR algorithm, but still is a considerable problem whenever the system possesses many Rx antennas or the channel is severely delay-spread.

## 6. Grouped Joint Gaussian Detector

In this section, we introduce a grouped joint Gaussian approximation (GJGA) of  $p(\mathbf{V}_m^k)$ , which brings a significant complexity reduction w.r.t. the joint Gaussian approximation.

6.1. *Grouped Joint Gaussian Approximation.* From Figure 2 we see that the average magnitude of correlation coefficient between  $v_{n_1,m}^{k,i}$  and  $v_{n_2,m}^{k,i}$  is constant for all  $n_1 \neq n_2$ , while the

average magnitude of correlation coefficient between  $v_{n,m}^{k,i}$  and  $v_{n,m}^{k,j}$  drops steadily as the distance  $(i-j)$  increments. This observation inspires us for a new approximation of  $p(\mathbf{V}_m^k)$  (Initial work has been presented in [15]). As illustrated in the following expression:

$$\mathbf{V}_m^k = \begin{bmatrix} v_{1,m}^{k,0} & v_{1,m}^{k,1} & \cdots & v_{1,m}^{k,L} \\ v_{2,m}^{k,0} & v_{2,m}^{k,1} & \cdots & v_{2,m}^{k,L} \\ \vdots & \vdots & \ddots & \vdots \\ v_{N_R,m}^{k,0} & v_{N_R,m}^{k,1} & \cdots & v_{N_R,m}^{k,L} \end{bmatrix}. \quad (23)$$

we assume that the columns of  $\mathbf{V}_m^k$  are linearly independent from each other while the elements in each column are jointly Gaussian distributed. Mathematically, this approximation can be written as

$$p(\mathbf{V}_m^k) \approx \prod_{l=0}^L p(\mathbf{v}_m^{k,l}) \quad (24)$$

with

$$\begin{aligned} \mathbf{v}_m^{k,l} &\doteq [v_{1,m}^{k,l}, v_{2,m}^{k,l}, \dots, v_{N_R,m}^{k,l}]^T, \\ p(\mathbf{v}_m^{k,l}) &\propto \exp\left(-(\mathbf{v}_m^{k,l} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{v}_m^{k,l} - \boldsymbol{\mu})\right), \end{aligned} \quad (25)$$

where  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  are the mean vector and the covariance matrix of  $\mathbf{v}_m^{k,l}$ , respectively. Note that the order of  $\boldsymbol{\Sigma}$  is now only  $N_R$ . In the following, we refer to the receiver algorithm based on this approximation as grouped joint Gaussian detector (GJGD).

6.2. *Factor Graph with Grouped Joint Gaussian Approximation.* Applying the grouped joint Gaussian approximation, we achieve the following factorization:

$$p(\mathbf{Y}[k] | x_m[k]) \approx \prod_{l=0}^L p(\mathbf{y}[k+l] | x_m[k]) \quad (26)$$

with

$$\mathbf{y}[k+l] \doteq [y_1[k+l], y_2[k+l], \dots, y_{N_R}[k+l]]^T. \quad (27)$$

The resulting factor graph will look like Figure 6. Now the observation matrix  $\mathbf{Y}[k]$  is split into observation vectors  $\mathbf{y}[k+l]$ . Compared to the factor graph with the JGA, the factor graph with the GJGA becomes more complicated, that is, there are more edges diverging from each symbol node. However, the corresponding detection complexity actually becomes much lower, as explained in Section 6.4.

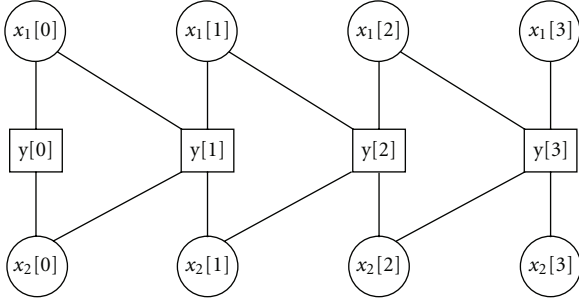


FIGURE 6: Factor graph of a MIMO-ISI channel with grouped joint Gaussian approximation,  $N_T = 2$ ,  $L = 1$ .

6.3. *Message Update Rule at Observation Nodes.* With the new approximation, the message updating rule at an observation node can be written as

$$\begin{aligned} \text{LLR}(x_m[k]) &= \ln \frac{p(\mathbf{y}[k+1] | x_m[k] = +1)}{p(\mathbf{y}[k+1] | x_m[k] = -1)} \\ &= -(\mathbf{v}_1 - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{v}_1 - \boldsymbol{\mu}) \\ &\quad + (\mathbf{v}_2 - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{v}_2 - \boldsymbol{\mu}) \end{aligned} \quad (28)$$

with

$$\begin{aligned} \mathbf{v}_1 &\doteq \mathbf{y}[k+1] - \mathbf{h}_m^l, \\ \mathbf{v}_2 &\doteq \mathbf{y}[k+1] + \mathbf{h}_m^l, \\ \mathbf{h}_m^l &\doteq [h_{1,m}^l, h_{2,m}^l, \dots, h_{N_R,m}^l]^T. \end{aligned} \quad (29)$$

The statistical properties  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  can be calculated by utilizing the incoming LLR messages from all relevant symbol nodes. Due to limited space, we would like to refer interested readers to [15] for more details on this topic.

6.4. *Computational Complexity.* By checking (26) and (28), and noting that the covariance matrix  $\boldsymbol{\Sigma}$  is now only of order  $N_R$ , we have

$$\mathcal{O}(\text{GJGD}) \propto N_T N_R^3 (L+1). \quad (30)$$

Comparing (30) with (22), it is clear that the computational complexity of the GJGD is much lower than that of the JGD, particularly for MIMO systems with severe delay spread. Nevertheless, a cubic term is still present due to matrix inversion.

## 7. Graph-Based Iterative Gaussian Detector

In this section, we introduce an independent Gaussian approximation (IGA) which completely eliminates matrix inversion and a graph-based iterative Gaussian detector (GIGD) based on that (Initial work has been presented in [14]).

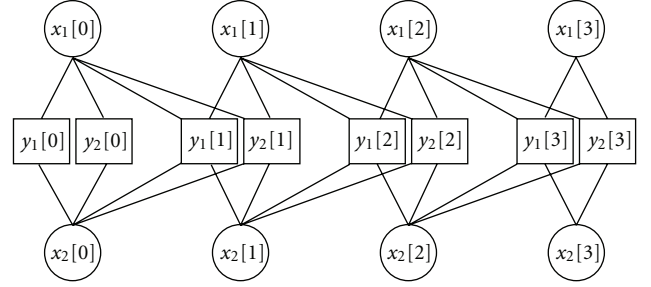


FIGURE 7: Factor graph of a MIMO-ISI channel with independent Gaussian approximation,  $N_T = N_R = 2$ ,  $L = 1$ .

7.1. *Independent Gaussian Approximation.* In Section 3.2, we mentioned that the cross-correlation between effective noise samples drops steadily as the product  $N_T(L+1)$  increases. Therefore, if  $N_T(L+1)$  is sufficiently large, we might completely neglect the mutual dependence, that is, to approximate all effective noise samples to be independently Gaussian distributed, as illustrated in the following:

$$\mathbf{V}_m^k = \begin{bmatrix} \textcircled{v_{1,m}^{k,0}} & \textcircled{v_{1,m}^{k,1}} & \dots & \textcircled{v_{1,m}^{k,L}} \\ \textcircled{v_{2,m}^{k,0}} & \textcircled{v_{2,m}^{k,1}} & \dots & \textcircled{v_{2,m}^{k,L}} \\ \vdots & \vdots & \ddots & \vdots \\ \textcircled{v_{N_R,m}^{k,0}} & \textcircled{v_{N_R,m}^{k,1}} & \dots & \textcircled{v_{N_R,m}^{k,L}} \end{bmatrix}. \quad (31)$$

Mathematically, we may write this approximation as

$$p(\mathbf{V}_m^k) \approx \prod_{n=1}^{N_R} \prod_{l=0}^L p(v_{n,m}^{k,l}) \quad (32)$$

with

$$p(v_{n,m}^{k,l}) \approx \frac{1}{\pi \sigma_v^2} \exp\left(-\frac{|v_{n,m}^{k,l} - \mu_v|^2}{\sigma_v^2}\right), \quad (33)$$

where  $\mu_v$  and  $\sigma_v^2$  are defined as

$$\mu_v \doteq \mathbb{E}\{v_{n,m}^{k,l}\}, \quad \sigma_v^2 \doteq \mathbb{E}\{|v_{n,m}^{k,l} - \mu_v|^2\}. \quad (34)$$

7.2. *Factor Graph with Independent Gaussian Approximation.* Revisiting (7) and applying (32), we achieve the following factorization:

$$p(\mathbf{Y}[k] | x_m[k]) \approx \prod_{n=1}^{N_R} \prod_{l=0}^L p(y_n[k+l] | x_m[k]). \quad (35)$$

The resulting factor graph will look like Figure 7. Now all observations are separately represented in the factor graph, and there are even more edges diverging from each symbol node. However, the corresponding detection complexity is again much lower than that of the GJGD.

7.3. *Message Update Rule at Observation Nodes.* Combining (4) with (33), the message updating rule at an observation node can be written as

$$\begin{aligned} \text{LLR}(x_m[k]) &= \ln \frac{p(y_n[k+l] | x_m[k] = +1)}{p(y_n[k+l] | x_m[k] = -1)} \\ &= - \frac{|y_n[k+l] - \mu_v - h_{n,m}^l|^2}{\sigma_v^2} \\ &\quad + \frac{|y_n[k+l] - \mu_v + h_{n,m}^l|^2}{\sigma_v^2} \\ &= 4 \frac{\text{Re}\{h_{n,m}^l (y_n[k+l] - \mu_v)^*\}}{\sigma_v^2}, \end{aligned} \quad (36)$$

with  $\mu_v$  and  $\sigma_v^2$  as defined in (34) and the way of calculating them described in the following.

Revisiting Figure 7, we see that each observation node is connected with  $G \doteq N_T(L+1)$  symbol nodes. Replacing complicated indices  $n, m, k$ , and  $l$  by a single index  $i$ , we may simplify the relationship between an observation and its associated data symbols as

$$\begin{aligned} y &= \sum_{i=1}^G h_i x_i + w \\ &= h_j x_j + \sum_{i=1, i \neq j}^G h_i x_i + w \\ &= h_j x_j + v_j, \end{aligned} \quad (37)$$

with  $v_j = \sum_{i=1, i \neq j}^G h_i x_i + w$  denoting the effective noise sample w.r.t.  $x_j$ . Since all data symbols are mutually independent, the following statement is straightforward:

$$\begin{aligned} \mu_{v_j} &= \sum_{i=1, i \neq j}^G h_i \cdot \mu_{x_i}, \\ \sigma_{v_j}^2 &= \sum_{i=1, i \neq j}^G |h_i|^2 \cdot \sigma_{x_i}^2 + \sigma_w^2, \end{aligned} \quad (38)$$

where  $\mu_{x_i}$  and  $\sigma_{x_i}^2$  are calculated by utilizing the incoming LLR message from the symbol node:

$$\mu_{x_i} = \frac{e^{\text{LLR}(x_i)} - 1}{e^{\text{LLR}(x_i)} + 1}, \quad \sigma_{x_i}^2 = 1 - \mu_{x_i}^2. \quad (39)$$

Note that the principle of extrinsic information is implicitly applied in this message updating operation.

7.4. *Computational Complexity.* The computational load of the GIGD comes from the message updating at the symbol nodes and the observation nodes. Revisiting Figure 7, we find that there are  $N_T$  symbol nodes per time index, each connected with  $N_R(L+1)$  edges. Since the complexity of

message exchange at a symbol node is always proportional to the amount of associated edges (c.f. Section 4.3), we have

$$\mathcal{O}(\text{operation at symbol nodes}) \propto N_T N_R (L+1). \quad (40)$$

In each iteration, an observation node needs to calculate the LLR values of  $G = N_T(L+1)$  data symbols associated with it. In practice, this task is accomplished in two steps. In step one,  $\mu_{x_i}$  and  $\sigma_{x_i}^2$  are first calculated for  $i = 1, 2, \dots, G$ , according to (39). Afterwards, the products  $h_i \mu_{x_i}$  and  $|h_i|^2 \sigma_{x_i}^2$  as well as the summations  $\sum_{i=1}^G h_i \mu_{x_i}$  and  $(\sum_{i=1}^G |h_i|^2 \sigma_{x_i}^2 + \sigma_w^2)$  are calculated and stored. Obviously, the complexity of this step is proportional to  $G$ . In step two, the following calculation:

$$\mu_{v_j} = \sum_{i=1}^G h_i \mu_{x_i} - h_j \mu_{x_j}, \quad (41)$$

$$\sigma_{v_j}^2 = \left( \sum_{i=1}^G |h_i|^2 \sigma_{x_i}^2 + \sigma_w^2 \right) - |h_j|^2 \sigma_{x_j}^2, \quad (42)$$

is performed and then  $\text{LLR}(x_j)$ ,  $j = 1, 2, \dots, G$ , is obtained according to (36). Since the two sums in (41) and (42) have already been stored in step one, the complexity of step two is proportional to  $G = N_T(L+1)$  as well. Given this explanation and noting that there are  $N_R$  observation nodes per time index, we may conclude that

$$\mathcal{O}(\text{operation at observation nodes}) \propto N_T N_R (L+1). \quad (43)$$

We may recognize that  $N_T N_R (L+1)$  actually gives the number of channel taps. In reality, however, the discrete-time channel model often has a sparse ISI structure, that is, many channel taps are quasi zero. In this case, the edges associated with zero taps can safely be removed from the factor graph (c.f. Figure 7). Given this knowledge, and combining (40) and (43), we obtain the following expression:

$$\begin{aligned} \mathcal{O}(\text{GIGD}) &\propto \text{number of nonzero channel taps} \\ &\leq N_T N_R (L+1). \end{aligned} \quad (44)$$

Due to the complete elimination of matrix inversion, the complexity of the GIGD is truly linear. Besides, the GIGD is very attractive for sparse ISI channels, where the maximum delay spread is large while many zero taps are present. Note that neither the JGD nor the GJGD is able to benefit from the sparse ISI channel structure in such a straightforward manner, because of multivariate Gaussian approximations.

## 8. Performance in Uncoded Systems

In previous sections, we have introduced three low-complexity Gaussian detection algorithms, namely, joint Gaussian detector (JGD), grouped joint Gaussian detector (GJGD), and graph-based iterative Gaussian detector (GIGD). In this section, we provide numerical results from Monte Carlo simulations to assess and compare the performance of these three algorithms in uncoded systems, and ultimately illustrate the merits and demerits of the GIGD algorithm.



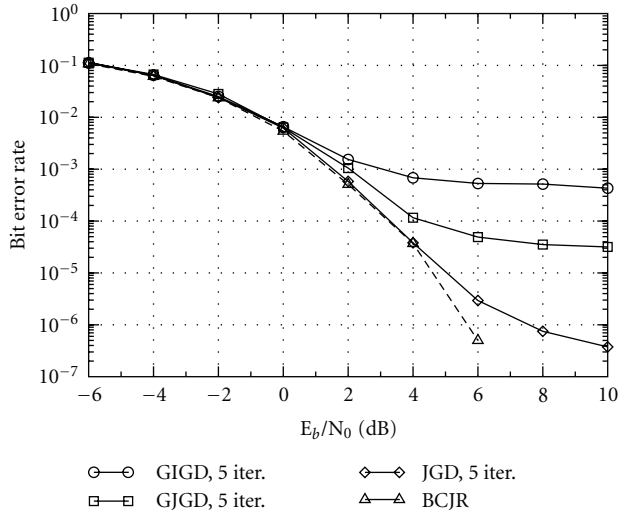


FIGURE 8: Performance of the three Gaussian detectors in an uncoded system,  $N_T = N_R = 4$ ,  $L = 4$ , and  $K = 400$ .

**8.1. Simulation Setup.** Each burst from each Tx antenna contains 400 data symbols. After one burst is transmitted, each Tx antenna ceases transmission for an interval of  $L$  symbol durations to avoid interburst interference, where  $L$  denotes the effective channel memory length. All Tx and Rx antennas are assumed to be perfectly synchronized. For simplicity, the signal mapping scheme is always BPSK. The channel coefficients  $h_{n,m}^l$  of every subchannel are normalized to form an equal delay power profile with an average sum power of one, that is,  $E\{|h_{n,m}^l|^2\} = 1/(L+1)$  with  $\sum_{l=0}^L E\{|h_{n,m}^l|^2\} = 1$ . For a fair comparison, for all three Gaussian detection algorithms, 5 iterations are performed, that is, the operations of message updating and message exchanging are repeated 5 times.

**8.2. Theoretical Performance Bound.** Observing the architecture of the JGD, the GJGD, and the GIGD, these three algorithms clearly fall into the class of symbol-by-symbol detectors, as they all try to maximize the global likelihood function w.r.t. individual symbols. Therefore, the symbol-by-symbol MAP detector provides a lower bound for the achievable BER performance in uncoded systems. Here, we use the BCJR algorithm [3] to implement the symbol-by-symbol MAP detector. Certainly, given the BCJR algorithm, no receiver iterations are necessary for uncoded transmission.

**8.3. Performance Comparison.** Figure 8 displays the BER performances of the three Gaussian detection algorithms. As can be seen, the JGD algorithm achieves a BER performance very close to that of the BCJR algorithm. It shows a trivial error floor at high SNRs due to the inaccuracy of (18) and feeding back intrinsic information as a priori information. (In an uncoded system, the factor graph for the JGD is cycle-free, c.f. Figure 5. Therefore, a self-feedback is enforced at all symbol nodes in order to implement an iterative detection.

The JGD algorithm for uncoded systems in fact falls into the class of probabilistic data association (PDA) algorithms [16]. Nevertheless, it is not necessary and also not proper to do so in a coded system, since the existence of code nodes enables rigorous extrinsic information exchange.) Compared to the JGD, the GJGD algorithm shows a performance loss of approximately 1 dB at  $\text{BER} = 10^{-4}$ . Due to the further inaccuracy introduced by (24), the error floor of the GJGD is higher than that of the JGD and is no longer trivial. The performance of the GIGD algorithm is undesirable in this scenario. It shows a significant error floor at  $\text{BER} \approx 5 \times 10^{-4}$  due to the coarseness of the approximation given in (32).

**8.4. Complexity Comparison.** As a matter of fact, the introduced three Gaussian detection algorithms do not really differ in the necessary number of iterations. Though applying different type of approximations, these algorithms never change the amount of channel outputs ( $y_n[k]$ ) that a symbol node can extract information from. Consequently, the speed of information aggregation does not change for these three algorithms, and the required number of iterations for a satisfactory BER performance basically stays constant for a fixed system setup. Given a reasonable burst length, 5 iterations are already good enough, empirically.

For the current system setup, the covariance matrices to be inverted are of order  $N_R(L+1) = 20$  in the JGD algorithm. The covariance matrices are only of order  $N_R = 4$  in the GJGD algorithm. Finally, matrix inversion is completely eliminated in the GIGD algorithm. Revisiting (22), (30), and (44), we will find that the complexity of the GJGD is about 25 times lower than that of the JGD, and the complexity of GIGD is about 16 times lower than that of the GJGD. In total, a complexity reduction of factor 400 is achieved by the GIGD algorithm w.r.t. the JGD algorithm. As such a complexity reduction is rather attractive, it is worthwhile to study the error floor behavior of the GIGD algorithm.

**8.5. Error Floor of the GIGD.** The error floor of the GIGD algorithm is mainly caused by approximating the elements of the effective noise matrix to be mutually independent. As mentioned in Section 3.2, the average correlation coefficient between effective noise samples drops when the product  $N_T(L+1)$  increases. Therefore, we may expect the error floor of the GIGD to drop when the channel memory length becomes larger or when the system deploys more antennas. To verify our conjecture, we again utilize Monte Carlo simulations.

Figure 9 demonstrates the behavior of the GIGD under different channel memory lengths. Since the complexity of the BCJR algorithm and the JGD algorithm both become prohibitive for severely delay spread MIMO channels, we use the BER bound of an AWGN channel as an asymptotic performance bound if  $L$  approaches infinity. As predicted, the error floor drops as the channel memory length increases and/or the number of antennas increases. This observation reveals two issues. First, the independent Gaussian approximation (32) benefits from a large amount of channel taps. Second, despite its extremely low complexity achieved by

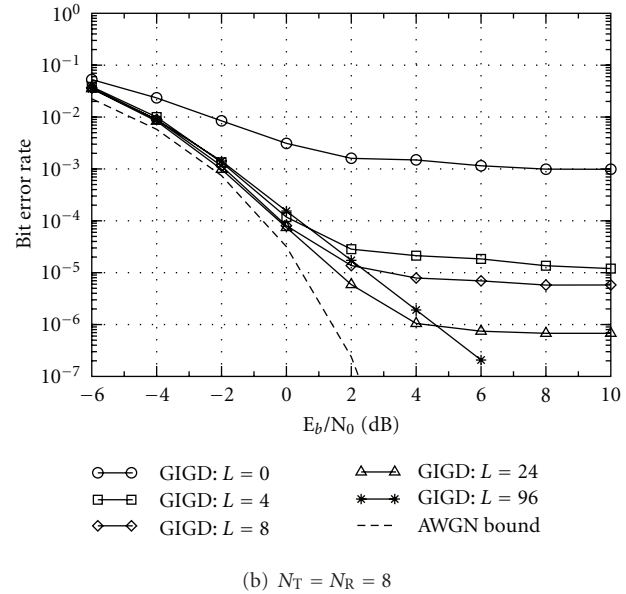
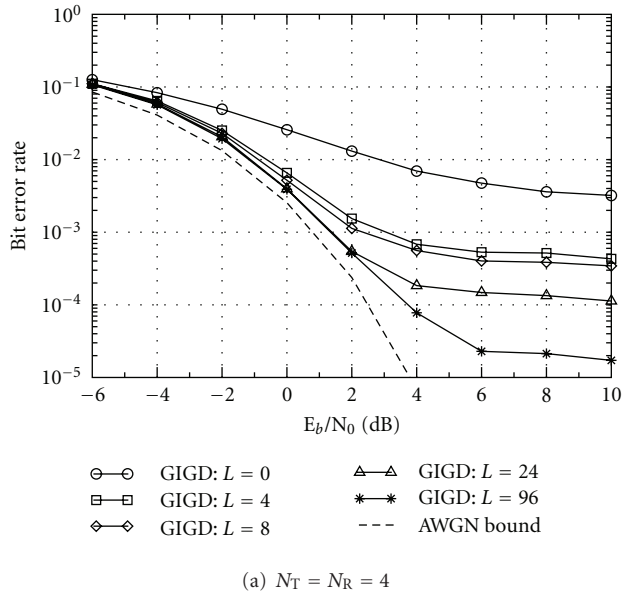


FIGURE 9: Performance of the GIGD in uncoded systems,  $K = 400$ , 5 iterations.

making a very coarse approximation, the GIGD is able to exploit the diversity provided by additional channel taps or receive antennas. The cross-over at  $N_T = N_R = 8$  and  $L = 96$  is mainly caused by the zero-padding burst structure. Both at the beginning and the end of the burst, the channel outputs are composed of few data symbols and a lot of zeros, which degrades the accuracy of the independent Gaussian approximation. This effect is significant at  $L = 96$ , given the burst length is  $K = 400$ . By applying a tail-biting burst structure or a cyclic prefix, this problem can be well eliminated, and the resulting performance will be very close to the AWGN bound.

The above results suggest that the GIGD algorithm is very attractive for large systems with severe delay spread. Nevertheless, the GIGD causes a significant error floor when  $N_T(L + 1)$  is not sufficiently large. The question that remains is if this error floor can be eliminated by means of channel coding.

## 9. Performance in Coded Systems

In this section, we check the BER performance of the three Gaussian detectors in coded systems.

**9.1. Simulation Setup.** For simplicity and for an easy derivation of performance bounds, we adopt repetition encoding with scrambling. The scrambling pattern is fixed, that is, every second bit of a code word is flipped. In case of short data bursts, scrambling is very helpful for the three Gaussian detectors, since they assume that all data symbols come with zero mean. Random interleaving is applied after scrambling in order to make neighboring data symbols as independent as possible. No matter which coding rate is used, the number of symbols per burst per antenna is always  $K = 400$ . Due to the

presence of channel decoding, local iterations in the graph of Gaussian detection are no longer desirable, particularly for the case of JGD. Hence, each receiver iteration contains the following sequential operations: message updating at observation nodes, message updating at symbol nodes, channel decoding, and message updating at symbol nodes. As the use of different Gaussian approximations does not really change the speed of information aggregation at symbol nodes, in the following we will always apply a fixed number of iterations for comparing the performance of using different Gaussian approximations.

**9.2. Performance Comparison.** Figure 10(a) illustrates the performance of the three Gaussian detectors in a rate 1/2 repetition encoded system. Surprisingly, all three Gaussian detectors as well as the BCJR algorithm show nearly the same performance at  $L = 4$ , regardless of their huge complexity difference. A purely theoretical analysis of this phenomenon appears difficult. An empirical answer is that the strongest detector for an uncoded system is not necessarily the best one for a coded system. From the JGD to the GJGD, and from the GJGD to the GIGD, more and more coarse approximations are made, which makes the detector outputs less and less accurate. However, this also makes the detector outputs less and less correlated, which is beneficial to the following channel decoder. From Figure 10(a), it seems that the effect of less accuracy is partially compensated by the effect of less correlation. Figure 10(b) further supports our supposition on this issue. In a rate 1/4 repetition encoded system, the performance of the BCJR algorithm is even worse than that of the three Gaussian detectors at  $L = 4$ . When the coding rate drops, the strong correlation of the outputs of the BCJR algorithm noticeably degrades the system performance, while the three Gaussian detectors

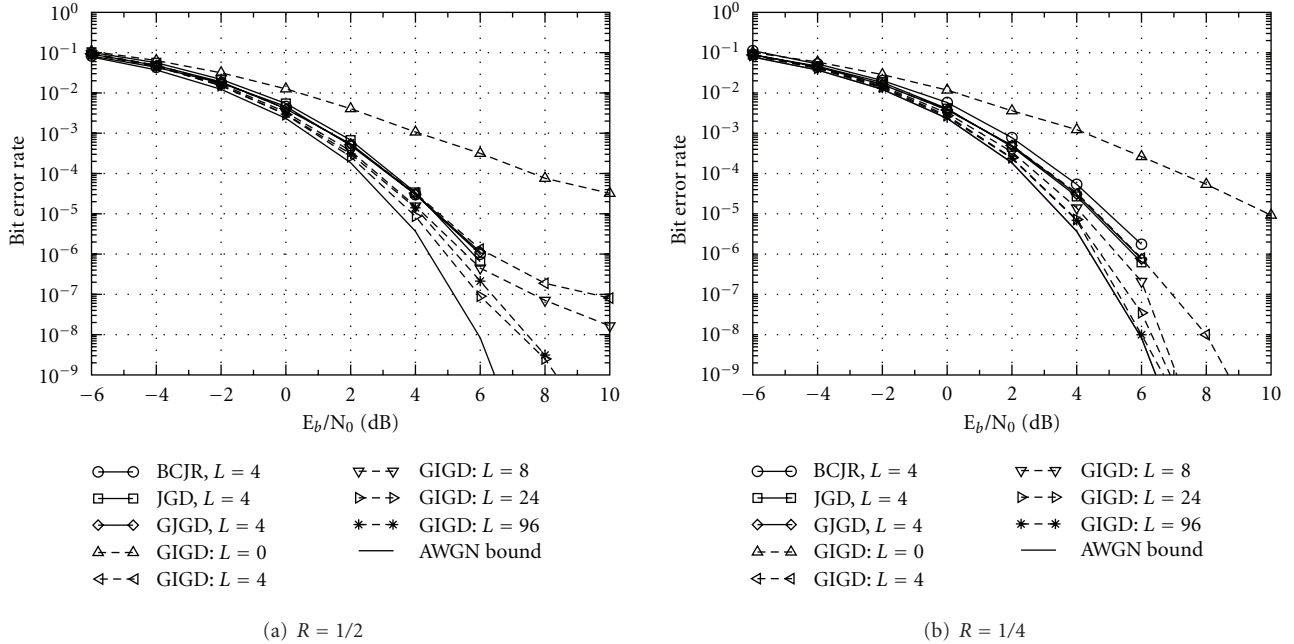


FIGURE 10: Performance comparison in repetition encoded systems,  $N_T = N_R = 4$ ,  $K = 400$ , 5 iterations.

stay robust. Among the four algorithms, the GIGD has a decisively lower complexity, meanwhile its BER performance is not worse than that of any other. Therefore, it is the most attractive solution.

**9.3. Error Floor of the GIGD.** Figures 10(a) and 10(b) also demonstrate the BER performance of the GIGD in repetition encoded systems with various channel memory lengths. Since a repetition code does not provide any coding gain, the AWGN bound still holds. At  $R = 1/2$ , error floors are still present, but are no longer significant. At  $R = 1/4$ , error floors nearly disappear, even for  $L = 0$ , that is, flat-fading channels. The reason of the cross-over at  $L = 96$  and  $R = 1/2$  is still the zero-padding burst structure. Nevertheless, this effect is well mitigated by the rate 1/4 repetition code. So far we may recognize that repetition encoding is really helpful in mitigating the estimation errors caused by the independent Gaussian approximation, and meanwhile the approximation errors do not present a problem to the convergence property of the repetition decoder. The asymptotic AWGN bound is quasi-approached at  $N_T = N_R = 8$ ,  $L = 96$ . Note that with this system setup, it is practically impossible to run the BCJR algorithm and it is computationally prohibitive to run the JGD. For systems with short memory lengths, repetition encoding is truly helpful for the GIGD. Necessary to be mentioned, the AWGN bound is only achievable for systems with very large channel memory lengths, since only then the channel instant power tends to be constant. By checking the performance of GIGD with small  $L$  values, we may recognize that these curves should also be quasi-bound approaching. Therefore, in repetition encoded systems, the GIGD is applicable for systems with moderate number of antennas and short channel memory lengths as well.

## 10. Conclusions and Future Work

In this paper, we revisited and slightly revised the joint Gaussian detection (JGD) algorithm, derived the grouped joint Gaussian detection (GJGD) algorithm, and proposed the graph-based iterative Gaussian detection (GIGD) algorithm. A mathematical derivation as well as a detailed performance analysis is provided. From the JGD to the GJGD and from the GJGD to the GIGD, the computational complexity dramatically decreases. The GIGD algorithm has a linear complexity and provides a promising performance for MIMO channels with severe delay spread. In [17], the incorporation of the GIGD algorithm with soft channel estimation has been studied.

The adopted channel model within this paper is very specific, in the sense that it presents the biggest challenge for conventional linear equalizers. Using such a channel model effectively exhibits the high potential of the proposed low-complexity Gaussian detection algorithms, particularly GIGD. Nevertheless, from an engineering standpoint, it deserves to be an interesting topic to test the performance of Gaussian detection with more realistic channel models. Repetition coding is considered within this paper for the sake of easy analysis as well as its strength in mitigating estimation errors due to approximation. Future work should also be targeted at more advanced code structures, particularly concatenations of a repetition code and a sparse graph code, for example, an LDPC code.

## Acknowledgments

The authors would like to thank Shan Jiang and Ying Yu for their effort on this topic during their master theses work.

This work has been supported by the German Research Foundation (DFG) under Contract nos. HO 2226/8-1 and HO 2226/10-1.

## References

- [1] G. D. Forney Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Transactions on Information Theory*, vol. 18, no. 3, pp. 363–378, 1972.
- [2] G. Ungerboeck, "Adaptive maximum-likelihood receiver for carrier-modulated data-transmission systems," *IEEE Transactions on Communications*, vol. 22, no. 5, pp. 624–636, 1974.
- [3] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, 1974.
- [4] A. Bahai, B. Saltzberg, and M. Ergen, *Multi Carrier Digital Communications: Theory and Applications of OFDM*, Springer, New York, NY, USA, 2004.
- [5] X. Wang and H. Vincent Poor, "Iterative (Turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Transactions on Communications*, vol. 47, no. 7, pp. 1046–1061, 1999.
- [6] S. Liu and Z. Tian, "Near-optimum soft decision equalization for frequency selective MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 3, pp. 721–733, 2004.
- [7] Y. Jia, C. Andrieu, R. J. Piechocki, and M. Sandell, "Gaussian approximation based mixture reduction for near optimum detection in MIMO systems," *IEEE Communications Letters*, vol. 9, no. 11, pp. 997–999, 2005.
- [8] X. Yuan, K. Wu, and L. Ping, "The jointly Gaussian approach to iterative detection in MIMO systems," in *Proceedings of the IEEE International Conference on Communications (ICC '06)*, Istanbul, Turkey, September 2006.
- [9] P. H. Tan and L. K. Rasmussen, "Asymptotically optimal nonlinear MMSE multiuser detection based on multivariate Gaussian approximation," *IEEE Transactions on Communications*, vol. 54, no. 8, pp. 1427–1438, 2006.
- [10] Y. Jia, C. Andrieu, R. J. Piechocki, and M. Sandell, "Gaussian approximation based mixture reduction for joint channel estimation and detection in MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2384–2389, 2007.
- [11] D. Divsalar, M. K. Simon, and D. Raphaeli, "Improved parallel interference cancellation for CDMA," *IEEE Transactions on Communications*, vol. 46, no. 2, pp. 258–268, 1998.
- [12] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.
- [13] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 28–41, 2004.
- [14] T. Wo and P. A. Hoeher, "A simple iterative Gaussian detector for severely delay-spread MIMO channels," in *Proceedings of the IEEE International Conference on Communications (ICC '07)*, pp. 24–28, Glasgow, Scotland, June 2007.
- [15] T. Wo, J. C. Fricke, and P. A. Hoeher, "A graph-based iterative Gaussian detector for frequency-selective MIMO channels," in *Proceedings of the IEEE Information Theory Workshop (ITW '06)*, Chengdu, China, October 2006.
- [16] J. Ch. Fricke, M. Sandell, J. Mietzner, and P. A. Hoeher, "Impact of the Gaussian approximation on the performance of the probabilistic data association MIMO decoder," *EURASIP Journal on Wireless Communications and Networking*, vol. 2005, no. 5, pp. 796–800, 2005.
- [17] T. Wo, C. Liu, and P. A. Hoeher, "Graph-based soft channel and data estimation for MIMO systems with asymmetric LDPC codes," in *Proceedings of the IEEE International Conference on Communications (ICC '08)*, pp. 620–624, Beijing, China, May 2008.





**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

