Graph-Based Iterative Gaussian Detection with Soft Channel Estimation for MIMO Systems

Tianbin Wo, Chunhui Liu†, and Peter Adam Hoeher
Information and Coding Theory Lab, Faculty of Engineering, University of Kiel, Germany

Abstract

Conventionally, the uncertainties of channel coefficients are neglected, that is the estimated values of channel coefficients are taken as the true values in the stage of data detection. In the communications community, it is still an open question how to take into account the channel uncertainty for data detection/decoding, especially in a low-complexity manner. In this paper, we propose a low-complexity receiver algorithm which utilizes soft channel information. Channel coefficients are treated as variables and estimated in an element-wise manner. Their uncertainties are represented by the variances. Instead of performing channel estimation and data detection in a separate manner, this algorithm does everything in one stage, i.e., channel estimation and data detection/decoding are carried out simultaneously over a general factor graph. The feasibility of this algorithm is verified by means of Monte-Carlo simulations both in bit error ratio (BER) and channel estimation mean squared error (MSE).

1 Introduction

For reliable data communication over the air interface, channel estimation is a task that we can never avoid. However, due to the existence of additive noise either from the front-end circuit or the surrounding environment, it is practically not possible for us to obtain the exact channel state information. Uncertainty will exist with the channel estimates as long as disturbances exist. Nevertheless, channel uncertainty is often neglected in conventional receivers. One reason is that channel coefficients are continuously valued, and it is not feasible to express their uncertainties by common metrics of reliability, such as log-likelihood ratio. Another reason lies in the fact that channel estimation and data detection are carried out in different stages in conventional receivers, which makes it troublesome to exchange soft information. However, channel uncertainty remains to be a critical issue for systems without accurate channel knowledge, such as systems with fast-fading channels and time-division duplexing (TDD) systems with channel evaluation only performed for the uplink.

Multi-input multi-output (MIMO) systems are today being widely deployed for wideband data communications. MIMO channels show dramatic capacity gain w.r.t. single-input single-output (SISO) channels in an environment rich of scattering [1]-[3]. On the other hand, an MIMO receiver is generally more sensitive to the channel estimation errors than an SISO one, due to multi-antenna interferences (MAI). Traditionally, people increase the power or the amount of time slots allocated to training (pilot) symbols to improve the channel estimation quality. Doing so however will degrade either the power efficiency or the bandwidth efficiency of the system. Hassibi *et al.* showed in [4] that pure training-based channel estimation can be highly suboptimal from the information theoretic point of view. In comparison, semiblind channel estimation (SBCE) tries to extract the channel state information carried by all observations, and is able to achieve very low mean squared error (MSE) with using just a few training symbols. Simulation results of SBCE can be found in a number of papers [5]-[9]. Even though, the uncertainty of channel estimates are all neglected by available SBCE algorithms, and moreover hard decisions of data symbols are used for iterative channel estimation. Therefore, it deserves to be an interesting topic to design a receiver which utilizes soft channel information for data detection and adopts soft data information during channel estimation.

In this paper, we will try to find a suitable metric to carry the reliability of a channel estimate, and furthemore utilize this reliability information in the stage of data detection. A graph-based iterative algorithm is proposed for the task of joint channel estimation and data detection in MIMO systems. Based on a general factor graph, channel estimation and data detection are carried out in the same stage. A characteristic feature of this algorithm is that channel coefficients are estimated element-wise instead of being jointly estimated by means of matrix inversion. This feature makes the algorithm a good prototype for systems whose channel coefficients are difficult to be jointly estimated, such as moving systems with very fast-fading channels. Gaussian approximation is applied both for channel estimation and data detection, which guarantees the algorithm with a complexity strictly linear in all system parameters, including the number of transmit antennas and the number of receive antennas.

The remainder of this paper is organized as follows. Sec. 2 introduces a flat-fading MIMO channel model and its corresponding factor graph. In Sec. 3, the concepts of soft channel estimation and Gaussian approximation are briefly outlined. Sec. 4 describes the proposed algorithm in details, and Sec. 5 verifies the feasibility of the proposed algorithm via simulation results. Finally, Sec. 6 concludes the paper.

This work has been supported by the German Research Foundation (DFG) under contract no. HO 2226/8-1.

†Mr. Chunhui Liu is now a Ph.D. student at RWTH Aachen University, Germany.

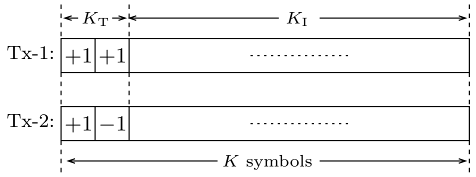

Figure 1: Burst structure

2 System Model

2.1 Channel Model

Let N_R denote the number of receive (Rx) antennas and N_T the number of transmit (Tx) antennas, the equivalent discrete-time model of an MIMO channel (including transmit and receive filter, physical channel and baud-rate sampling) is given by

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

where $k \in \{0, 1, \ldots, K-1\}$ is the discrete time index with K denoting the burst length. $y_n[k] \in \mathbb{C}$ is the channel output at the n-th ($1 \leq n \leq N_R$) Rx antenna at time index k, and $x_m[k] \in \{\pm 1\}$[1] is the channel input at the m-th Tx antenna at time index k. $h_{n,m}[k] \in \mathbb{C}$ marks the coefficient of the sub-channel connecting the n-th Rx antenna and the m-th Tx antenna at time index k. $w_n[k]$ represents an additive white Gaussian noise (AWGN) sample with zero mean and variance σ_w^2.

For the sake of simplicity, block fading is assumed throughout this paper, that is, all channel coefficients keep constant within each data burst while vary independently from burst to burst. We also assume that K_{T} training symbols are transmitted per burst per Tx antenna, as illustrated in Fig. 1, where K_{I} denotes the amount of data symbols per burst.

2.2 Factor Graph

A factor graph [10] is a bipartite graph visualizing the factorization of certain global functions subject to minimization or maximization. It is often helpful in the design of low-complexity iterative processing algorithms. Admitting the fact that uncertainties exist in channel coefficients, a general factor graph of an MIMO channel should include channel coefficients as variables as well. Making the following independence approximation[2]

$$p(\mathbf{Y} \,|\, x_m[k]) \approx \prod_{n=1}^{N_R} p(y_n[k] \,|\, x_m[k]) \ , \quad (2)$$

where $\mathbf{Y}$ is the matrix which collects all channel outputs of the current data burst, a factor graph of an MIMO channel will look like Fig. 2. The mark $\boxdot$ stands for the relationship between $h_{n,m}[k]$ and $h_{n,m}[k+1]$.

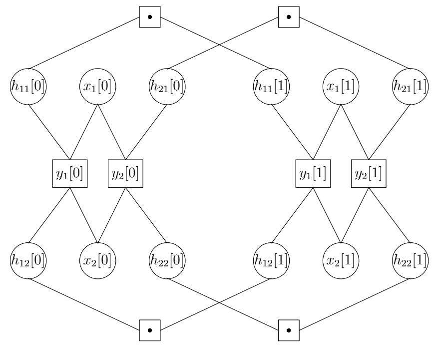

Figure 2: Factor graph of a block-fading MIMO channel, $N_T = N_R = 2$

Since we assume that the channel is constant within each data burst, the relationship between $h_{n,m}[k]$ and $h_{n,m}[k+1]$ is simply an equality. Nevertheless, for fast fading channels, this $\boxdot$ should be a suitable transfer function which describes the degree of variation of a channel coefficient between two neighboring time indices. Due to limited space, details on this topic will not be elaborated in this paper.

3 Preliminary Remarks

3.1 Soft Channel Estimation

Consider a real-valued SISO flat-fading channel model:

$$y = h \cdot x + w, \quad (3)$$

where w is a zero-mean AWGN sample with variance σ_w^2. If symbol x is known at the receiver side, a least-squares channel estimator (LSCE) calculates

$$\hat{h} = y \cdot x^{-1} \ , \quad (4)$$

and passes $\hat{h}$ to the data detector as the estimated value of the channel coefficient. LSCE is widely applied in modern communication systems due to its low complexity and desirable performance. Nevertheless, $\hat{h}$ in fact does not fully represent the information of h contained in y. Connecting (3) and (4), we can easily find that

$$h = \hat{h} - w \cdot x^{-1}. \quad (5)$$

Since x is known, the product $w \cdot x^{-1}$ is a zero-mean Gaussian variable with variance $\sigma_w^2/|x|^2$. Hence, the complete information of h which we can extract from y is that, h is a Gaussian distributed variable with mean $\mu_h = \hat{h}$ and variance $\sigma_h^2 = \sigma_w^2/|x|^2$, as illustrated in Fig. 3. In this case, we may say that σ_h^2 carries the reliability information of $\hat{h}$. Certainly, the smaller σ_h^2 is, the more reliable $\hat{h}$ is.

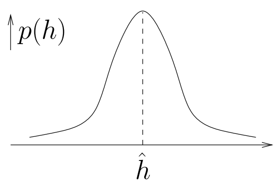

Figure 3: Probability density function of h

[1]BPSK mapping is assumed throughout the paper for the sake of simplicity. The extension of the proposed algorithm to higher-order modulation formats can be done in a straightforward manner.

[2]This approximation is made to reduce the data detection complexity. The authors would refer interested readers to [11], [12] for detailed explanation on this issue.

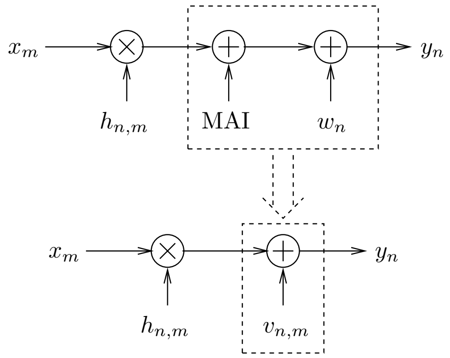

Figure 4: Relationship between a data symbol and a channel output

3.2 Gaussian Approximation

Considering an arbitrary data symbol x_m[3] and one of its channel outputs y_n, we have the following equation

$$\begin{aligned} y_n &= \sum_{i=1}^{N_T} h_{n,i} \cdot x_i + w_n \\ &= h_{n,m} \cdot x_m + \underbrace{\sum_{i=1, i \neq m}^{N_T} h_{n,i} \cdot x_i}_{\text{MAI}} + \underbrace{w_n}_{\text{AWGN}} \ , \end{aligned} \quad (6)$$

where MAI stands for multi-antenna interference. We define $v_{n,m}$ as the effective noise sample in the observation y_n w.r.t. the symbol x_m:

$$v_{n,m} \doteq y_n - h_{n,m} \cdot x_m \ , \quad (7)$$

as illustrated in Fig. 4. If the probability density function (PDF) of $v_{n,m}$ can be approximated by a complex Gaussian function:

$$p(v_{n,m}) \approx \frac{1}{\pi \sigma^2_{v_{n,m}}} \exp\Big(- \frac{|v_{n,m} - \mu_{v_{n,m}}|^2}{\sigma^2_{v_{n,m}}} \Big) \ , \quad (8)$$

with

$$\begin{aligned} \mu_{v_{n,m}} &\doteq \mathrm{E}\{v_{n,m}\} \\ \sigma^2_{v_{n,m}} &\doteq \mathrm{E}\left\{|v_{n,m} - \mu_{v_{n,m}}|^2\right\} \ , \end{aligned} \quad (9)$$

the computation of likelihood function will be dramatically simplified into

$$p(y_n \,|\, x_m) \approx \frac{1}{\pi \sigma^2_{v_{n,m}}} \exp\Big(- \frac{|y_n - h_{n,m} \, x_m - \mu_{v_{n,m}}|^2}{\sigma^2_{v_{n,m}}} \Big) \ . \quad (10)$$

Hereafter, we will refer to a data detector using the above Gaussian approximation as a Gaussian detector. If all channel coefficients are exactly known, this approximation is indeed not accurate enough at high signal-to-noise ratios (SNRs) due to the discreteness of data symbols. However, if channel coefficients are not exactly known, this approximation works pretty well even at high SNRs. Simulation results will show that the uncertainty of channel coefficients is in fact beneficial for the accuracy of this Gaussian approximation.

[3] In this section, time index k is omitted for simplicity.

4 Graph-Based Iterative Gaussian Detection with Soft Channel Estimation (GIGD-SCE)

In this section, joint channel estimation and data detection will be carried out in a general factor graph. Soft channel estimation will be applied to improve the system performance, while Gaussian approximation will be applied to reduce the detection complexity. Due to the common relationships between data symbols, channel coefficients, and channel observations, a more abstract notation is adopted for ease of explanation. In the following, we use y to denote an observation node, x to denote a symbol node, and h to denote a channel coefficient node. The complicated indices n, m, and k are in general replaced by a single index i.

4.1 Starting Point

To start the algorithm, we assume that all channel coefficients are zero-mean Gaussian distributed, and all subchannels have the same average power. This assumption is valid for most of the practical applications. Nevertheless, this assumption does not necessarily to be accurate since this initial setup is discarded as soon as the algorithm starts to run. Furthermore, we assume that all data symbols are uniformly distributed within a finite alphabet, which is always true.

4.2 Message Update in Observation Nodes

Revisiting Fig. 2, we will find that the relationship between an observation node and its associated variable nodes can be written as

$$y = \sum_{i=1}^{Q} h_i \cdot x_i + w, \quad (11)$$

where $Q = N_T$ is the amount of associated symbol nodes or channel coefficient nodes, h_i is the channel coefficient linking y and x_i, and w is the additive white Gaussian noise sample. This relationship is visualized in Fig. 5. In each iteration, one observation node will receive messages from its variable nodes in the form of probability functions. Then new messages should be generated and redistributed to these variable nodes according to the Turbo principle, that is only extrinsic information should be exchanged. A sketch diagram of message propagation is given in Fig. 6.

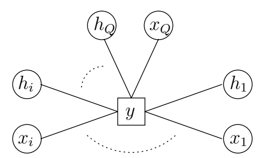

Figure 5: An observation node and its variable nodes

4.2.1 Data Detection with Soft Channel Information

Based on (11), we define v_m ($1 \leqslant m \leqslant Q$) as the effective noise sample in the observation y w.r.t. the

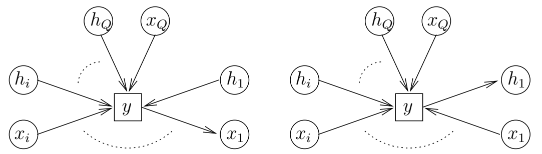

Figure 6: Message exchange at an observation node

symbol x_m:

$$v_m \doteq y - h_m \cdot x_m = \sum_{i=1, i\neq m}^{Q} h_i \cdot x_i + w \ . \tag{12}$$

Approximating $p(v_m)$ as $v_m \sim \mathcal{CN}(\mu_{v_m}, \sigma^2_{v_m})$, and knowing that $h_m \sim \mathcal{CN}(\mu_{h_m}, \sigma^2_{h_m})$, the log-likelihood ratio (LLR) of symbol x_m should be calculated as

$$\mathrm{LLR}(x_m) = 4\ \mathrm{Re}\{\mu^*_{h_m}(y - \mu_{v_m})\}/(\sigma^2_{h_m} + \sigma^2_{v_m}) \ . \tag{13}$$

In each iteration, LLR of each symbol is calculated according to (13) and distributed over the factor graph. For the readability of the main body, the mathematical derivation of (13) is provided in Appendix.

To obtain the statistics of the effective noise sample, we define

$$\xi_i \doteq h_i \cdot x_i \ , \tag{14}$$

so that

$$v_m = \sum_{i=1, i\neq m}^{Q} \xi_i + w \ . \tag{15}$$

For short-hand notation, we also define

$$\begin{aligned} \mathrm{P}(x_i = +1) &= \mathrm{P}_{i,+1} \\ \mathrm{P}(x_i = -1) &= \mathrm{P}_{i,-1} \ , \end{aligned} \tag{16}$$

which can be obtained from the LLR message of each data symbol in a straightforward manner. Noting that h_i and x_i are statistically independent, the PDF of ξ_i can be written as

$$p(\xi_i) = \mathrm{P}_{i,+1} \cdot p(h_i = \xi_i) + \mathrm{P}_{i,-1} \cdot p(h_i = -\xi_i) \ , \tag{17}$$

which is indeed a mixed Gaussian function with two peaks. Applying $h_i \sim \mathcal{CN}(\mu_{h_i}, \sigma^2_{h_i})$ and after some mathematical derivation, we will obtain

$$\begin{aligned} \mu_{\xi_i} &= \mu_{h_i} \cdot (\mathrm{P}_{i,+1} - \mathrm{P}_{i,-1}) \\ \sigma^2_{\xi_i} &= \sigma^2_{h_i} + 4\ \mathrm{P}_{i,+1}\mathrm{P}_{i,-1}|\mu_{h_i}|^2 \ . \end{aligned} \tag{18}$$

Finally, the mean and variance of v_m can be calculated as

$$\begin{aligned} \mu_{v_m} &= \sum_{i=1, i\neq m}^{Q} \mu_{\xi_i} \\ \sigma^2_{v_m} &= \sum_{i=1, i\neq m}^{Q} \sigma^2_{\xi_i} + \sigma^2_w \ . \end{aligned} \tag{19}$$

As a matter of fact, the effective noise sample v_m is a summation of $Q-1$ independent mixed Gaussian variables plus an independent Gaussian variable. Since all these component variables are continuously valued, the Gaussian approximation of v_m works pretty well even at high SNRs.

4.2.2 Channel Estimation with Soft Data Decision

Let us rewrite (12) into

$$y = h_m \cdot x_m + v_m \ . \tag{20}$$

The information of h_m contained in y is fully represented by the conditional probability density function $p(y|h_m)$, which may be computed as follows:

$$\begin{aligned} p(y|h_m) &= \sum_{x_m \in \{\pm 1\}} p(y|h_m, x_m)\mathrm{P}(x_m) \\ &= \mathrm{P}_{m,+1}\frac{1}{\pi\sigma^2_{v_m}} \exp\left(-\frac{|h_m - (y - \mu_{v_m})|}{\sigma^2_{v_m}}\right) + \\ &\quad \mathrm{P}_{m,-1}\frac{1}{\pi\sigma^2_{v_m}} \exp\left(-\frac{|h_m + (y - \mu_{v_m})|}{\sigma^2_{v_m}}\right) \ . \end{aligned} \tag{21}$$

Excluding a priori information, we have

$$p(h_m) = p(h_m|y) \propto p(y|h_m) \ , \tag{22}$$

and after considering the issue of normalization, the following statement holds:

$$p(h_m) = p(y|h_m) \ . \tag{23}$$

Clearly, it is again a mixed Gaussian function, which is troublesome to be utilized in the stage of data detection, particularly when the data symbol x_m is of higher-order modulation formats other than BPSK. Therefore, suitable approximation is necessary to simplify this channel knowledge. Note that if the data detection is carried out successfully, we will have

$$\mathrm{P}_{m,+1} \gg \mathrm{P}_{m,-1} \tag{24}$$

or

$$\mathrm{P}_{m,+1} \ll \mathrm{P}_{m,-1} \tag{25}$$

as the iteration goes on. It is reasonable to make the approximation $h_m \sim \mathcal{CN}(\mu_{h_m}, \sigma^2_{h_m})$ with

$$\begin{aligned} \mu_{h_m} &= (y - \mu_{v_m})(\mathrm{P}_{m,+1} - \mathrm{P}_{m,-1}) \\ \sigma^2_{h_m} &= \sigma^2_{v_m} + 4\mathrm{P}_{m,+1}\mathrm{P}_{m,-1}|y - \mu_{v_m}|^2 \ , \end{aligned} \tag{26}$$

which are calculated according to (21) and (23).

4.3 Message Update in Symbol Nodes

Observing Fig. 2, it is clear that each symbol node is connected with $Q = N_R$ observation nodes. This relationship is concisely depicted in Fig. 7. The message

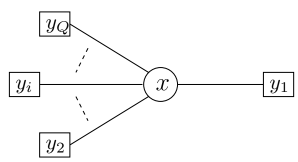

Figure 7: A symbol node and its observation nodes

update rule in a symbol node is very simple and in fact nothing more than linear additions. The underlying principles are that LLR messages from independent observations are additive and only the extrinsic information should be propagated in a factor graph. For example, if a symbol node receives LLR messages from its observation nodes as shown in the left part of Fig. 8, then the updated LLR messages are generated just by linear additions as shown in the right part of Fig. 8.

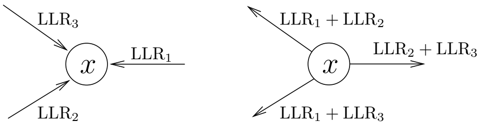

Figure 8: Message exchange at a symbol node

4.4 Message Update in Coefficient Nodes

For block-fading channels, the channel coefficients keep constant within each data burst. Therefore, each channel coefficient node is associated with $Q = K$ observation nodes as depicted in Fig. 9. The message

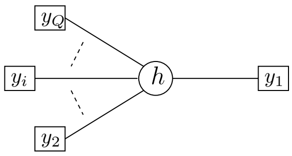

Figure 9: A channel coefficient node and its observation nodes

update rule in a channel coefficient node is very similar with that in a symbol node. The only difference is that we replace the linear addition by a product. For example, if a channel coefficient node receives PDF messages from its observation nodes as shown in the left part of Fig. 10, then the updated messages are generated in a way shown in the right part of Fig. 10. Note that the product of two Gaussian PDFs gives a

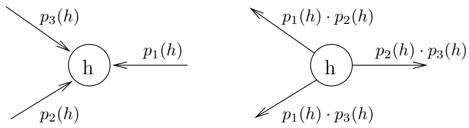

Figure 10: Message exchange at a channel coefficient node

new Gaussian PDF. Suppose we have the following two messages:

$$\begin{aligned} p_1(h) &\quad : \quad h \sim \mathcal{CN}(\mu_1, \sigma_1^2) \\ p_2(h) &\quad : \quad h \sim \mathcal{CN}(\mu_2, \sigma_2^2)\ , \end{aligned} \tag{27}$$

then the product of these two messages will be given by

$$p_1(h) \cdot p_2(h) \quad : \quad h \sim \mathcal{CN}(\mu_h, \sigma_h^2) \tag{28}$$

with

$$\begin{aligned} \mu_h &= (\sigma_2^2 \mu_1 + \sigma_1^2 \mu_2)/(\sigma_1^2 + \sigma_2^2) \\ \sigma_h^2 &= \sigma_1^2 \sigma_2^2/(\sigma_1^2 + \sigma_2^2)\ . \end{aligned} \tag{29}$$

4.5 Scheduling

Before the first iteration, the observation nodes connected with training symbols update their messages to channel coefficient nodes in order to provide a reasonable starting point for the iterative processing algorithm. Afterwards in each iteration, the message updating operation is performed once per observation node, per symbol node, and per channel coefficient node. Using the same analysis as in [11], we will easily find that the complexity of this algorithm is strictly linear in the number of Tx antennas and Rx antennas.

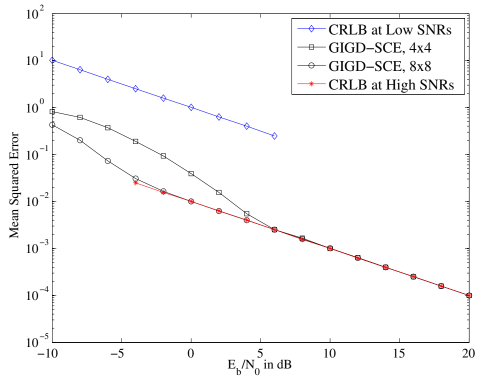

Figure 11: MSE vs. E_b/N_0, $K_\mathrm{T} = N_T$, $K_\mathrm{I} = 100$, $R = 1/4$, 5 iterations

5 Numerical Results

5.1 Simulation Setup

For the numerical results provided in this section, a block-Rayleigh-fading MIMO channel model is used. The coefficient of each subchannel is normalized to have an average power of 1. The SNR per bit E_b/N_0 is calculated as $1/\sigma_w^2$, where σ_w^2 denotes the variance of the additive noise.

We adopt a repetition code for the sake of simplicity, and channel coding is done separately for each transmit antenna. Coding rate is set to be $1/4$. Scrambling with fixed pattern is applied, that is, every second bit of a code word is flipped. In case of short data bursts, scrambling is mandatory for GIGD-SCE, since it assumes 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.

5.2 Mean Squared Error

The theoretical tight bound of the MSE of unbiased semi-blind channel estimation (SBCE) is given by the Cramer-Rao lower bound (CRLB) [8], [9]. Detailed explanation of CRLB is out of the scope of this paper. Nevertheless, the CRLB for SBCE at high SNRs can be understood in a quite intuitive way. The SBCE-CRLB at high SNRs is given by the MSE of a least-squares channel estimator which has full knowledge of all data symbols and uses the whole burst to perform channel estimation.

Fig. 11 gives a comparison between the MSE performances of GIGD-SCE with the corresponding Cramer-Rao lower bound. Here, $n \times n$ stands for a system with n transmit antennas and n receive antennas. GIGD-SCE demonstrates a superior performance in the sense of channel estimation MSE. It approaches the CRLB at high SNRs from rather early points. For a 4×4 system, it approaches CRLB at SNR ≈ 6 dB, while for an 8×8 system, it approaches CRLB at SNR ≈ -2 dB. Necessary to be mentioned, GIGD-SCE outputs the mean value and variance of each channel coefficient. Here in Fig. 11, the mean values are taken as hard channel estimates for the calculation of channel estimation MSE. Unlike an LSCE or MMSE channel

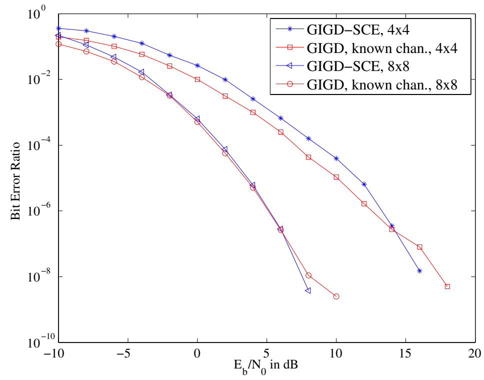

Figure 12: BER vs. E_b/N_0, $K_{\mathrm{T}} = N_T$, $K_{\mathrm{I}} = 100$, $R = 1/4$, 5 iterations

estimator, GIGD-SCE performs no matrix inversion at all. It evaluates channel coefficients one by one with the help of Gaussian approximation. Nevertheless, it works pretty well as shown by the results.

At middle-range SNRs, GIGD-SCE shows a non-linear behaviour just as all available semi-blind channel estimation algorithms. At low SNRs, an interesting phenomenon is that the MSE curves of GIGD-SCE are even lower than the CRLB. This however tells us that GIGD-SCE is a biased channel estimator at low SNRs. For details on these two issues, the authors would refer interested readers to [8], [9].

5.3 Bit Error Ratio

In this section, we check the BER performances of GIGD-SCE in a 4×4 system and an 8×8 system. As a comparison, we take the BER performances of graph-based iterative Gaussian detector (GIGD) with known channel coefficients. The algorithm of GIGD can be simply obtained by removing the channel estimation part of GIGD-SCE. Interested readers may refer to [11] for more direct descriptions of GIGD.

Fig. 12 demonstrates the BER performances of GIGD-SCE in repetition coded MIMO systems. As we can see, the BER performances of GIGD-SCE are desirable. The gap between GIGD-SCE and GIGD with known channel is less than 2 dB in a 4×4 system. Such a performance is already very good, since only 4 training symbols are inserted into each burst in this case and we have 4 transmit antennas. When the system deploys more antennas, the situation gets even better. The gap between GIGD-SCE and GIGD with known channel nearly vanishes in an 8×8 system. This excellent BER performance in an 8×8 system can be well explained by Fig. 11, as the quality of channel estimation is optimal for SNR $\gtrapprox -2$ dB.

A very interesting phenomenon is that GIGD-SCE outperforms GIGD with known channel at high SNRs, which is contradictory to our preconceived ideas. This is due to the fact that the Gaussian approximation of MAI is still valid at high SNRs with uncertain channel coefficients. Nevertheless, if all channel coefficients are exactly known, the summation of several interfering BPSK symbols will be far away from a Gaussian distribution.

6 Conclusions

In this paper, we proposed a graph-based iterative Gaussian detection and soft channel estimation algorithm. This algorithm treats channel coefficients as variables as well as data symbols. Data detection and channel estimation are performed over a general factor graph. Channel coefficients are estimated element-wise with the help of a Gaussian approximation. The feasibility of this algorithm is verified by various numerical results both in channel estimation mean squared error and bit error ratio.

Appendix

$$\begin{aligned}
p(y|x_m) &= \int p(y|h_m, x_m)\ p(h_m)\ \mathrm{d}h_m \\
&= \int \frac{1}{\pi\sigma_{v_m}^2} e^{-\frac{|y-h_m x_m-\mu_{v_m}|^2}{\sigma_{v_m}^2}} \\
&\quad \cdot \frac{1}{\pi\sigma_{h_m}^2} e^{-\frac{|h_m-\mu_{h_m}|^2}{\sigma_{h_m}^2}}\ \mathrm{d}h_m \\
&= \frac{1}{\pi(\sigma_{h_m}^2|x_m|^2+\sigma_{v_m}^2)} e^{-\frac{|y-\mu_{h_m}x_m-\mu_{v_m}|^2}{\sigma_{h_m}^2|x_m|^2+\sigma_{v_m}^2}} .
\end{aligned}$$

For BPSK symbols, $|x_m|^2 \equiv 1$. We have

$$\Rightarrow \mathrm{LLR}(x_m) = 4\ \mathrm{Re}\{\mu_{h_m}^*(y-\mu_{v_m})\}/(\sigma_{h_m}^2+\sigma_{v_m}^2) .$$

References

[1] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” *Euro. Trans. Telecommun.*, vol. 10, pp. 585-595, 1999.

[2] 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, pp. 311-335, 1998.

[3] B. Hochwald and T. L. Marzetta, “Capacity of a mobile multiple-antenna communication link in a Rayleigh flat-fading environment,” *IEEE Trans. Inform. Theory*, vol. 45, pp. 139-157, Jan. 1999.

[4] B. Hassibi and B.M. Hochwald, “How much training is needed in multiple-antenna wireless links?” *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 951-963, Apr. 2003.

[5] C. Cozzo and B.L. Hughes, “Joint channel estimation and data detection in space-time communications,” *IEEE Trans. Commun.*, vol. 51, pp. 1266-1270, Aug. 2003.

[6] M. Loncar, R.M.J. Wehinger, and T. Abe, “Iterative joint detection, decoding, and channel estimation for dual antenna arrays in frequency selective fading,” in *Proc. 5th Int. Symp. Wireless Personal Multimedia Communication*, pp. 125-129, Honolulu, USA, Oct. 2002.

[7] T. Wo and P.A. Hoeher, “Semi-blind channel estimation for frequency-selective MIMO systems,” in *Proc. 14-th IST Mobile & Wireless Communications Summit*, Dresden, Germany, June 2005.

[8] T. Wo, P.A. Hoeher, A. Scherb, and K.-D. Kammeyer, “Performance analysis of maximum-likelihood semiblind estimation of MIMO channels,” in *Proc. IEEE 63rd Vehic. Techn. Conf. (VTC)*, Melbourne, Australia, May 2006.

[9] T. Wo, A. Scherb, P.A. Hoeher, and K.-D. Kammeyer, “Analysis of semiblind channel estimation for FIR-MIMO systems,” in *Proc. 4th Int. Symp. Turbo Codes & Related Topics*, Munich, Germany, Apr. 2006.

[10] F.R. Kschischang, J. Frey, and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 498-519, Feb. 2001.

[11] T. Wo and P.A. Hoeher, “A simple iterative Gaussian detector for severely delay-spread MIMO channels,” in *Proc. Int. Conf. Commun. (ICC)*, Glasgow, Scotland, June 2007.

[12] T. Wo, J.Ch. Fricke, and P.A. Hoeher, “A graph-based iterative Gaussian detector for frequency-selective MIMO channels,” in *Proc. IEEE Inform. Theory Workshop (ITW)*, Chengdu, China, Oct. 2006.