Improving Multi-Dimensional Graph-Based Soft Channel Estimation

Christopher Knievel Peter Adam Hoeher
Information and Coding Theory Lab
Faculty of Engineering, University of Kiel
Kaiserstr. 2, 24143 Kiel, Germany
Email: {chk,ph}@tf.uni-kiel.de

Alexander Tyrrell Gunther Auer
DOCOMO Euro-Labs
80687 Munich, Germany
Email: gunther.auer@ieee.org

***Abstract*—The principles and benefits of soft decisions are well known and widely applied. The advantage of knowing that a decision is reliable or not is obvious. Belief propagation within a factor graph enables a unified use of soft information for both channel estimation and data detection. However, if soft information does not reflect the true reliability of a decision, typically the achievable performance degrades. In this paper, the calculation of reliability information is refined to consider the event of unreliable soft decisions. The proposed solution is based on the average bit error probability calculated after each iteration and integrates nicely within existing factor graphs. Simulation results are provided to illustrate the performance gains.**

I. Introduction

In general, soft values represent the reliability of an estimated value. In case of data detection, soft values are typically represented by log-likelihood ratios (LLRs). True LLRs are unbiased and fair, i.e., given a large magnitude the estimated value is reliable. Given perfect channel state information (CSI) and a single-input single-output (SISO) system, the soft values are directly related to the variance of the noise process. Obviously, soft values are no longer only related to the noise process when channel estimation is performed. In [1] a soft decision metric is derived for a SISO system to calculate LLRs taking channel estimation errors into account. A significant performance improvement may be achieved when channel estimation errors are considered appropriately. However, the additional error variance due to the channel uncertainty is assumed to be known.

Iterative receivers based on factor graphs have been developed, which elegantly integrate soft information for data detection in a multiple-input multiple-output (MIMO) environment [2–5]. By utilizing the turbo principle, reliably detected data symbols may be used as pseudo training symbols, and thus, iteratively refine the channel estimates. The graph-based soft iterative receiver (GSIR) proposed in [2, 5] utilizes soft information consistently for both data detection and channel estimation.Due to a suboptimum detection strategy and imperfect channel estimation, the GSIR does not produce true LLRs for the data symbols anymore. Subsequently, we refer to approximate LLRs, defined as ALLRs. ALLRs represent the reliability of data symbols whereas the variance of a Gaussian function reflects the reliability of a channel coefficient estimate. Messages, which represent the estimates of data symbols and channel coefficients, are exchanged within the factor-graph according to the sum-product algorithm. Hereby, reliable messages have a stronger impact on the message generation than others.

In general, detection/estimation errors cannot be identified for sure. Hence, reliability information provides an important measure about an estimated variable. For suboptimum receiver algorithms and/or mismatches like imperfect knowledge of channel state information at the receiver, it is likely that the soft information is not correct, e.g. a large magnitude may be assigned to an unreliably estimated value. Wrongly detected data symbols which are considered to be reliable may significantly deteriorate the performance.

The calculation of the reliability information of a message is mainly influenced by noise, multi-antenna interference, and fading characteristics of a channel. In this paper the calculation of the reliability information of a message is refined to additionally consider detection errors based on the average bit error probability. Monte-Carlo simulations are provided to illustrate the performance gain for the refined variance model.

The remainder of this paper is organized as follows. Section II introduces the system and channel model. The multi-dimensional graph-based receiver and its extension to consider detection errors for the message generation is elaborated in Section III. Simulation results are presented in Section IV. Finally, Section V draws the conclusion.

II. System and Channel Model

The MIMO-OFDM system under consideration consists of L OFDM subcarriers and K OFDM symbols per frame. The equivalent discrete-time model of a MIMO channel with N_{T} transmit (Tx) and N_{R} receive (Rx) antennas can be represented after OFDM demodulation as

$$\begin{aligned} y_n[l,k] &= \sum_{m=1}^{N_{\mathrm{T}}} h_{n,m}[l,k]x_m[l,k] + w_n[l,k], \\ &= h_{n,m}[l,k]x_m[l,k] + \underbrace{\sum_{\substack{i=1 \\ i\neq m}}^{N_T} h_{n,i}[l,k]x_i[l,k]}_{\text{MAI}} + \underbrace{w_n[l,k]}_{\text{AWGN}}, \end{aligned} \tag{1}$$

978-1-4673-0990-5/12/$31.00 ©2012 IEEE

where $l \in \{0, 1, \ldots, L-1\}$ represents the OFDM subcarrier index and $k \in \{0, 1, \ldots, K-1\}$ the OFDM symbol index. The received symbol $y_n[l,k] \in \mathbb{C}$ consists of N_T superimposed signals. The link between the nth Rx antenna and the mth Tx antenna is described by the channel coefficient $h_{n,m}[l,k] \in \mathbb{C}$. The channel coefficients are assumed to be wide-sense-stationary (WSS), complex Gaussian variables with zero mean. $x_m[l,k] \in \mathbb{C}$ represents the channel input at the mth Tx antenna. $w_n[l,k]$ is an additive white Gaussian noise sample with zero mean and variance σ_w^2. The desired signal of the mth transmit antenna is thus superimposed with $(N_\mathrm{T}-1)$ interfering signals, termed multi-antenna interference (MAI), and the additive white Gaussian noise (AWGN) term.

III. Multi-Dimensional Graph-Based Receiver

In recent years, factor graphs [6] attracted great interest as a viable solution for low complexity iterative receivers. They provide a powerful graphical framework, which has been applied to a wide range of problems in digital communications.

The GSIR of [2, 5] is used in this paper as a framework to implement a refined variance model. The corresponding factor graph structure of a 2×2 MIMO system is shown in Fig. 1. The message exchange starts at the symbol nodes where symbol probabilities are calculated. Subsequently, the probabilities are used in the observation nodes to calculate channel estimates. One of the main features of the GSIR is the unified usage of soft information throughout the receiver. Reliability information is utilized for both data detection and channel estimation. Approximate log-likelihood ratios provide soft information of the detected symbols, while estimates of a channel coefficient are represented by a Gaussian pdf $\left(p(h) \sim \mathcal{CN}\left(\mu_h, \sigma_h^2\right)\right)$, i.e., the mean value is the hard estimate and the variance refers to the reliability. Accordingly, a large variance indicates an unreliable estimate and vice versa. Thus, a message from a coefficient node contains the estimate of a channel coefficient of one time and frequency index and is distributed to the neighboring coefficient nodes in order to further improve the channel estimation accuracy (cf. Fig. 2). The message exchange between coefficient nodes is established by transfer nodes.

A. Transfer Nodes

The purpose of a transfer node is to model the relation of neighboring channel coefficients, i.e., it converts the message

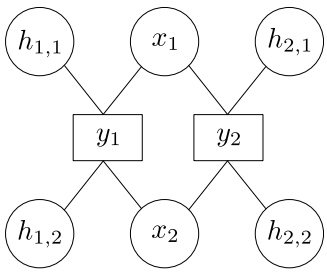


Fig. 1: Factor graph structure of a 2×2 MIMO system

of a channel coefficient $h_{n,m}[l,k]$ to $h_{n+n',m+m'}[l+l', k+k']$, which will be denoted with h and h' in the following to improve readability:

$$\Delta_{n',m'}[l',k'] \doteq h - \omega h', \qquad |\omega| = 1. \tag{2}$$

The tuning factor $\omega \in \mathbb{C}$ depends on the correlation properties between adjacent transfer nodes. Given a wide-sense-stationary (WSS) channel model, the transfer node is approximated by a zero-mean Gaussian pdf:

$$\Delta_{n',m'}[l',k'] \sim \mathcal{N}\left(0, \sigma^2_{\Delta,n',m'}[l',k']\right). \tag{3}$$

Based on (2), information between adjacent channel coefficients is exchanged as follows:

$$\begin{aligned} \mu_{h'} &= \omega \cdot \mu_h, \\ \sigma^2_{h'} &= \sigma^2_h + \sigma^2_{\Delta,n',m'}[l',k']. \end{aligned} \tag{4}$$

A message exchanged via a transfer node scales its mean μ_h by a factor of ω, whereas its variance is increased by the variance of the domain-specific transfer node $\sigma^2_{\Delta,n',m'}[l',k']$. This ensures that a message has only a finite influence to the overall message generation, since the variance determines the influence of the message.

The calculation of the variance is given in detail in [5] and is briefly summarized in the following. The variance of a transfer node is calculated as follows:

$$\begin{aligned} \sigma^2_{\Delta,n',m'}[l',k'] &= \mathbf{E}\left\{|h - \omega h'|^2\right\} \\ &= 2\left(1 - \mathrm{Re}\left[\omega \mathbf{E}\left\{h^* h'\right\}\right]\right), \end{aligned} \tag{5}$$

where $\mathbf{E}\{h^* h'\}$ corresponds to the multi-dimensional autocorrelation function between the two channel coefficients h and h'. The domain-specific tuning factor ω is chosen such that the real part in (5) is maximized, hence, the resulting variance is minimized. Often, the autocorrelation function can be calculated independently for each domain. Given a Jakes power spectral density for the distribution of the Doppler frequencies, the variance in time domain is given by

$$\sigma^2_{\Delta,\mathrm{t}} = 2\left(1 - J_0\left(2\pi f_{\mathrm{D,max}} T_\mathrm{s}\right)\right), \tag{6}$$

where $f_{\mathrm{D,max}}$ is the maximum Doppler frequency, T_s the symbol duration, and $J_0(\cdot)$ refers to the Bessel function of the first kind and order zero. The tuning factor for the time domain results to $\omega_t = 1$ due to the real-valued autocorrelation function. In the frequency domain, the power delay profile is typically described by an exponential distribution within $[0, \tau_{\max}]$, where $\tau_{\max}$ denotes the maximum propagation delay. The variance required for the transfer nodes in frequency domain amounts to

$$\sigma^2_{\Delta,\mathrm{f}} = 2\left(1 - \frac{1}{1 + 4\pi^2 \sigma_\tau^2 F^2}\right), \tag{7}$$

where σ_τ represents the root mean squared delay spread and F refers to the OFDM subcarrier spacing. The tuning factor is hereby set to $\omega_f = \frac{1}{1 - j2\pi\sigma_\tau F}$ in order to minimize the resulting variance. Assuming a uniform distribution of the angular spread Θ around a mean angle of departure ϕ within

the interval $[\phi{+}\Theta/2, \phi{-}\Theta/2]$ the variance in spatial domain is expressed by

$$\sigma^2_{\Delta,\mathrm{s}} = 2\left(1 - \frac{|A|\sqrt{A^2+B^2}}{A}\right), \tag{8}$$

where

$$A = J_0(z) + 2\sum_{m=1}^{\infty} J_{2m}(z)\cos(2m\phi)\frac{\sin(2m\Theta)}{2m\Theta}, \tag{9}$$

$$B = 2\sum_{m=0}^{\infty} J_{2m+1}(z)\sin\left((2m+1)\phi\right)\frac{\sin\left((2m+1)\Theta\right)}{(2m+1)\Theta}, \tag{10}$$

and $z{=}2\pi d_{\mathrm{Tx}}/\lambda$. The tuning factor for the spatial domain is set to

$$\omega_s = \exp\left(j\tan\left(-\frac{B}{A}\right)\right). \tag{11}$$

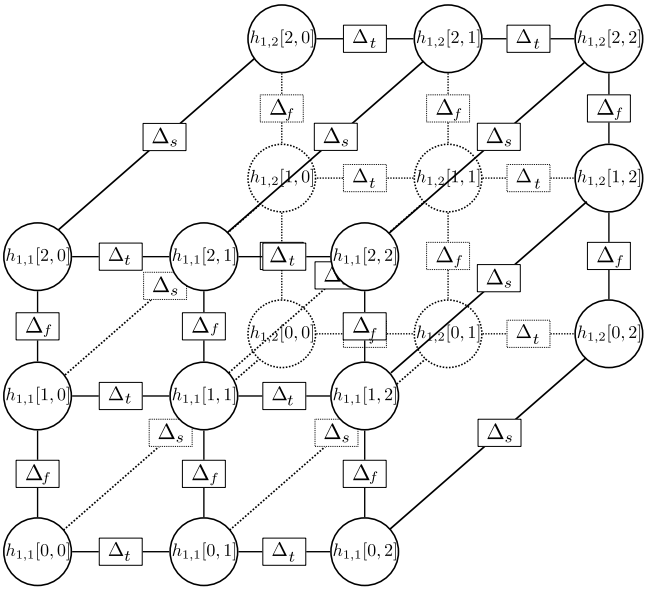


Fig. 2: 3D structure of the factor graph connecting coefficient nodes in time, frequency, and spatial domain.

B. Message Exchange

A coefficient node is connected to two transfer nodes in each domain and a symbol node. In order to ensure that only extrinsic messages are exchanged, a message generated at a certain node needs to consider all other connected nodes, except the node for which the message is generated. Suppose a channel coefficient receives the messages $p_i \sim \mathcal{CN}\left(\mu_i, \sigma_i^2\right), i \in \{1,\ldots,N\}$, where N is the number of nodes connected to the coefficient nodes. $N{-}1$ incoming messages have to be considered in order to generate one outgoing message. The product of Gaussian pdfs results in a complex-valued normal distribution:

$$\left(\prod_{j=1,j\neq i}^{N} p_j(h)\right) = p_i(h), \tag{12}$$

with mean and variance

$$\mu_i = \frac{\sum\limits_{j=1,j\neq i}^{N} \frac{\mu_j}{\sigma_j^2}}{\sum\limits_{j=1,j\neq i}^{N} \frac{1}{\sigma_j^2}}, \qquad \sigma_i^2 = \frac{1}{\sum\limits_{j=1,j\neq i}^{N} \frac{1}{\sigma_j^2}}. \tag{13}$$

As can be seen from (13), the variance of a message determines the influence of this message to the overall message generation. Messages with a small variance have a stronger impact than messages with a high variance.

Given a fast fading channel, the variance of a message traversing the 3D-graph is rapidly increased after a few nodes, which means that this message will only contribute marginally to the combined message. Whereas for a quasi time-invariant channel ($\sigma^2_{\Delta,\mathrm{t}} \approx 0$) the same message may contribute for the whole burst.

C. Influence of Detection Errors

The BER/MSE performance achieved by the GSIR partially depends on the accuracy of the calculated reliability information. Since the estimates of data symbols and channel coefficients are iteratively refined, it is important to differentiate between reliable and unreliable estimates.

The uncertainty of a channel coefficient introduced by the variance of the noise process of the AWGN channel, σ_w^2, together with the multi-antenna interference is summarized by σ_h^2 and illustrated in Fig. 3a. The hard estimate μ_h in the vicinity of the true channel coefficient depends on the noise and interference of the channel represented by a circle of radius σ_h^2. The additional uncertainty of a fading channel, when a message of a coefficient node is exchanged via transfer nodes, is taken into account by σ_Δ^2, which further expands the radius to $\sigma_h^2{+}\sigma_\Delta^2$ (cf. Fig. 3b). A third variance related to estimation errors is introduced by σ_e^2 (cf. Fig. 3c) and is explained in detail in the following.

It is commonly assumed that the soft value of an estimate is true and corresponds to the estimates' reliability. Nevertheless, a system setup with low training density, higher-order modulation, and multiple transmit antennas is sensitive to detection errors at low SNRs. In combination with channel coding, erroneous data symbols with high ALLRs may be generated.

Error propagation may occur due to the message exchange and the fact that messages which are considered to be reliable contribute more to the newly generated message. Given a slow fading channel, detection errors with large magnitude ALLRs easily influence the message generation of neighboring channel coefficients because of the low increase of variance by the transfer nodes, as described in Section III-B. One origin of wrongly estimated reliability information are the extrinsic ALLRs which are generated during iterations by the channel decoder and are used subsequently to refine the channel estimates. Since a priori information of data symbols are not available during initialization, only training symbols can be used to calculate initial channel estimates. On the basis of these initial estimates, the unknown data symbols are detected. Their information is sent to the channel decoder

where the extrinsic information of the decoded data symbols is generated. Given unreliable initial channel estimates, the extrinsic ALLRs are not necessarily correct. Furthermore, a strong code may generate ALLR values with large magnitude, which are not correct, i.e., the magnitude of an ALLR is large and the corresponding symbol is detected wrongly.

Assuming that a soft value is wrong implies several challenges. How to identify wrong soft information and how to calculate a reliability information of a soft value?

Although single soft values might be wrong, it is assumed that soft values on average do represent the true reliability of an estimate. Based on the extrinsic aLLRs, an average bit error probability can be calculated [7]. If the average bit error probability is low, it is unlikely that high magnitude ALLR errors are generated and vice versa.

Based on the average bit error probability, the error variance σ_e^2 is introduced to the message exchange at a transfer node. Accordingly, the variance of a channel coefficient estimate is extended to $\sigma_h^2+\sigma_\Delta^2+\sigma_e^2$, which accounts for the increase of uncertainty due to detection errors (cf. Fig 3c). Since all estimates of the factor graph are iteratively refined, also the average bit error probability needs to be recalculated after each completed iteration. The extrinsic information from the channel decoder of the ith iteration $\mathcal{L}^{(i)}$ is used to calculate the average bit error probability $\overline{P}_b^{(i)}$ [7]:

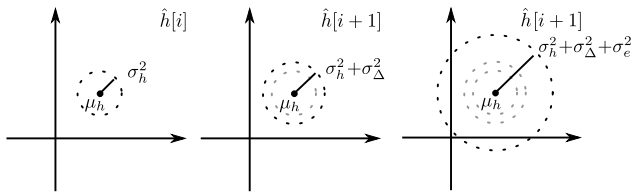
$$\overline{P}_b^{(i)} = \frac{1}{L \cdot K} \sum_{l=0}^{L-1} \sum_{k=0}^{K-1} \frac{1}{1+\exp\left(|\mathcal{L}^{(i)}[l,k]|\right)}. \quad (14)$$

Alternatively, in case of higher-order modulation a symbol error probability can be considered as well [7]:

$$\overline{P}_s^{(i)} = \frac{1}{L \cdot K} \sum_{l=0}^{L-1} \sum_{k=0}^{K-1} \left(1 - \prod_{n=0}^{N_m-1} \frac{1}{1+\exp(-|\mathcal{L}_n^{(i)}[l,k]|)}\right) \quad (15)$$

where N_m is the modulation order.

In the worst case scenario, an estimated coefficient is inverted in its phase, which is most likely to happen with BPSK/QPSK modulation, since only one or a few bit errors are required. For higher-order modulation it is more likely that the phase and/or amplitude is shifted by a certain degree, depending on the erroneously detected symbol. To calculate



(a) *Estimate of a channel coefficient at time index i.* (b) *Estimate of adjacent channel coefficient at time index $i+1$.* (c) *Estimate of adjacent channel coefficient at time index $i+1$ taking detection errors into account.*

Fig. 3: *Influence of variance on the estimation of a channel coefficient.*

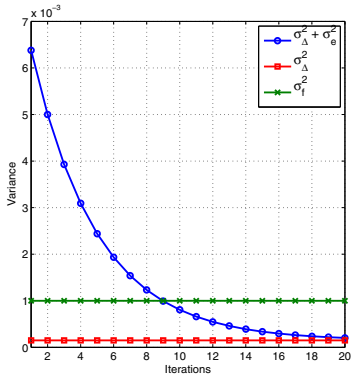


Fig. 4: Evolution of error variance as a function of iterations

σ_e^2 it is assumed that the event of an inverted phase occurs with the average probability $\overline{P}_b$:

$$\sigma_{e,n',m'}^2[l',k'] = \overline{P}_b^{(i)} \cdot 2 \cdot |\mu_{h,n',m'}[l',k']|^2. \quad (16)$$

The problem of wrong reliability information is that they cannot be identified for sure and, thus, distribute their information to neighboring nodes. The average bit error probability provides a measure to estimate how likely it is that wrong soft values are generated, which is used to prevent their distribution accordingly.

Such a constraint is consistently integrated into the existing factor graph by increasing the variance of message within a transfer node. The message update of a transfer node (4) is changed accordingly:

$$\mu_{h'} = \omega \cdot \mu_h, \quad (17)$$
$$\sigma_{h'}^2 = \sigma_h^2+\sigma_{\Delta,n',m'}^2[l',k']+\sigma_{e,n',m'}^2[l',k']. \quad (18)$$

In subsequent iterations the average bit error probability has to be re-calculated with the refined ALLR values of the data estimates. If the 3D-GSIR is able to converge, the error variance σ_e^2 is decreasing to zero over iterations, as illustrated in Fig. 4. As can be seen, the error variance σ_e^2 is converging to the variance of a transfer node σ_Δ^2 over the course of iterations. In order to reduce computational complexity a deterministic function could also be applied to calculate the evolution of the error variance. A fixed increased variance σ_f^2 is shown for comparison, and its effect is investigated in the numerical results.

IV. Numerical Results

A MIMO-OFDM system with $L = 64$ OFDM subcarriers, $K = 32$ OFDM symbols, $N_\mathrm{T} = 2$ transmit, and $N_\mathrm{R} = 4$ receive antennas is considered in the following. A turbo code with code rate $R = 1/3$ and 16-QAM modulation is used. The WINNER C2 NLOS channel model [8] is applied. A velocity of 10 km/h at a carrier frequency of 4 GHz is chosen, resulting in a normalized maximum Doppler shift $f_{\mathrm{D,max}}T_\mathrm{s} \approx 1.3815 \cdot 10^{-4}$. The maximum propagation delay in this model is $\tau_\mathrm{max} = 1.845\mu s$ [8]. A training spacing of

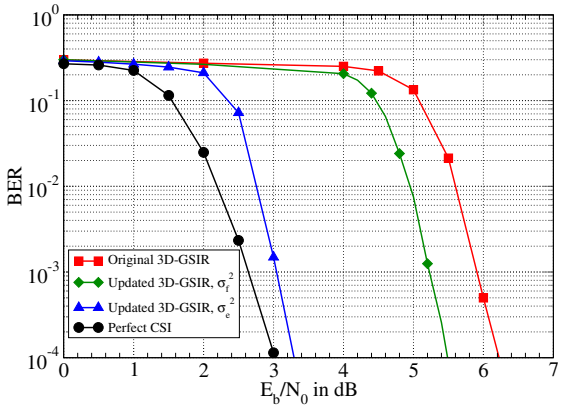


Fig. 5: *Comparison of BER performance of the original 3D-GSIR, the refined 3D-GSIR, and a 3D-GSIR with perfect CSI.*

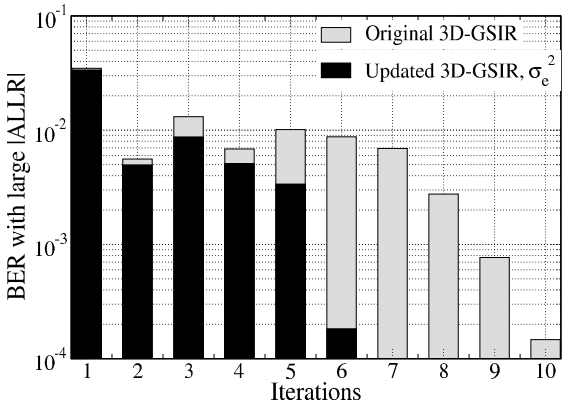


Fig. 6: BER with large |ALLR| at a SNR of 6 dB.

D_t=16 and D_f=16 in time and frequency domain, respectively, is used. The training symbols are orthogonally separated in time and frequency, which minimizes the interference on training symbols and simplifies the initialization.

BER results are given in Fig. 5. All results are obtained with a fixed number of 10 global iterations, which include one iteration for the GSIR and one turbo code iteration. The 3D-GSIR of [5] is termed 'Original 3D-GSIR'. For comparison, the GSIR with perfect channel state information (CSI) is included. As can be seen, the original 3D-GSIR is about 3.2 dB worse than the 3D-GSIR with perfect CSI at an BER of 10^{-4}. The deterioration of the original GSIR to the GSIR with perfect CSI can be mitigated by taking detection errors into account. The corresponding factor graph which considers the error variance based on the average bit error probability is denoted with 'Updated 3D-GSIR, σ_e^2'. The gain of the updated 3D-GSIR is nearly 3 dB compared to the original 3D-GSIR. The loss due to channel estimation is only about 0.4 dB. The gap to the perfect channel knowledge can be further reduced by increasing the number of iterations or by means of an improved initialization [9]. The 'Updated 3D-GSIR, σ_f^2', increases the variance of a message by a fixed value. As can be seen, the performance improves compared to the original 3D-GSIR, but is far away from optimum. Furthermore, if the value for σ_f^2 is chosen too large, the performance of the updated 3D-GSIR can get worse than that of the original 3D-GSIR.

Figure 6 shows the bit error rate which has been estimated with a large magnitude of LLR at an SNR of 6 dB for the original 3D-GSIR as well as for the updated GSIR, σ_e^2. Note here, that the results are superimposed for the two receivers. It can be seen that the updated 3D-GSIR which takes detection errors into account can correct all large magnitude LLR errors within 6 iterations, whereas bit errors remain for the original 3D-GSIR. Interestingly, the amount of the bit errors for the first iteration are nearly equal but can be corrected. It is worth mentioning that the miscalculation of reliability information appears not only for the studied case but in general. However, the influence on the performance is most pronounced for higher-order modulation.

V. Conclusions

A refined model for the message exchange within a graph-based soft iterative receiver is presented in this paper. The effect of detection errors on the overall performance is studied. In case of imperfect channel state information, the reliability information of the estimated variables does not necessarily represent the true reliability. Thus, error propagation occurs due to the message exchange. Based on the average bit error probability an additional variance is introduced which mitigates this effect and significantly improves the performance.

References

[1] M. M. Wang, W. Xiao, and T. Brown, "Soft decision metric generation for QAM with channel estimation error," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1058–1061, Jul. 2002.

[2] C. Knievel, Z. Shi, P. A. Hoeher, and G. Auer, "2D graph-based soft channel estimation for MIMO-OFDM," in *Proc. IEEE Int. Conf. on Communications (ICC)*, Cape Town, South Africa, May 2010.

[3] G. E. Kirkelund, C. N. Manchon, L. P. B. Christensen, E. Riegler, and B. H. Fleury, "Variational message-passing for joint channel estimation and decoding in MIMO-OFDM," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Miami, USA, Dec. 2010.

[4] X. Xu and R. Mathar, "Low complexity joint channel estimation and decoding for LDPC coded MIMO-OFDM systems," in *Proc. IEEE Vehicular Technology Conf. (VTC Spring)*, Budapest, Hungary, May 2011.

[5] C. Knievel, P. Hoeher, A. Tyrrell, and G. Auer, "Multi-dimensional graph-based iterative receiver for MIMO-OFDM," *IEEE Trans. Commun.*, submitted for publication.

[6] H.-A. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang, "The factor graph approach to model-based signal processing," *Proc. IEEE*, vol. 95, no. 6, pp. 1295–1322, Jun. 2007.

[7] J. Fricke and P. Hoeher, "Reliability-based retransmission criteria for hybrid ARQ," *IEEE Trans. Commun.*, vol. 57, no. 8, pp. 2181–2184, Aug. 2009.

[8] *IST WINNER II Channel Model, D1.1.2 V1.2*, 2008.

[9] C. Knievel, P. A. Hoeher, G. Auer, and A. Tyrrell, "Particle swarm enhanced graph-based channel estimation for MIMO-OFDM," in *Proc. IEEE Vehicular Technology Conf. (VTC Spring)*, Budapest, Hungary, May 2011.