Low Complexity Joint Channel Estimation and Decoding for LDPC Coded MIMO-OFDM Systems

Xiang Xu and Rudolf Mathar
Institute for Theoretical Information Technology, RWTH Aachen University
Aachen, Germany 52074
Email: {xu,mathar}@ti.rwth-aachen.de

***Abstract*—In this paper, a joint iterative channel estimation and low-density parity-check (LDPC) decoding algorithm based on factor graphs and the sum-product algorithm is proposed for orthogonal frequency division multiplexing (OFDM) systems employing multiple transmit and receive antennas (MIMO). By modeling time-varying frequency-selective fading channels as autoregressive (AR) processes and approximating messages as Gaussian pdf, this receiver algorithm is able to maintain a low complexity. Moreover, with the help of strong channel coding, pilot overhead can be significantly reduced.**

***Index Terms*—MIMO, OFDM, LDPC, channel estimation, iterative receiver, factor graph.**

I. Introduction

Powerful error-correcting codes based on iterative decoding, epitomized by LDPC codes and turbo codes, exhibit near-capacity performance over Rayleigh fading channels when perfect channel state information (CSI) is available at the receiver [1]. In practice, the knowledge of CSI is often acquired through channel estimation. To achieve better decoding performance, pilot symbols are periodically inserted [2] to extract CSI at the receiver. However, pure pilot based channel estimation is approved to be highly sub-optimal [3]. Therefore, soft channel estimators using detected data symbols as a-priori information are adopted[4]. Unluckily, joint detection and channel estimation algorithms show good performance but generally have high computational complexity, especially for MIMO systems. Recently, low complexity message passing algorithms designed on factor graphs are developed to track fast-varying frequency-selective MIMO channels with known channel statistics [5].

For OFDM systems, pilot symbols are allocated on the 2D time-frequency grid. The distance between pilot symbols is upper bounded by sampling theorem. While multiple transmit (Tx) antennas are employed, number of pilot symbols grows proportionally to the number of Tx antennas [3]. With LDPC codes, time and frequency diversity is considerably exploited by encoding the whole data frame to a long code word. Thus, the pilot overhead can be reduced.

In this paper, the message passing algorithms are represented by factor graphs. The variances of estimated channel frequency response (CFR) are used as reliability information for detection as well as refined channel estimation. Time and frequency correlations are considered in two kinds of correlation nodes. By approximating the correlation functions as low order AR processes, low complexity can be preserved while the number of pilots is reduced.

This paper is organized as follows. The system and channel models are described in Section II. Factor graphs are briefly explained in Section III. Section IV illustrates the details about the iterative channel estimation and decoding algorithm. Simulation results are presented in Section V. Finally, conclusions are made in Section VI.

II. System Model

Consider a MIMO system with N_T and N_R antennas for transmitter and receiver, respectively. On sth Tx antenna, info word $\mathbf{a}_s \in \{0,1\}^{K_u}$ is non-systematically encoded to $\mathbf{c}_s = e(\mathbf{a}_s) \in \{0,1\}^{K_c}$, where $s = 1, 2, \cdots, N_\mathrm{T}$ is Tx antenna index. Interleavers are avoided due to the non-systematic encoding. The coded bit streams are modulated to sequence of symbols $\mathbf{x}_s$ chosen from an M-ary constellation $\mathscr{X}$. The encoding and mapping process can be expressed as $\mathbf{x}_s = \mathcal{C}(\mathbf{a}_s)$. With parallelization and properly inserted pilot symbols, the modulated symbol stream are then modulated by orthogonal subcarriers via inverse fast Fourier transform (IFFT), and an additional cyclic prefix is added on the tail of OFDM symbols to remove inter-symbol interferences (ISI). The parallel signals are converted to serial form and transmitted through different antennas.

A. Channel model

With appropriate cyclic prefix insertion and sampling, the MIMO-OFDM system with K subcarriers decouples frequency-selective channel into K correlated flat-fading channel. The received signal on each receive (Rx) antenna at the kth tone of the nth OFDM block is the superposition of N_T distorted transmitted signals, thus, can be expressed as

$$y_u[n,k] = \sum_{s=1}^{N_\mathrm{T}} h_{u,s}[n,k] x_s[n,k] + w_u[n,k], \tag{1}$$

where $u = 1, 2, \cdots, N_\mathrm{R}$ is Rx antenna index. Additive white Gaussian noise (AWGN) with zero mean and variance σ_w^2 is denoted by $w_u[n,k]$, while the symbol power is normalized to 1. The CFR between the sth Tx and the

This work is partially supported by UMIC, a research project in the frame work of the German excellence initiative.

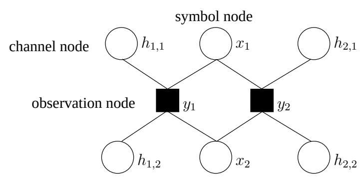

Fig. 1. Factor graph of 2×2 MIMO system for one subcarrier at one time index.

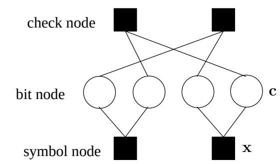

Fig. 2. Factor graph of QAM mapping and LDPC coding.

uth Rx antenna is denoted by $h_{u,s}[n,k]$. The average power of the CFR in each subchannel is normalized to 1. Define $\mathbf{y}_u = [y_u[0,0], y_u[0,1], \cdots, y_u[N-1, K-1]]$, and $\mathbf{h}_{u,s} = [h_{u,s}[0,0], h_{u,s}[0,1], \cdots, h_{u,s}[N-1, K-1]]$, (1) can also be written in vector form:

$$\mathbf{y}_u = \sum_{s=1}^{N_{\mathrm{T}}} \mathbf{h}_{u,s}\mathbf{x}_s + \mathbf{w}_u, \tag{2}$$

Assuming there is no spatial correlation, the correlation function of CFR for different times and frequencies can be decoupled as

$$\mathrm{E}\{h[n+\Delta n, k+\Delta k]h^*[n,k]\} = r_t[\Delta n]r_f[\Delta k], \tag{3}$$

where r_t is the time domain correlation depending on the Doppler frequency [6] and r_f is the frequency domain correlation depending on the power delay profile [7].

B. Pilot allocation

Known pilot symbols are inserted to the two dimensional time-frequency grid for each Tx antenna. To keep the orthogonal separation of pilots in both time and frequency, during a pilot grid, only one antenna is allowed to transmit. The position of a pilot can be denoted by a vector $\mathbf{p} = [s, n, k]^{\mathrm{T}}$. The position vector can be uniquely determined by [8]

$$\mathbf{p} = \mathbf{D}\tilde{\mathbf{p}} + \mathbf{p}_0 \pmod{\mathbf{m}} \tag{4}$$

with

$$\mathbf{D} = \begin{pmatrix} D_{\mathrm{s}} & \delta_{\mathrm{st}} & \delta_{\mathrm{sf}} \\ d_{\mathrm{st}} & D_{\mathrm{t}} & \delta_{\mathrm{ft}} \\ d_{\mathrm{sf}} & d_{\mathrm{ft}} & D_{\mathrm{f}} \end{pmatrix}, \mathbf{p}_0 = \begin{pmatrix} s_0 \\ n_0 \\ k_0 \end{pmatrix}, \mathbf{m} = \begin{pmatrix} N_{\mathrm{T}} \\ N \\ K \end{pmatrix} \tag{5}$$

where $\tilde{\mathbf{p}}$ is the pilot index, $\mathbf{p}_0$ is the position of the first pilot, D_{s}, D_{t} and D_{f} are distance between pilots in space, time and frequency, respectively. The off-diagonal entries in $\mathbf{D}$ indicates the shift on time-frequency grid. Due to the double selectivity, pilots must be scattered in both time and frequency domain. Thus, these off-diagonal elements should be carefully chosen, such that pilots are properly separated as well as the time and frequency orthogonality can still be preserved.

According to Nyquist sampling theorem, to perfectly reconstruct a time varying frequency selective channel from pilot symbols, the maximum spacings of pilot symbols are bounded by [9]

$$D_{\mathrm{f}} \leq \frac{\tau_{\max}}{T_{\mathrm{s}}}, \quad D_{\mathrm{t}} \leq \frac{1}{2f_{\mathrm{d}}T}. \tag{6}$$

Later we will show this bound can be surpassed with the help of channel coding.

III. Factor graphs

Consider a global function $g(x_1, \cdots, x_n)$ which can be factorized into the product of several local functions

$$g(x_1, \cdots, x_n) = \prod_{j \in J} f_j(X_j) \tag{7}$$

where J is the discrete index set, X_j is a subset of $\{x_1, \cdots, x_n\}$, and $f_j(X_j)$ is a function having the elements of X_j as arguments. The factor graph corresponding to the factorization consists of variable nodes for x_i, function nodes for f_j and edges connecting function nodes to their associated variable nodes. For example, the channel model described in (1) can be represented by factor graph depicted in Fig. 1, where transmitted symbols and CFR are both treated as variables. The factor graph of QAM mapping and LDPC coding is illustrated in Fig. 2, where the mapping and parity check functions take coded bits as variables.

A general method for computing the global function g is to apply the sum-product algorithm [10]. In the sum-product algorithm, function nodes and variable nodes send messages to their associated nodes, the messages can be described as follows:
messages from variable nodes to function nodes:

$$\mu_{x \to f}(x) = \prod_{h \in S_x \setminus \{f\}} \mu_{h \to x}(x), \tag{8}$$

message from function nodes to variable nodes:

$$\mu_{f \to x}(x) = \sum_{\sim x} \left(f(X) \prod_{y \in S_f \setminus \{x\}} \mu_{y \to f}(y) \right), \tag{9}$$

where S_v denote the set of neighbors of a node v in a factor graph, $\sum_{\sim x}$ means summation over all unknown variables in the summand except x. To get reasonable complexity, low factor node degrees are usually preferred [11]. For instance, the correlation function in (3) can be described by factor graph in Fig. 3. A further simplification is to consider only first order channel correlation, corresponding factor graph is shown in Fig. 4.

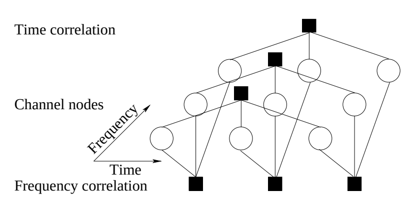

Fig. 3. Factor graph for decoupled channel correlations

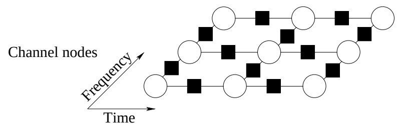

Fig. 4. Factor graph for first order channel correlations

IV. Iterative Receiver Design

A. Receiver structure

For simplicity reasons, time and subcarrier indices are omitted in this section. Consider uncoded bits a_s, define $\mathbf{Y}$ as the matrix which collects all the received symbols, the optimal receiver with respect to bit-error probability is:

$$\hat{a}_s \triangleq \arg \max_{a_s \in \{0,1\}} p(a_s|\mathbf{Y}) \tag{10}$$

Assuming a_s uniformly distributed, according to Bayes' rule, we obtain:

$$p(a_s|\mathbf{Y}) = \sum_{\sim a_s} p(\mathbf{A}|\mathbf{Y}) \propto \sum_{\sim a_s} p(\mathbf{Y}|\mathbf{A}) \tag{11}$$

where $\mathbf{A}$ is the matrix consisting of all the uncoded bits, $\propto$ means equality up to irrelevant additive and multiplicative constants. We can factor $p(\mathbf{Y}|\mathbf{A})$ as

$$p(\mathbf{Y}|\mathbf{A}) = \sum_{\mathbf{X}} p(\mathbf{Y}|\mathbf{X}) \prod_{s=1}^{N_\mathrm{T}} I(\mathbf{x}_s = \mathcal{C}(\mathbf{a}_s)), \tag{12}$$

because of the one-to-one encoding and mapping procedure. The indicator function $I(statement)$ is one if the statement is true and zero otherwise. Consider unknown channel coefficients also as variables, we have

$$p(\mathbf{Y}|\mathbf{X}) = \int_{\mathbf{H}} p(\mathbf{Y}|\mathbf{X}, \mathbf{H}) p(\mathbf{H}) \mathrm{d}\mathbf{H} \tag{13}$$

where $\mathbf{H}$ is a matrix collecting all the channel frequency response, which has $N_\mathrm{T} N_\mathrm{R} K N$ entries. Since spatial correlation is not considered in this work, noticing the fact that received signals on different Rx antennas or at different indices have weak correlation[12], the following approximation can be made to reduce complexity:

$$p(\mathbf{Y}|\mathbf{X}, \mathbf{H}) \approx \prod_{u=1}^{N_\mathrm{R}} \prod_{s=1}^{N_\mathrm{T}} p(\mathbf{y}_u|\mathbf{x}_s, \mathbf{h}_{u,s}), \tag{14}$$

the integration in (13) becomes

$$p(\mathbf{Y}|\mathbf{X}) \approx \int_{\mathbf{h}_{u,s}} \prod_{u=1}^{N_\mathrm{R}} \prod_{s=1}^{N_\mathrm{T}} p(\mathbf{y}_u|\mathbf{x}_s, \mathbf{h}_{u,s}) p(\mathbf{h}_{u,s}) \mathrm{d}\mathbf{h}_{u,s} \tag{15}$$

Furthermore, using (1), with inserted pilot symbols, we have

$$p(\mathbf{y}_u|\mathbf{x}_s, \mathbf{h}_{u,s}) = \prod_{x_s \notin \mathcal{P}_s} p(y_u|x_s, h_{u,s}) \prod_{x_s \in \mathcal{P}_s} p(y_u|h_{u,s}), \tag{16}$$

where $\mathcal{P}_s$ is the set of pilots on antenna s. Till this step, the conditional pdf in (10) has been noticeably simplified. Further simplifications are possible with suitable approximations. The message passing algorithm is described in detail below.

B. Gaussian approximation

To simplify MIMO detection, the Gaussian approximation is made. Consider an arbitrary symbol x_s and its observation y_u, all the signal components contributed by symbols other than x_s can be regarded as effective noises:

$$y_u = h_{u,s} x_s + v_u, \tag{17}$$

where the effective noise sample defined as

$$v_u = \sum_{\varsigma=1, \varsigma \neq s}^{N_\mathrm{T}} h_{u,\varsigma} x_\varsigma + w_u \tag{18}$$

is approximated by a Gaussian distribution. Due to the inaccuracy of the channel estimation, the noisy estimate of CFR can be written as:

$$\widehat{h}_{u,s} = h_{u,s} + e_{u,s}, \tag{19}$$

where the estimation error $e_{u,s}$ can be assumed to have a Gaussian distribution $\mathcal{CN}(0, \sigma^2_{h_{u,s}})$ [13]. Therefore, mean and variance of effective noise $v_{u,s}$ can be calculated via

$$\begin{aligned} \mu_{v_{u,s}} &= \sum_{\varsigma=1, \varsigma \neq s}^{N_\mathrm{T}} \widehat{h}_{u,\varsigma} \widehat{x}_\varsigma \\ \sigma^2_{v_{u,s}} &= \sum_{\varsigma=1, \varsigma \neq s}^{N_\mathrm{T}} \sigma^2_{h_{u,\varsigma}} + \sigma^2_{x_\varsigma} |\widehat{h}_{u,\varsigma}|^2 + \sigma^2_w, \end{aligned} \tag{20}$$

where $\widehat{x}_\varsigma$ and $\sigma^2_{x_\varsigma}$ are mean and variance of the detected symbol x_ς. Statistics of effective noise samples are calculated during every iteration for both detection and channel estimation purposes.

C. Message update for channel estimation

The message from observation nodes to channel variable nodes are

$$\begin{aligned} \mu_{y_u \to h_{u,s}}(h_{u,s}) = & \sum_{\mathbf{x}_s} \int_{\sim h_{u,s}} p(y_u|x_s, h_{u,s}) \prod_{\varsigma=1}^{N_\mathrm{T}} \mu_{x_\varsigma \to y_u}(x_\varsigma) \\ & \prod_{\nu=1, \nu \neq s}^{N_\mathrm{T}} \mu_{h_{u,\nu} \to y_u}(h_{u,\nu}) \mathrm{d}\mathbf{h}_{u,s}, \end{aligned} \tag{21}$$

Since no close form integration can be deduced, we use the Gaussian approximation to simplify (21) to

$$\mu_{y_u \to h_{u,s}}(h_{u,s}) = \sum_{x_s \in \mathscr{X}} p(y_u|x_s, h_{u,s}) P(x_s) \tag{22}$$

where $\mathscr{X}$ is the set of all constellation points. With a-priori information $P(x_s)$, which is feeded by the symbol detector, this conditional pdf is mixed Gaussian. Since mixed Gaussian distribution cannot be calculated in a fast and accurate manner, proper approximation has to be made. With a successful detection, we have $P(x_s = q_i) \gg P(x_s = q_\nu), \nu \neq i$, hence $p(y_u|h_{u,s})$ can be well approximated by a Gaussian distribution $\mathcal{CN}(\mu_{h_{u,s}}, \sigma^2_{h_{u,s}})$. The message passing from observation nodes to channel variable nodes can be expressed as

$$\mu_{y_u \to h_{u,s}}(h_{u,s}) = \frac{1}{\pi \sigma^2_{h_{u,s}}} \exp\left(-\frac{|h_{u,s} - \widehat{h}_{u,s}|^2}{\sigma^2_{h_{u,s}}}\right), \quad (23)$$

where the estimator $\widehat{h}_{u,s}$ is obtained by least-square estimation

$$\widehat{h}_{u,s} = (y_u - \mu_{v_{u,s}})\widehat{x}_s^* \quad (24)$$

where $(\cdot)^*$ denotes complex conjugate. The variance is calculated from

$$\sigma^2_{h_{u,s}} = \sigma^2_{v_{u,s}} + \sigma^2_{x_s}|y_u - \mu_{v_{u,s}}|^2. \quad (25)$$

As the distribution of $h_{u,s}$ is already known, only $\widehat{h}_{u,s}$ and $\sigma^2_{h_{u,s}}$ actually need to be passed.

D. Refined channel estimation

The estimated CFR are further refined using channel statistics. Define $h'_{u,s} = h_{u,s}[n + \Delta n, k + \Delta k]$, to simplify this procedure, we make the following approximation:

$$h'_{u,s} = r h_{u,s} + \sqrt{1 - r^2}\xi = r\widehat{h}_{u,s} + r e_{u,s} + \sqrt{1 - |r|^2}\xi \quad (26)$$

where $r = \mathrm{E}\{h'_{u,s} h^*_{u,s}\}$, $\xi \sim \mathcal{N}(0, 1)$. Because of the noisy channel estimation, $h'_{u,s}$ follows the Gaussian distribution with mean $r\widehat{h}_{u,s}$ and variance $r^2\sigma^2_{h_{u,s}} + (1 - |r|^2)$. The refined information is generated in correlation nodes and passed to channel nodes. Only first order correlation is considered in this work, i.e. $\Delta n + \Delta k = \pm 1$.

As variable nodes, channel nodes update messages following the rule in (8)

$$\mu_{h_{u,s} \to f}(h_{u,s}) = \prod_{g \in S_{h_{u,s}} \setminus \{f\}} \mu_{g \to h_{u,s}}(h_{u,s}) \quad (27)$$

where f and g can be either observation nodes or correlation nodes, $S_{h_{u,s}}$ is the set of all function nodes connected to $h_{u,s}$. Since $p(h_{u,s})$ is Gaussian, the product will also be Gaussian with mean and variance:

$$\mu_f = \frac{\sum_g \frac{\mu_g}{\sigma_g^2}}{\sum_g \frac{1}{\sigma_g^2}}, \quad \sigma_f^2 = \frac{1}{\sum_g \frac{1}{\sigma_g^2}}, \quad (28)$$

where μ_g and σ_g^2 are mean and variance of message $\mu_{g \to h_{u,s}}(h_{u,s})$.

E. Message update for detection

Based on the same consideration of (21) The message passing from observation nodes to symbol nodes can be written as

$$\begin{aligned} \mu_{y_u \to x_s}(x_s) &= p(y_u|x_s) \\ &\propto \exp\left(-\frac{|y_u - \widehat{h}_{u,s}x_s - \mu_{v_{u,s}}|^2}{|x_s|^2\sigma^2_{h_{u,s}} + \sigma^2_{v_{u,s}}}\right) \end{aligned} \quad (29)$$

Generally, log likelihood ratios (LLR) are used for LDPC decoding. Since for M-ary modulation, coded bit c_s^i, $i \in \{1, \cdots, \log_2 M\}$ has N_R independent observations. The output the output LLRs of the data detector is given as [14]

$$\begin{aligned} \mathrm{LLR}(c_s^i) &= \sum_{u=1}^{N_\mathrm{R}} \log \frac{P(c_s^i = 0|y_u)}{P(c_s^i = 1|y_u)} \\ &= \sum_{u=1}^{N_\mathrm{R}} \max_{x_s \in \mathscr{X}_0^i} \log p(y_u|x_s) - \max_{s_s \in \mathscr{X}_1^i} \log p(y_u|x_s) \end{aligned} \quad (30)$$

where $\mathscr{X}_b^i$ denotes the set of constellation points corresponding to $c_s^i = b$. LLR messages are combined at the parity check function nodes and feed back to bit variable nodes following LDPC decoding constraint [1]. The feed back LLR messages are additive. The probability of each bit can be recovered from the summed LLR of each bit variable nodes by

$$\begin{aligned} P(c_s^i = 0) &= \frac{\exp(\mathrm{LLR}(c_s^i))}{1 + \exp(\mathrm{LLR}(c_s^i))} \\ P(c_s^i = 1) &= \frac{1}{1 + \exp(\mathrm{LLR}(c_s^i))}. \end{aligned} \quad (31)$$

For symbol by symbol detection, the decision is yield by taking average among all constellation points

$$\widehat{x}_s = \sum_{l=1}^{M} (\prod_{i=1}^{\log_2 M} P(c_s^i) I(x_s = q_l)) q_l \quad (32)$$

where q_l are M-ary constellation points. The corresponding variance is

$$\sigma^2_{x_s} = 1 - |\widehat{x}_s|^2, \quad (33)$$

as symbol energy is normalized to 1.

F. Scheduling and complexity

The message updating starts from the initial pilot based channel estimation. The CFR corresponding to data symbols are updated by channel correlation nodes. Afterwards, in each iteration, observation nodes simultaneously update messages to each symbol node, bit metrics are generated from symbol nodes and forwarded to LDPC decoder. After decoding, symbol estimators are generated according to LLR values and utilized in observation nodes to perform data aided channel estimation. The data aided channel estimator are further refined using channel correlations for the next iteration. Since in

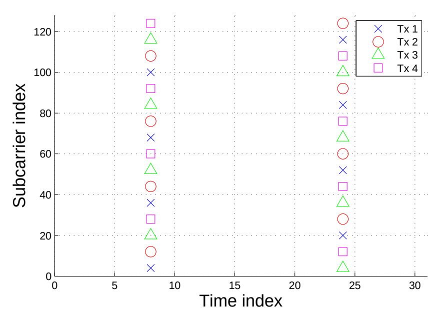

Fig. 5. Pilot allocation for simulation, $N_\mathrm{T} = 4$, $N = 32$, $K = 128$, $D_\mathrm{f} = 32$, $D_\mathrm{t} = 16$.

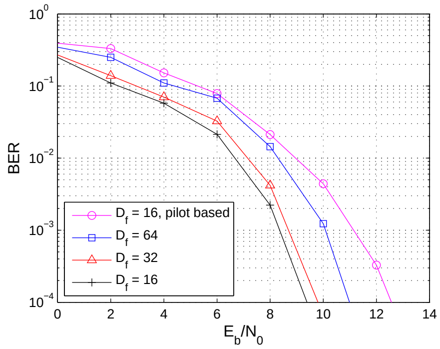

Fig. 6. Comparison of BER versus E_b/N_0 for different pilot spacing.

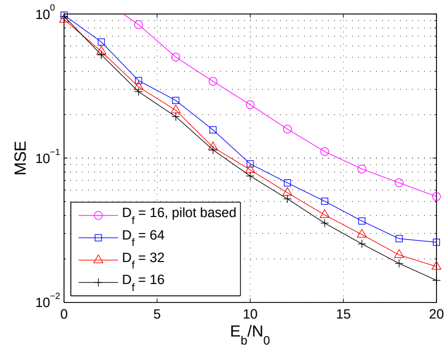

Fig. 7. Comparison of MSE versus E_b/N_0 for different pilot spacing.

every iteration, each message updating operation is performed only once, the complexity of this receiver grows linearly with respect to the number of Tx and Rx antennas, number of OFDM symbols and subcarriers, modulation order and degree of correlations.

V. Simulation Results

We consider a 4×4 MIMO systems with block transmission. Each block consists of 32 symbols. Baseband modulation is 4-QAM. A total bandwidth of 10 MHz is divided into 128 subchannels, resulting in a symbol length of 12.8 μs. A non-systematic LDPC code with coding rate $1/4$ and column weight 3 is applied to each Tx antenna. The carrier frequency is 5 GHz. Normalized Doppler frequency is $f_\mathrm{d}\mathrm{T} = 0.02$, corresponds to a moving speed around 80 km/h. The maximum delay spread is 2 μs. According to sampling theorem, pilot spacing is bounded by $D_\mathrm{t} \le 25$ and $D_\mathrm{f} \le 20$. Setting positioning parameters$D_\mathrm{s} = 1$, $d_\mathrm{sf} = D_\mathrm{f}/N_\mathrm{T}$, $\delta_\mathrm{ft} = 2d_\mathrm{sf}$ and the other off-diagonal elements in $\mathbf{D}$ to 0 results in a diamond pilot allocation [15] depicted in Fig. 5.

The performance is evaluated by both bit error ratio (BER) in Fig. 6 and mean square error (MSE) in Fig. 7. While the pilot spacing in time is fixed to $D_\mathrm{t} = 16$, systems with different pilot spacing in frequency domain are compared. The experimental results show that both BER and MSE goes down when the pilot spacing gets smaller. However, comparing to systems with $D_\mathrm{f} = 16$, systems with $D_\mathrm{f} = 32$ lose only 0.5 dB in signal to noise ratio (SNR), while only half of the pilots are used. Furthermore, even when the pilot spacing is larger than the upper bound set by sampling theorem, the system still works with a penalty in SNR. Performance of systems with pilot based channel estimation using linear interpolation is also plotted for comparison. In the whole SNR range, our receiver evidently outperforms systems with linear interpolation even if a much smaller number of pilots is employed.

VI. Conclusion

In this paper, we address the problem of joint channel estimation and LDPC decoding for MIMO-OFDM systems with factor graphs and the sum-product algorithm. Using different simplifications of the graph, an iterative receiver with low complexity is obtained. Even when the number of pilots is smaller than the limit according to sampling theorem, this receiver is able to track the time-varying frequency-selective channel.

References

[1] D. J. C. MacKay, *Information theory, inference, and learning algorithms*, Jan. 2003.

[2] Z. Tang, R. Cannizzaro, G. Leus, and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems," *Signal Processing, IEEE Transactions on*, vol. 55, no. 5, pp. 2226–2238, May 2007.

[3] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *Information Theory, IEEE Transactions on*, vol. 49, no. 4, pp. 951 – 963, apr. 2003.

[4] H. Niu, M. Shen, J. Ritcey, and H. Liu, "A factor graph approach to iterative channel estimation and LDPC decoding over fading channels," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 4, pp. 1345–1350, July 2005.

[5] X. Xu and R. Mathar, "Factor graph based detection and channel estimation for MIMO-OFDM systems in doubly selective channel," in *7th International Symposium on Wireless Communication Systems (IEEE ISWCS)*, York, United Kingdom, Sep. 2010, pp. 1–5.

[6] J. W. C. Jakes, *Microwave Mobile Communications*. New York: IEEE Press, 1994.

[7] Y. Li, J. Cimini, L.J., and N. Sollenberger, "Robust channel estimation for ofdm systems with rapid dispersive fading channels," *Communications, IEEE Transactions on*, vol. 46, no. 7, pp. 902–915, Jul 1998.

[8] G. Auer, "3D pilot aided channel estimation," in *IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2009, pp. 1–6.

[9] H. Arslan and T. Ycek, "Estimation of frequency selectivity for OFDM based new generation wireless communication systems," vol. 1, San Francisco, Ca, May 2003.

[10] F. Kschischang, B. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *Information Theory, IEEE Transactions on*, vol. 47, no. 2, pp. 498–519, Feb 2001.

[11] A. Worthen and W. Stark, "Unified design of iterative receivers using factor graphs," *Information Theory, IEEE Transactions on*, vol. 47, no. 2, pp. 843 –849, feb. 2001.

[12] T. Wo and P. Hoeher, "A simple iterative Gaussian detector for severely delay-spread MIMO channels," in *IEEE International Conference on Communications (ICC)*, June 2007, pp. 4598–4603.

[13] R. Narasimhan, "Performance of diversity schemes for OFDM systems with frequency offset, phase noise and channel estimation errors," in *Communications, 2002. ICC 2002. IEEE International Conference on*, vol. 3, 2002, pp. 1551 – 1557.

[14] M. Wang, W. Xiao, and T. Brown, "Soft decision metric generation for QAM with channel estimation error," *Communications, IEEE Transactions on*, vol. 50, no. 7, pp. 1058 – 1061, jul. 2002.

[15] J.-W. Choi and Y.-H. Lee, "Optimum pilot pattern for channel estimation in OFDM systems," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 5, pp. 2083 – 2088, sep. 2005.