Expectation Propagation for Near-Optimum Detection of MIMO-GFDM Signals

Dan Zhang*, Luciano Leonel Mendes†*, Maximilian Matthe*, Ivan Simões Gaspar*, Nicola Michailow* and Gerhard Fettweis*, Fellow, IEEE

*Vodafone Chair Mobile Communications Systems, Technische Universität Dresden, Germany
†Instituto Nacional de Telecomunicações, Telecommunication Engineering Department, Sta. Rita do Sapucai, MG, Brazil

Abstract—Generalized frequency division multiplexing (GFDM) as a non-orthogonal waveform aims at diverse applications in future mobile networks. To evaluate its performance, its capacity limits are of particular importance. Therefore, this paper analyzes its constellation-constrained capacities for cases where the channel state information (CSI) is unknown at the transmitter and perfectly known at the receiver. In frequency selective channels, GFDM may provide advantage over the conventional orthogonal frequency division multiplexing (OFDM) scheme. In order to achieve near-capacity performance, the interaction of data symbols in time and frequency combined with multiple antennas (MIMO) challenges the design of GFDM receivers. This paper therefore applies expectation propagation (EP) for systematic receiver design. It is shown that the resulting iterative MIMO-GFDM receiver with affordable complexity can approach optimum decoding performance and outperform MIMO-OFDM in a rich multipath environment. Simulations are also used to illustrate the impact of channel delay spread on the constellation-constrained capacities and on the performance of the novel receiver algorithm.

Index Terms—5G, non-orthogonal waveforms, GFDM, MIMO, mutual information, expectation propagation

I. INTRODUCTION

Future fifth-generation (5G) communications systems need to handle a wide range of applications [1]. The demand for throughput requires a high capacity modulation scheme. Low out-of-band (OOB) emission is mandatory to spectrum aggregation and dynamic spectrum allocation. Tactile Internet needs ultra-low latency [2]. Internet of Things (IoT), with a multitude of devices, must allow fast coarse synchronization or even completely asynchronous transmission to prevent energy waste [3]. As orthogonal frequency division multiplexing (OFDM) falls short of fulfilling these requirements, a new and more flexible waveform is needed for 5G networks. In fact, it was believed in [1] that the orthogonality of OFDM constitutes a major obstacle for the envisioned 5G applications. This understanding has been motivating the invention and also rediscovery of non- and quasi-orthogonal waveforms, such as filterbank multi-carrier [4], universal-filtered multi-carrier [5] and generalized frequency division multiplexing (GFDM) [6].

This paper considers GFDM because of its flexibility to cover different scenarios foreseen for 5G networks [6]. Within one GFDM block, each subcarrier carries multiple data symbols in different time slots, known as subsymbols, and they are individually filtered using circular convolution. This means each GFDM block has a fixed length and all necessary pulse-shapes are obtained by circularly shifting a prototype filter in time and frequency. The block structure of GFDM without filter tails favors latency constrained applications. By choosing a pulse shaping filter with good spectral localization, GFDM can achieve low OOB emission and its subcarriers can be easily arranged to provide fragmented and opportunistic spectrum allocation. As a non-orthogonal waveform, GFDM can tolerate loose time and frequency synchronization for IoT applications. Last but not least, GFDM can emulate OFDM and single carrier frequency domain equalization (SC-FDE) [6], supporting a smooth migration from 4G to 5G.

Besides the above mentioned advantages of GFDM, its non-orthogonality may provide additional freedom of exploiting frequency selectivity of rich multipath channels. Assume no knowledge of channel state information (CSI) at the transmitter but CSI perfectly known by the receiver. This paper presents an information theoretic analysis of GFDM, showing its constellation-constrained capacity can be higher than that of OFDM under frequency selective channels. In order to achieve the performance offered by GFDM, the optimum receiver needs to deal with both intercarrier interference (ICI) and intersymbol interference (ISI). The receiver task becomes even more challenging for multiple-input multiple-output (MIMO) systems using spatial multiplexing, where interantenna interference (IAI) also takes place. Given such three-dimensional interference, the complexity of optimal decoding is prohibitively high. Therefore, approximations are inevitable and losses compared with the theoretic limit are expected.

Expectation propagation (EP) has been successfully applied in the literature for iterative receiver designs, e.g., [7]–[14] and references therein. It can be visualized as a message passing algorithm operating on a factor graph (FG). To use it efficiently, it is critical to explore special features of the underlying system and also understand practical constraints. For instance, the authors of [10] have revealed how to link the EP algorithm for small-scale MIMO detection to conventional MIMO detection using a linear minimum mean square error (LMMSE) based interference cancellation (LMMSE-IC) algorithm. Considering large-scale MIMO systems, the goal of [12], [13] was to approximate the message update equations of EP for achieving a manageable complexity at the receiver. In [7], [9], [14], the application of EP is based on a vector-form FG representation of the probabilistic model of the considered ISI channels. Another example is [8], which is based on a flat fading channel. As the time-varying feature of the channel is modeled as an autoregressive moving-average process, the
system can be written as a state-space model. The correlation between two state vectors decreases as their time spacing increases. Due to this property, the authors of [8] developed a window-based EP algorithm such that a subset of observations is sufficient to yield a reliable estimate.

In this paper, we tailor EP for near-optimum detection of MIMO-GFDM signals. That includes FG construction, derivation of message update equations and also scheduling. Briefly, the FG is first constructed based on the channel input-output relation in frequency domain (FD). This is not only because the presence of cyclic prefix (CP) eases equalization in FD, but also because GFDM adopts circular filters with good spectral localization. Due to the latter, the variable nodes in the FG are coupled in a localized manner. Next, we derive the message update equations to iteratively resolve ICI, ISI and IAI. The update equations to iteratively resolve ICI, ISI and IAI. The presence of cyclic prefix (CP) eases equalization in FD, but contrary, the application of EP for GFDM requires to take into account the spatial, spectral and temporal correlation of data symbols. The order of message passing influences not only the decoding performance, but also the degree of parallelism in computation. Taking advantage of spectrally-localized correlation in GFDM, this paper proposes a hybrid scheduling mechanism, showing near-optimum decoding performance can be achieved with a high level of parallelism in computation.

In summary, this paper contributes with the following aspects. First, we introduce the FD model of the considered MIMO-GFDM system in Section II. Based on this model, Section III provides an information theoretic analysis of the MIMO-GFDM system. The detailed derivation of the EP-based iterative equalization and detection algorithm is presented in Section IV, followed by remarks on its complexity and scheduling in Section V. Finally, Section VI evaluates its performance by means of simulation. In particular, the evaluation includes 1) comparison against a lower bound (LB) of optimal maximum likelihood (ML) decoding and 2) investigation of impact of imperfect CSI on the decoding performance. Section VII concludes the paper. A brief introduction of EP is presented in Appendix A.

**Notations:** In general, letters in normal type stand for scalars. The letter η is reserved for \(\sqrt{-1}\). Letters in bold type represent vectors and matrices, where \(I_N\) represents the \(N \times N\) identity matrix. Calligraphic letters are used to denote sets. The cardinality of a discrete set \(A\) is denoted as \(|A|\).

The order of message passing influences not only the decoding performance, but also the degree of parallelism in computation. Taking advantage of spectrally-localized correlation in GFDM, this paper proposes a hybrid scheduling mechanism, showing near-optimum decoding performance can be achieved with a high level of parallelism in computation.

In summary, this paper contributes with the following aspects. First, we introduce the FD model of the considered MIMO-GFDM system in Section II. Based on this model, Section III provides an information theoretic analysis of the MIMO-GFDM system. The detailed derivation of the EP-based iterative equalization and detection algorithm is presented in Section IV, followed by remarks on its complexity and scheduling in Section V. Finally, Section VI evaluates its performance by means of simulation. In particular, the evaluation includes 1) comparison against a lower bound (LB) of optimal maximum likelihood (ML) decoding and 2) investigation of impact of imperfect CSI on the decoding performance. Section VII concludes the paper. A brief introduction of EP is presented in Appendix A.

The order of message passing influences not only the decoding performance, but also the degree of parallelism in computation. Taking advantage of spectrally-localized correlation in GFDM, this paper proposes a hybrid scheduling mechanism, showing near-optimum decoding performance can be achieved with a high level of parallelism in computation.

In summary, this paper contributes with the following aspects. First, we introduce the FD model of the considered MIMO-GFDM system in Section II. Based on this model, Section III provides an information theoretic analysis of the MIMO-GFDM system. The detailed derivation of the EP-based iterative equalization and detection algorithm is presented in Section IV, followed by remarks on its complexity and scheduling in Section V. Finally, Section VI evaluates its performance by means of simulation. In particular, the evaluation includes 1) comparison against a lower bound (LB) of optimal maximum likelihood (ML) decoding and 2) investigation of impact of imperfect CSI on the decoding performance. Section VII concludes the paper. A brief introduction of EP is presented in Appendix A.

The order of message passing influences not only the decoding performance, but also the degree of parallelism in computation. Taking advantage of spectrally-localized correlation in GFDM, this paper proposes a hybrid scheduling mechanism, showing near-optimum decoding performance can be achieved with a high level of parallelism in computation.

In summary, this paper contributes with the following aspects. First, we introduce the FD model of the considered MIMO-GFDM system in Section II. Based on this model, Section III provides an information theoretic analysis of the MIMO-GFDM system. The detailed derivation of the EP-based iterative equalization and detection algorithm is presented in Section IV, followed by remarks on its complexity and scheduling in Section V. Finally, Section VI evaluates its performance by means of simulation. In particular, the evaluation includes 1) comparison against a lower bound (LB) of optimal maximum likelihood (ML) decoding and 2) investigation of impact of imperfect CSI on the decoding performance. Section VII concludes the paper. A brief introduction of EP is presented in Appendix A.
the $n_{t}$th transmit and $n_{r}$th receive antenna is expressed as

$$h_{n_{t},n_{r}}[n] = \sum_{i=0}^{L_{ch}-1} h_{v,n_{t},n_{r}} \delta[n - v]. \tag{6}$$

For each antenna pair, the $L_{ch}$ path gains $\{h_{v,n_{t},n_{r}}\}$ are zero mean independent complex Gaussian random variables with the total power normalized to one, i.e., $\sum_{v=0}^{L_{ch}-1} E[\{h_{v,n_{t},n_{r}}\}^2] = 1$. To combat the multipath fading channel, a CP with length $N_{CP} \geq L_{ch}$ is appended to the head of each GFDM block. After CP removal at the receiver, the $N$ samples obtained at the $n_{r}$th receive antenna are equal to

$$r_{n_{r}}[n] = \sum_{n_{t}=1}^{N_{t}} s_{n_{t}}[n] \otimes h_{n_{t},n_{r}}[n] + w_{n_{r}}[n], \tag{7}$$

where $w_{n_{r}}[n]$ is spatially and temporally white proper Gaussian noise with variance $\sigma_w^2$. Circular convolution $\otimes$ in above rather than linear convolution is due to the use of CP.

By applying $N$-point discrete Fourier transform (DFT) of the received $N$ samples per receive antenna, the time domain (TD) input-output relation in (7) is converted to a FD representation

$$R_{n_{r}}[\nu] = \sum_{n_{t}=1}^{N_{t}} S_{n_{t}}[\nu] H_{n_{t},n_{r}}[\nu] + W_{n_{r}}[\nu], \tag{8}$$

where $R_{n_{r}}[\nu]$, $S_{n_{t}}[\nu]$, $H_{n_{t},n_{r}}[\nu]$ and $W_{n_{r}}[\nu]$ are DFTs of $r_{n_{r}}[n]$, $s_{n_{t}}[n]$, $h_{n_{t},n_{r}}[n]$ and $w_{n_{r}}[n]$, respectively. In particular, the noise $W_{n_{r}}[\nu]$ in FD remains to be white and follows the Gaussian distribution, i.e., $CN(0, N\sigma_w^2)$. By the definition of $s_{n_{t}}[n]$ given in (5), the GFDM block in FD is expressed as

$$S_{n_{t}}[\nu] = \sum_{n=0}^{N-1} s_{n_{t}}[n] e^{-j \frac{2\pi \nu n}{N}},$$

$$= \sum_{k=-K_{m}/2}^{K_{m}/2-1} \sum_{m=0}^{M-1} G(\nu - kM) \times d_{n_{t},k,m} e^{-j \frac{2\pi \nu m}{N}} \tag{9}$$

with $G[\nu] \overset{\Delta}{=} \sum_{n=0}^{N-1} g[n] e^{-j \frac{2\pi \nu n}{N}}$. Substituting (9) back into (8), we reach to

$$R_{n_{r}}[\nu] = \sum_{n_{t}=1}^{N_{t}} \sum_{k=-K_{m}/2}^{K_{m}/2-1} \sum_{m=0}^{M-1} d_{n_{t},k,m} H_{n_{t},n_{r}}[k,m] + W_{n_{r}}[\nu] \tag{10}$$

with the effective channel gain in FD defined as $\tilde{H}_{n_{t},n_{r}}[k,m,\nu] = e^{-j \frac{2\pi \nu k}{N}} G(\nu - kM) H_{n_{t},n_{r}}[\nu]$.

For notational convenience, we respectively concatenate the data symbols $\{d_{n_{t},k,m}\}$ transmitted via $N_{r}$ transmit antennas, the observations $\{R_{n_{r}}[\nu]\}$, and noise $\{W_{n_{r}}[\nu]\}$, into column vectors, i.e., $\mathbf{d}_{k,m} \in \mathbb{C}^{N_{t}, \nu}$, $\mathbf{R}[\nu] \in \mathbb{C}^{N_{r}, \nu}$ and $\mathbf{W}[\nu] \in \mathbb{C}^{N_{r}, \nu}$. Analogously, the channel gains in FD $\{H_{n_{t},n_{r}}[\nu]\}$ are formed into a matrix $\mathbf{H}[\nu] \in \mathbb{C}^{N_{r}, \nu \times \nu}$. We also define $\hat{\mathbf{H}}_{k,m}[\nu] = e^{-j \frac{\pi \nu k}{N}} G(\nu - kM) \mathbf{H}[\nu] \in \mathbb{C}^{N_{r}, \nu \times \nu} \tag{11}$

And then, the input-output relations with respect to all receive antennas can be written as

$$\mathbf{R}[\nu] = \sum_{k=\pm K_{m}/2}^{K_{m}/2-1} \sum_{m=0}^{M-1} \hat{\mathbf{H}}_{k,m}[\nu] \mathbf{d}_{k,m} + \mathbf{W}[\nu] \tag{12}$$

The received samples are subject to ICI, ISI and IAI. In order to have low OOB emission, it is desirable to limit ICI to few adjacent subcarriers. To this end, we shall select a pulse shaping filter with good spectral localization. This means the corresponding $G(\nu - kM) \delta$ equals zero at most values of $k$ for an arbitrary frequency bin $\nu$. Fig. 1 shows an example, in which $g(t)$ is constructed from the raised cosine (RC) filter with the roll-off factor $\alpha = 0.5$. By the definition of $\hat{\mathbf{H}}_{k,m}[\nu]$ given in (11), $G(\nu - kM) \delta$ equal to zero results in a zero-valued channel matrix. As such, eq. (12) reduces to

$$\mathbf{R}[\nu] = \sum_{k \in K(\nu)} \sum_{m=0}^{M-1} \hat{\mathbf{H}}_{k,m}[\nu] \mathbf{d}_{k,m} + \mathbf{W}[\nu], \tag{13}$$

where the set $K(\nu)$ contains the indices of the subcarriers yielding $G(\nu - kM) \delta \neq 0$. For pulse shaping filters with spectrum narrower than $[\frac{1}{4}, \frac{3}{4}]$, e.g., Fig. 1, the cardinality of $K(\nu)$ cannot exceed two, meaning ICI is limited to the very adjacent subcarrier on either left-hand side or right-hand side. For later use, $R$ and $H$ are used to compactly denote the observations $\{R[\nu]\}$, and channel matrices $\hat{\mathbf{H}}_{k,m}[\nu] \in \mathbb{C}^{N_{r}, \nu \times \nu}$.

III. INFORMATION THEORETIC ANALYSIS

Based on the system model presented above, this section analyzes GFDM from an information theoretic point of view. Namely, the capacity of the coded modulation (CM) system using GFDM and also the capacity of the binary-input continuous-output (BICO) memoryless channel as illustrated in Fig. 2 will be derived. Before proceeding to the detailed...
derivations, some remarks are necessary. First, the code bits \{c_t\} input to the channel are assumed to be i.i.d. uniform random variables. Second, we assume no CSI at the transmitter and perfect CSI at the receiver. Third, the rate loss due to CP removal is neglected. Fourth, GFDM in a MIMO system is effectively a memoryless modulator (MOD) over an \(N_C N_t\) dimensional signal set. The duration of one GFDM block can be regarded as one channel use. Consider an ergodic channel, where codewords can extend over an arbitrary number of GFDM blocks. The information rates derived below can then be interpreted as ergodic capacities. Last, the BICO memoryless channel depicted in Fig. 2 is composed of \(N_C N_t N_{\text{bps}}\) parallel binary-input sub-channels. The ideal interleaver (II) acting as a switch randomly selects a sub-channel to transmit each code bit. The demodulator (DEM) in Fig. 2 maps the outputs of the MIMO channel to a sequence of \(L\)-values (i.e., each code bit). The demodulator (DEM) in Fig. 2 maps the outputs of the MIMO channel to a sequence of \(L\)-values (i.e., each code bit). The demodulator (DEM) in Fig. 2 maps the outputs of the MIMO channel to a sequence of \(L\)-values (i.e., each code bit). The demodulator (DEM) in Fig. 2 maps the outputs of the MIMO channel to a sequence of \(L\)-values (i.e., each code bit).

\[
\lambda_{\text{opt},j} = \log \left[ \frac{P(c_t = 1 | \mathcal{R}, \tilde{H})}{P(c_t = 0 | \mathcal{R}, \tilde{H})} \right] - \log \left[ \frac{P(c_t = 1)}{P(c_t = 0)} \right],
\]

where \(P(c_t)\) represents the prior distribution of \(c_t\).

**A. CM Capacity**

As can be seen in (15) and (16) at the top of next page, the CM capacity for the MIMO-GFDM system is given by the conditional average mutual information (MI) between the data symbol vectors per MIMO-GFDM block and the MIMO channel output. In both (15) and (16), the vector \(d_t \in \mathcal{X}^{N_t M}\) is obtained by concatenating the \(M\) data symbol vectors \(d_k, m \in \mathcal{X}^{N_t}\) on the subcarrier \(k\) one after another and \(\{d_k\}\) are i.i.d. uniform random variables. Here the CM capacity is normalized by \(K_{\text{CM}} M\), since one MIMO-GFDM block carries \(K_{\text{CM}} M\) data symbol vectors. For evaluating (16), the most computational intensive part is to marginalize the likelihood function \(p(\mathcal{R} | \{d_k\}, \tilde{H})\) over the \(K_{\text{CM}}\) data vectors. For an orthogonal waveform, e.g., OFDM, the likelihood function can be factorized into a form such that each factor function only depends on a single data symbol vector \(d_k, m \in \mathcal{X}^{N_t}\). Relying on such factorization, the computational complexity of marginalization is exponential only in the number \(N_t N_{\text{bps}}\) of bits per transmitted data symbol vector. On contrary, GFDM is subject to ICI and ISI. Its factor functions can consist of more than one data symbol vector, leading to a significant increment on the computational complexity. The following part contributes to build a trellis diagram and apply a modified Viterbi algorithm for marginalization with complexity \(O(K_{\text{CM}}^{2N_t N_{\text{bps}}})\). Additionally, we derive an upper bound (UB) on \(C_{\text{CM}}\) by assuming perfect interference cancellation. Since its complexity reduces to \(O(K_{\text{CM}}^{2N_t N_{\text{bps}}} - 1)\), it can be used as an approximate solution when the trellis-based approach becomes unaffordable.

1) **Trellis-based Approach:** Fig. 1 illustrates that ICI is limited to the very adjacent subcarrier either on the left-hand side or the right-hand side. This implies the set \(\mathcal{K}(v)\) in (13) can only have three possibilities, i.e., \(\{k\}\), \(\{k-1,k\}\) and \(\{k,k+1\}\). Under this identification, we can group the observation vectors and corresponding channel matrices into two different cases: \(\{\mathcal{R}_k \cdot \tilde{H}_k\}\) and \(\{\mathcal{R}_{k+1} \cdot \tilde{H}_{k+1}\}\). Here, \(\mathcal{R}_k\) and \(\mathcal{R}_{k+1}\) are the compact representation of all observation vectors that only depend on the data vector \(d_k \in \mathcal{X}^{N_t M}\) modulated on the \(k\)th subcarrier and \(\tilde{H}_k\) compactly denotes the relevant channel matrices. Analogously, \(\mathcal{R}_{k+1}\) and \(\tilde{H}_{k+1}\) are associated to the data vectors \(d_k\) and \(d_{k+1}\). In terms of \(\mathcal{R}_{k} \cdot \tilde{H}_k\) and \(\mathcal{R}_{k+1} \cdot \tilde{H}_{k+1}\), the likelihood function \(p(\mathcal{R} | \{d_k, m\}, \tilde{H})\) can be factorized as (17). Note that besides the data vectors on the active subcarriers \(k = -K_{\text{CM}}/2, \ldots, K_{\text{CM}}/2 - 1\) the above factorization also involves the data vector \(d_{K_{\text{CM}}/2}\) associated to the subcarriers \(K_{\text{CM}}/2\). If \(K_{\text{CM}} < K\), the unused subcarrier \(K_{\text{CM}}/2\) implies \(d_{K_{\text{CM}}/2}\) equal to zero vector. If all subcarriers are used, \(d_{K_{\text{CM}}/2}\) becomes equivalent to \(d_{K_{\text{CM}}/2}\) by noting \(\langle K_{\text{CM}}/2\rangle K = \langle -K_{\text{CM}}/2\rangle K\) when \(K = K_{\text{CM}}\).

First, let us consider the situation \(K_{\text{CM}} < K\). Fig. 3(a) depicts a trellis diagram in accordance with (17). It has \(K_{\text{CM}} + 1\) stages. On the first \(K_{\text{CM}}\) stages, each realization of the data vector \(d_k \in \mathcal{X}^{N_t M}\) defines one state whose metric is calculated from \(p(\mathcal{R}_{k} | \tilde{H}_k, d_k)\). The transition between states on two consecutive stages is quantified by \(p(\mathcal{R}_{k,k+1} | \tilde{H}_{k,k+1}, d_{k+1}, d_k)\). The trellis is terminated at the zero state on the final stage, representing the unique realization of \(d_{K_{\text{CM}}/2}\). On top of the trellis diagram, the standard Viterbi algorithm identifies the sequence of data vectors \(d_k\) that maximizes the likelihood function by always selecting the maximum of all messages incoming to each state. In order to marginalize the likelihood function over all data vectors, we modify the Viterbi algorithm such that the sum of all messages incoming to a state rather than the maximum is computed. After traversing all stages in the trellis, the value attained at the final stage is the result. Plugging it into (16), the remaining terms for evaluating \(C_{\text{CM}}\) are easy to compute. The computational complexity of the
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2482479, IEEE Transactions on Wireless Communications

\[ C''_{cm} \triangleq \frac{1}{K_{on}M} I\{\{d_k\}; R|\tilde{H}\} = \frac{1}{K_{on}M} \left[ h(\{d_k\}) - h(\{d_k\}|R, \tilde{H}) \right] \]

\[ = N_t N_{bps} \left[ \log_2 \left( \frac{\sum_{a_{K_{on}t} e \in \mathcal{X}_{K_{on}t}^{M}} \sum_{a_{K_{on}t-1} e \in \mathcal{X}_{K_{on}t-1}^{M}} p(\mathcal{R}|\{d_k\}, \tilde{H})}{p(\mathcal{R}|\{d_k\}, \tilde{H})} \right) \right]. \]

\[ p(\mathcal{R}|\{d_k\}, \tilde{H}) = \prod_{k=-K_{on}/2}^{K_{on}/2-1} p \left( \{R_k^{[t]}|\tilde{H}_k^{[t]}\}, \{d_k\} \left| \{R_{k+1}^{[t]}|\tilde{H}_{k+1}^{[t]}\}, \{d_k\}, \{d_{k+1}\} \right. \right). \]

\[ I(\{d_k\}; R|\tilde{H}) = \sum_{k=-K_{on}/2}^{K_{on}/2-1} I(\{d_k\}; R|\tilde{H}, d_{-K_{on}/2}, \ldots, d_{k-1}) \]

\[ \leq \sum_{k=-K_{on}/2}^{K_{on}/2-1} \left( I(\{d_k\}; R|\tilde{H}, d_{-K_{on}/2}, \ldots, d_{k-1}, d_{k+1}, \ldots, d_{K_{on}/2-1}) \right) \]

\[ \leq \sum_{k=-K_{on}/2}^{K_{on}/2-1} \sum_{m=0}^{M-1} I(\{d_k, m\}; R|\tilde{H}, \{d_k\}' \neq k, \{d_{k+1}\}' \neq k, \ldots, \{d_{k+m}\}' \neq k). \]

\[ \mathcal{R}_{k,m}[\nu] \triangleq R[\nu] - \sum_{k'=k}^{K_{on}/2} \sum_{m'=0}^{M-1} \tilde{H}_{k,m}[\nu]d_{k,m'} - \sum_{m'=0}^{M-1} \tilde{H}_{k,m}[\nu]d_{k,m'} = \tilde{H}_{k,m}[\nu]d_{k,m} + W[\nu], \]

where the effective channel matrix \( \tilde{H}_{k,m}[\nu] \) is non-zero for any \( \nu \in V(k) \).

Figure 3: Trellis diagram based on the factorization in (17).

modified Viterbi algorithm is linear in the number of states on each stage, i.e., exponential in \( N_t N_{bps} M \).

Next, we proceed to the situation \( K_{on} = K \) which implies \( d_{K_{on}/2} = d_{-K_{on}/2} \). Under such tail-biting situation, marginalizing the likelihood function over the \( K_{on} \) data vectors is equivalent to repeatedly applying the modified Viterbi algorithm \( 2N_t N_{bps} M \) times and then adding up the results. Each time the trellis traversal starts from one realization of \( d_{-K_{on}/2} \) and terminates at the identical realization on the final stage \( k = K_{on}/2 \), see Fig. 3(b).

2) Upper Bound: Next, an UB on \( C''_{cm} \) is derived from a set of equalities and inequalities listed in (18), where the equalities (a) and (c) are outcomes of applying the chain rule; and the inequalities (b) and (d) are because conditioning on extra data symbols increases information rates. In (18), the MI between \( d_{k,m} \) and \( \mathcal{R} \) is computed given the perfect knowledge of CSI and the rest transmit symbol vectors. This motivates the construction of an interference-free observation of \( d_{k,m} \), see (19). Defining \( \mathcal{R}_{k,m} \triangleq \{\mathcal{R}_{k,m}[\nu]\}_{\nu \in V(k)} \) and \( \tilde{H}_{k,m} \triangleq \{\tilde{H}_{k,m}[\nu]\}_{\nu \in V(k)} \), we obtain (20), where the equality (a) is because \( \mathcal{R}_{k,m} \) is deterministic given the knowledge of \( \mathcal{R}, \tilde{H}, \{d_{k'}\}' \neq k \) and \( \{d_{k,m'}\}' \neq m \); and the equality (b) is because apart from \( \mathcal{R}_{k,m} \) the conditional observation \( \mathcal{R} \) contains no more information of \( d_{k,m} \). Substituting (20) back into (18) and then replacing \( I(\{d_k\}; R|\tilde{H}) \) in (15) by the derived UB, we consequently reach to (21). The complexity for computing (21) grows exponentially with the number \( N_t \).
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2482479, IEEE Transactions on Wireless Communications

\[ I(d_{k,m}; R|\tilde{H}) > I(d_{k,m}; R|\tilde{H}) \]

For evaluating \( C_{bicm} \), we resort to the histogram approach presented in [17].

\[ C_{bicm}^\prime \leq N_t N_{bps} = bico \sum_{l=0}^{K_{on} M N_t N_{bps} - 1} I(c_l; \lambda_k,l) = N_t N_{bps} - bico \sum_{l=0}^{K_{on} M} l E \left\{ \log_2 \left( \frac{\sum_{d' \in X^N_{i(l)} \cup X^N_{i(l)}} p(R|d', \tilde{H}_{k,l})}{p(R|c_l, \tilde{H})} \right) \right\} \]

IV. EP FOR DETECTION OF MIMO-GFDM SIGNALS

In this section, we apply the EP algorithm for detection of MIMO-GFDM signals. It consists of 1) constructing a FG and 2) deriving the message update equations.
observation vectors and channel matrices in FD reads
\[
p(c, d_{k,m} | \mathcal{R}, \mathcal{H}) \propto \prod_{\nu=0}^{N-1} f_\nu(\{d_{k,m}\}_{k \in K(\nu), m}) \prod_{k,m} p(d_{k,m} | c_{k,m}) \mathcal{P}(c_{k,m}),
\]
where the factor function \(f_\nu(\{d\}_{k \in K(\nu), m})\) is defined as
\[
f_\nu(\{d\}_{k \in K(\nu), m}) \triangleq p(\mathbf{R}[\nu])\left\{\mathbf{H}_{k,m}[\nu], \mathbf{d}_{k,m}\}_{k \in K(\nu), m}\right\}
\]
\[
\propto e^{-\frac{1}{2} \| \mathbf{R}[\nu] - \sum_{k \in K(\nu), m} \mathbf{H}_{k,m}[\nu] \mathbf{d}_{k,m} \|^2}.
\]
Note that we now treat \(d_{k,m}\) as a vector of continuous variables. Its bijective relation to the code bit vector \(c_{k,m}\), i.e., the constellation constraint, is conveyed by the conditional distribution \(p(d_{k,m}|c_{k,m}) \propto \delta(d_{k,m} - \phi(c_{k,m}))\). Since the data symbol \(d_{n_t,k,m} \in \mathcal{X}\) transmitted per antenna is modulated independently, the Dirac delta function effectively consists of \(N_t\) factor functions
\[
\delta(d_{k,m} - \phi(c_{k,m})) = \sum_{n_t=1}^{N_t} \delta(d_{n_t,k,m} - [\phi(c_{k,m})]_{n_t}).
\]
The prior distribution of the code bit vector \(c_{k,m}\) can be further factorized as
\[
\mathcal{P}(c_{k,m}) = \prod_{n_t=1}^{N_t,N_{\text{bps}}} \mathcal{P}([c_{k,m}]_{n_t}).
\]
Fig. 4 is a graphical representation of the factorization in (27). It consists of \(K_{\text{con}}\) clusters, each of which corresponds to one active subcarrier. Due to the spectrally localized ICI, each cluster only interacts with its neighbors. Their interaction depends on the pulse shaping filter, see Fig. 1. Such localized coupling justifies an effective use of message passing algorithms. Taking the cluster \(k\) as an example, it consists of \(M\) data symbol vectors carried by the subcarrier \(k\) and their corresponding code bit vectors. It also consists of \(M\) factor nodes \(f_\nu\), representing the factor functions \(f_\nu(\{d_{k,m}\}_{k \in K(\nu), m})\) with \(\nu \in \{\nu | M + m'\}_{m'=-\lfloor(M-1)/2\rfloor, \ldots, 0, \ldots, \lfloor(M+1)/2\rfloor}\). Due to ISI, each factor node \(f_\nu\) is connected to all variable nodes \(\{d_{k,m}\}_{m=0, \ldots, M-1}\) in the same cluster. As a Dirac delta function, \(p(d_{k,m}|c_{k,m})\) is denoted by the factor node \(\delta_{k,m}\). The factor node \(\mathcal{P}(c_{k,m})\) represents the prior distribution of the code bit vector \(c_{k,m}\).

As a generalization of belief propagation (BP), EP allows to project beliefs onto an exponential family, see Appendix A. This is particularly useful to statistical inference problems in which BP suffers from computational intractability, e.g., when dealing with continuous random variables. Here we constrain the belief of \(d_{k,m}\) to lie in a family of proper Gaussians, i.e.,
\[
\mathcal{G} \triangleq \left\{ q(z) = \prod_{n_t=1}^{N_t} \mathcal{N}(z_{n_t};\mu_{n_t},\sigma_{n_t}^2), z \in \mathbb{C}^{N_t} \right\}.
\]
For code bit vectors, their beliefs are members of
\[
\mathcal{B} \triangleq \left\{ q(z) = \prod_{i=1}^{N_t,N_{\text{bps}}} \exp[\frac{\|z_i - \lambda_i\|^2}{1 + \exp(\lambda_i)}], z \in \{0,1\}^{N_t,N_{\text{bps}}} \right\},
\]
where the L-values \(\lambda_i \in \mathbb{R}\) are natural parameters.

A. ICI, ISI and IAI Equalization

This part derives the message update equation at the factor node \(f_\nu\) and further links it to the LMMSE-based equalizer, aiming to separate superimposed data streams.

Considering the family \(\mathcal{G}\) of proper Gaussians, it is convenient to parameterize the messages incoming to the factor node \(f_\nu\) as
\[
m_{d_{k,m} \rightarrow f_\nu}(d_{k,m}) \propto \prod_{n_t=1}^{N_t} \mathcal{N}\left(d_{n_t,k,m};\mu_{d_{n_t,k,m} \rightarrow f_\nu},\sigma_{d_{n_t,k,m} \rightarrow f_\nu}^2\right).
\]

---

**Figure 4:** Graphic illustration of the factorization given in (27).
where the mean vector \( \mu_{d_{k,m}} \) and covariance matrix \( \Sigma_{d_{k,m}} \) are equal to
\[
\begin{align*}
\mu_{d_{k,m}} &= \mu_{d_{k,m} \rightarrow f_{\nu}} + \Sigma_{d_{k,m} \rightarrow f_{\nu}} \mathbf{H}^H_{k,m}[\nu] \Sigma_{R,\nu}^{-1} \left( \mathbf{R}[\nu] - \sum_{k',t' \in \mathcal{K}(\nu),m'} \mathbf{H}_{k',t'}[\nu] \mu_{d_{k',t'} \rightarrow f_{\nu}} \right), \\
\Sigma_{d_{k,m}} &= \Sigma_{d_{k,m} \rightarrow f_{\nu}} - \Sigma_{d_{k,m} \rightarrow f_{\nu}} \mathbf{H}^H_{k,m}[\nu] \Sigma_{R,\nu}^{-1} \mathbf{H}_{k,m}[\nu] \Sigma_{d_{k,m} \rightarrow f_{\nu}}
\end{align*}
\]
with \( \Sigma_{R,\nu} \triangleq \sum_{k' \in \mathcal{K}(\nu),m'} \mathbf{H}_{k',t'}[\nu] \Sigma_{d_{k',t'} \rightarrow f_{\nu}} \mathbf{H}^H_{k',t'}[\nu] + N \sigma^2 \mathbf{I}_{N_t} \).

\[
m_{f_{\nu} \rightarrow d_{k,m}}(d_{k,m}) \triangleq \prod_{n=1}^{N_t} \mathcal{C} \mathcal{N} \left( d_{n,k,m}; [\mu_{d_{n,k,m}}]_{n}, [\Sigma_{d_{n,k,m}}]_{n} \right) / m_{d_{k,m} \rightarrow f_{\nu}}(d_{k,m}),
\]

\[
\propto \prod_{n=1}^{N_t} \mathcal{C} \mathcal{N} \left( d_{n,k,m}; \mu_{d_{n,k,m}}, \sigma^2_{d_{n,k,m} \rightarrow f_{\nu}} \right),
\]
where \( \mu_{d_{n,k,m}} \) and \( \sigma^2_{d_{n,k,m} \rightarrow f_{\nu}} \) are given as
\[
\begin{align*}
\mu_{d_{n,k,m}} &= \mu_{d_{n,k,m} \rightarrow f_{\nu}} + \mathbf{h}^H_{n,k,m}[\nu] \Sigma_{R,\nu}^{-1} \left( \mathbf{R}[\nu] - \sum_{k',t' \in \mathcal{K}(\nu),m'} \mathbf{H}_{k',t'}[\nu] \mu_{d_{k',t'} \rightarrow f_{\nu}} \right), \\
\sigma^2_{d_{n,k,m} \rightarrow f_{\nu}} &= \mathbf{h}^H_{n,k,m}[\nu] \Sigma_{d_{n,k,m} \rightarrow f_{\nu}} \mathbf{h}_{n,k,m}[\nu] - \sigma^2_{d_{n,k,m} \rightarrow f_{\nu}}
\end{align*}
\]
with \( \mathbf{h}_{n,k,m}[\nu] \) standing for the \( n \)th column of the effective channel matrix \( \mathbf{H}_{k,m}[\nu] \in \mathbb{C}^{N_t \times N_t} \).

For compact notation, we also introduce a mean vector and a diagonal covariance matrix
\[
\begin{align*}
\mu_{d_{k,m} \rightarrow f_{\nu}} &\triangleq \left[ \mu_{d_{1,k,m} \rightarrow f_{\nu}}, \cdots, \mu_{d_{N_t,k,m} \rightarrow f_{\nu}} \right], \\
\Sigma_{d_{k,m} \rightarrow f_{\nu}} &\triangleq \text{diag} \left( \left[ \sigma^2_{d_{1,k,m} \rightarrow f_{\nu}}, \cdots, \sigma^2_{d_{N_t,k,m} \rightarrow f_{\nu}} \right] \right)
\end{align*}
\]
In terms of \( m_{f_{\nu} \rightarrow d_{k,m}}(d_{k,m}) \), the message outgoing from \( f_{\nu} \) computed according to (A.1) is expressed as (35). In (35), the marginal with respect to \( d_{k,m} \) is proportional to a Gaussian function as given in (36). Projecting the Gaussian function onto the family \( \mathcal{G} \), eq. (A.10) yields (38). Plugging (33) into (38) and then following (A.8), we obtain (39).

An intuitive way to derive (36) is presented as follows. By the definition of the factor function \( f_{\nu}(.) \) in (28), it is the likelihood function of \( \{d_{k,m}\}_{k \in \mathcal{K}(\nu),m} \) under a Gaussian channel. The product of all incoming messages can be treated as the prior distribution of \( \{d_{k,m}\}_{k \in \mathcal{K}(\nu),m} \). Proper normalization, the product of the likelihood function and prior distribution yields a posterior distribution of \( \{d_{k,m}\}_{k \in \mathcal{K}(\nu),m} \), which is Gaussian as well. Under this identification, the Gauss-Markov Theorem [19, pp.391] can lead us to (36). Moreover, the mean vector \( \mu_{d_{k,m}} \) in (37) is an LMMSE estimate and its error covariance matrix is \( \Sigma_{d_{k,m}} \). In other words, LMMSE equalization has been adopted for resolving ISI, ICI and IAI. However, the result is not directly applied to generate the message outgoing from \( f_{\nu} \) to \( d_{k,m} \). According to its update equation in (35), the obtained posterior distribution of \( d_{k,m} \) is divided by its prior distribution. In the literature of message passing algorithms, this step has been understood as a way of removing the prior information from the posterior distribution. One well-known instance under the framework of BP is the so-called extrinsic information exchange in iterative decoding. In our context, the message \( m_{f_{\nu} \rightarrow d_{k,m}}(d_{k,m}) \) therefore represents an extrinsic information of \( d_{k,m} \).

**B. Soft Detection**

This part derives the message update equations at the factor node \( \delta_{k,m} \). From \( \delta_{k,m} \), messages propagate in two directions, namely towards the variable node \( d_{k,m} \) and \( c_{k,m} \). Since the factor node \( \delta_{k,m} \) captures the constellation constraint, the former one effectively conveys this information to the equalization nodes \( f_{\nu} \). On the contrary, the latter one aims to produce the L-values for code bits. Therefore, it corresponds to the task of soft detection.

1) **Message Propagation from \( \delta_{k,m} \) to \( d_{k,m} \):** According to (A.1), the message update equation for \( m_{\delta_{k,m} \rightarrow d_{k,m}}(d_{k,m}) \) is written as (41). The message incoming from \( d_{k,m} \) equals the product of messages outgoing from the equalization nodes \( \{f_{\nu}\} \) that are connected to \( d_{k,m} \), i.e., any \( f_{\nu} \) with \( \nu \in \mathcal{V}(k) \).

\[
m_{d_{k,m} \rightarrow \delta_{k,m}}(d_{k,m}) = \prod_{\nu \in \mathcal{V}(k)} m_{f_{\nu} \rightarrow d_{k,m}}(d_{k,m}).
\]
Based on (39), the product of Gaussians \( m_{\delta_k,m}\rightarrow d_{k,m}(d_{k,m}) \) is still a Gaussian, i.e., (43). The message \( m_{c_{k,m}}\rightarrow \delta_{k,m}(c_{k,m}) \) represents the prior distribution of the code bits

\[
m_{c_{k,m}}\rightarrow \delta_{k,m}(c_{k,m}) \propto \prod_{i=1}^{N_iN_{hit}} P(\{c_{k,m}\}) \]

\[
= \prod_{i=1}^{N_iN_{hit}} \frac{\exp(\{c_{k,m}\}|\lambda^{[e]}_{k,m,i})}{1 + \exp(\{c_{k,m}\}|\lambda^{[e]}_{k,m,i})},
\]

where \( \lambda^{[e]}_{k,m,i} \) is the prior L-value equal to \( \log \frac{P(\{c_{k,m}\}|\lambda^{[e]}_{k,m,i})}{P(\{c_{k,m}\}|\lambda^{[e]}_{k,m,i})} \). Based on (43), (44) and (29), we can easily compute the projection in (41) by following (A.7). Denoting the outcome as \( m_{\delta_k,m}\rightarrow d_{k,m}(d_{k,m}) \) we obtain

\[
m_{\delta_k,m}\rightarrow d_{k,m}(d_{k,m}) = \prod_{i=1}^{N_i} \mathcal{N}(d_{k,m}; \mu_{\delta_{k,m},m} \rightarrow \delta_{k,m}, \sigma^2_{\delta_{k,m},m} \rightarrow \delta_{k,m}, m_{c_{k,m}} \rightarrow \delta_{k,m}(c_{k,m}))
\]

Based on (A.8), the division of two Gaussians in (41) yields (46). Note that we may encounter the “negative variance” problem [7] when computing (47). In the literature, several heuristic approaches, e.g., in [7, 10, 14], have been presented to tackle this problem. Here, we adopt the approach in [10] because of its connection to the conventional LLMSSE-based equalizer working with posterior feed-

back [20]. Whenever the variance \( \sigma^2_{\delta_{k,m},m} \rightarrow \delta_{k,m}, m \) is negative, we simply set \( \mu_{\delta_{k,m},m} \rightarrow \delta_{k,m} = \mu_{\delta_{k,m},m} \rightarrow \delta_{k,m} \) and \( \sigma^2_{\delta_{k,m},m} \rightarrow \delta_{k,m} = \sigma^2_{\delta_{k,m},m} \rightarrow \delta_{k,m} \). This is equivalent to omit the division in (45), meaning the posterior rather than extrinsic information is propagated to the equalization nodes \( f_{\delta} \).

2) Message Propagation from \( \delta_{k,m} \) to \( c_{k,m} \): The message update equation for \( m_{\delta_{k,m}} \rightarrow d_{k,m}(c_{k,m}) \) is given by (48). Plugging (43), (44) and (29) into (48), we can use the parameter \( \lambda^{[e]}_{k,m,i} \) determined as (49) to parameterize the message

\[
m_{\delta_{k,m}} \rightarrow d_{k,m}(c_{k,m}) \propto \prod_{i=1}^{N_iN_{hit}} \exp(\{c_{k,m}\}|\lambda^{[e]}_{k,m,i}).
\]

From above, \( \lambda^{[e]}_{k,m,i} \) effectively corresponds to the extrinsic L-value of the \( i \)th code bit in the bit vector \( c_{k,m} \).

C. BICM Decoding

After several iterations between the equalization and detection nodes, we pass the resulting extrinsic L-values \( \lambda^{[e]}_{k,m,i} \) through the de-interleaver. The output is a sequence of L-values \( \lambda_i \) that is sorted in accordance with the code bit sequence \( c_i \). Based on \( \lambda_i \), we can perform BICM decoding [15] to retrieve the transmitted message.
D. BICM with Iterative Decoding (BICM-ID)

The output of decoder can be fed back to update the prior L-values of code bits, i.e., \( \{ \lambda[k,m] \} \). Relying on this new information, we can re-calculate the messages between the equalization and detection nodes. The presence of such feedback loop yields the structure of BICM-ID [21].

V. SCHEDULING AND COMPLEXITY ANALYSIS

In the framework of EP, the order of message passing, i.e., scheduling, is not specified. In particular for a cyclic graph, scheduling can impact not only the performance but also the convergence behavior. It has been empirically shown that updating messages in a sequential rather than parallel way yields better convergence results, e.g., in the context of low-density parity-check (LDPC) decoding [22]. However, one disadvantage of sequential scheduling is latency, as messages have to be updated one after another. This problem becomes even more severe if we have to refine each message several times. This section aims to introduce a hybrid scheduling scheme that can deliver good decoding performance for latency sensitive cases. On top of it, the computational complexity is analyzed.

A. Scheduling

The FG in Fig. 4 is divided into \( K_{on} \) clusters. The scheduling scheme presented below consists of two levels, namely, inter- and intra-scheduling. The task of inter-scheduling is to decide the activation order of the \( K_{on} \) clusters. Due to the coupling of adjacent clusters, we propose to first activate the clusters with even indices and then activate those with odd indices. By doing so, each cluster is able to use the newly updated messages from its neighbors. Equipped with sufficient computation resources, the even(odd) clusters can simultaneously accomplish their message updating. Within each cluster, a sequential scheduling scheme is adopted for intra-scheduling. Taking the cluster \( k \) in Fig. 4 as an example, the messages are first propagated from the bottom of the cluster \( k \) to the top and then propagated in the reverse direction:

1. \( P(c_{k,m}) \rightarrow c_{k,m} \rightarrow \delta_{k,m} \rightarrow d_{k,m} \) for \( m = 0, 1, \ldots, M-1 \).
2. for each \( f_{\nu} \) in the cluster do
   3. \( \{ d_{k,m} \}_{k \in K(\nu), m = 0, \ldots, M} \rightarrow f_{\nu} \rightarrow \{ d_{k,m} \}_{k \in K(\nu), m = 0, \ldots, M} \)
4. end for
5. \( d_{k,m} \rightarrow \delta_{k,m} \rightarrow c_{k,m} \rightarrow P(c_{k,m}) \) for \( m = 0, 1, \ldots, M-1 \).

Note that the computations at the step 1 and 5 for different \( m \) can be executed in parallel. With respect to the above-proposed hybrid scheduling scheme, we can estimate the processing time for \( J \) iterations. Assume we have the computation resources to activate \( \gamma \) clusters at the same time and the execution time of each cluster equals \( T_{cluster} \). Then, the total processing time for \( J \) iterations equals \( 2J \frac{\gamma}{2} | T_{cluster} \). In the following, we will link \( T_{cluster} \) to the most computational intensive task in one cluster.

B. Complexity of Equalization and Detection

Within each cluster, every message update effectively boils down to the calculation of the mean and variance of a Gaussian function. From Section IV, we can observe that most calculations only require a few number of arithmetic operations, except (40) in Section IV-A. In other words, the most computational intensive task within each cluster is to compute the messages outgoing from the factor nodes \( \{ f_{\nu} \} \) as illustrated in Fig. 4.

The complexity of (40) is mainly determined by the number of multiplications. Here, we assume one complex-valued multiplication has complexity \( O(1) \). Then, the total complexity for computing all messages outgoing from one factor node \( f_{\nu} \) is \( O(N_{c}^{2} + N_{c}^{2}N_{t} + N_{c}^{2}N_{i} + N_{c}M) \). In particular, the matrix inversion \( \Sigma_{R_{\nu}}^{-1} \in C^{N_{c} \times N_{c}} \) contributes to \( O(N_{c}^{3}) \), while the other matrix-vector operations yield \( O(N_{c}^{2}N_{t} + N_{c}^{2}N_{i} + N_{c}M) \). Since the number of factor nodes \( \{ f_{\nu} \} \) residing in one cluster is \( M \), the total complexity for \( K_{on} \) clusters equals \( O(K_{on}MN_{c}^{2} + K_{on}MN_{c}^{2}N_{t} + K_{on}MN_{c}^{2}N_{i} + K_{on}MN_{c}N_{c}M^{2}) \), which is linear in the number \( K_{on} \) of active subcarriers and quadratic in the number \( M \) of subsymbols.

Hardware implementation of the proposed algorithm is beyond the scope of this paper. However, the remark given at the end of Section IV-A has linked the computational intensive part of the message update equation in (40) to standard LMMSE equalization. Such connection indicates that we can actually rely on existing energy efficient implementations of LMMSE equalization in the literature, e.g., [23]–[25] for both small and large scale MIMO systems. In addition, the execution time \( T_{LMMSE} \) of LMMSE filtering dominates the execution time \( T_{cluster} \) per cluster, meaning \( T_{cluster} \approx M T_{LMMSE} \).

VI. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed EP-based iterative equalization and detection algorithm in MIMO-GFDM systems with Gray-mapped modulation schemes. The obtained performance is also compared to that of conventional MIMO-OFDM systems. For all simulations, the Rayleigh fading channels associated to all transmit-receive antenna pairs are modeled as i.i.d. random processes, i.e., no spatial correlation is present. Each fading channel has a uniform power-delay profile characterized by the number \( L_{ch} \) of independent delay taps. The length of CP in both the GFDM and OFDM system is identical and greater than \( L_{ch} \). With the channel gain normalized to one, the signal-to-noise ratio (SNR) is defined as \( (N_{c}E_{s})/N_{0} \) where \( E_{s} \) is the transmit energy per symbol duration \( E_{s} = E(|d_{k,m}|^{2}) \) and the \( N_{0} \) represents the noise spectral density. The configuration for GFDM is given in Table I, where the corresponding setting for OFDM is listed as well.

In all plots, the labels, i.e., Par-EP, Seq-EP, Hyb-EP, denote the derived EP algorithm combined with parallel, sequential and hybrid scheduling, respectively. Hyb-EP has been described in Section V-A. In the context of sequential scheduling, the \( K_{on} \) clusters depicted in Fig. 4 are executed one after another. Par-EP simultaneously activates the \( K_{on} \) clusters. Within each cluster, the messages from the \( M \) factor nodes \( f_{\nu} \) are also computed in parallel. For OFDM systems, detection is conducted on subcarrier basis and the order of message passing among subcarriers is not a concern. Therefore, we simply
use the label $EP$. The well-known sphere decoding (SD) algorithm [26] achieving the maxlog maximum a posteriori (MAP) optimum is available for MIMO-OFDM. It is labeled as $SD$ in the corresponding plots.

A. Efficiency of $EP$

The BICM capacity $C_{bicm}'$ and BICO capacity $C_{bico}'$ are performance metrics independent of channel coding schemes. By comparing $C_{bico}'$ achieved by Hyb-EP with $C_{bicm}'$, or its UB, this part assesses the efficiency of Hyb-EP in approaching the optimum performance.

Fig. 5(a) considers a $2 \times 2$ MIMO-GFDM system. In this case, the BICM capacity can be evaluated by the trellis-based approach introduced in Section III. Analogous to the observation of applying EP for MIMO detection [10], EP shows near-optimum performance in the low and high SNR regions, e.g., below $0 \, \text{dB}$ and above $12 \, \text{dB}$. The observed $1 \, \text{dB}$ performance loss in the medium SNR region can be explained by noting the quality of Gaussian approximation. In the process of tailoring EP for detection of MIMO-GFDM signals, we have treated data symbols as continuous variables and constrain their beliefs to be Gaussians. Such approximation is good if the true beliefs tend to be uniform or peak at a single constellation point. On contrary, the existence of multiple peaks can yield a loose approximation. This situation often occurs when the channel condition is not good enough to generate an unequivocal decision. Fig. 5(a) also reveals GFDM benefits from large delay spread. The gain increases along with the SNR until the maximum normalized information rate is reached, i.e., $N_t N_{bps} = 2$. At low SNRs, the noise plus the interference determines the performance.

Fig. 5(b) resorts to an UB of the BICM capacity for a $4 \times 4$ MIMO system. Relying on it, we tend to estimate the performance loss of Hyb-EP in a pessimistic manner, meaning its true performance loss must be smaller than $4 \, \text{dB}$ as observed in Fig. 5(b). Since the Gaussian family $G$ used in the derivation of EP ignores the spatial correlation of data symbols for the sake of low complexity, it is expected to experience a larger performance loss in a system with more transmit antennas. Furthermore, we also observe the impact of the delay spread on $C_{bicm}'$ and the UB of $C_{bicm}'$ is consistent with that in Fig. 5(a).

B. OFDM vs. GFDM

For the comparison between GFDM and OFDM, we start from the BICM capacity, CM capacity, and their UBs as depicted in Fig. 6. Being more specific, Fig. 6(a) shows the BICM and CM capacities of GFDM and OFDM. OFDM is an orthogonal waveform. Ignoring the information loss due to CP removal, it is optimal to conduct MIMO detection on subcarrier basis. This means its ergodic capacities actually equal the sum of MIs with respect to each active subcarriers. Since the MIs depend on the statistics of channel matrices on individual subcarriers rather than the channel correlation between subcarriers [27], frequency selectivity does not affect the BICM and CM capacities of OFDM. In other words, both $C_{bicm}'$ and $C_{cm}'$ of OFDM depicted in Fig. 6(a) are invariant to the number of delay taps. On contrary, GFDM is a non-orthogonal waveform. The presence of ICI and ISI on the other hand, since the information of each data symbol is contained by multiple FD observations, it opens an opportunity to exploit frequency selectivity for an increased information rate. Therefore, we can observe the gain of GFDM over OFDM in Fig. 6(a) increases as the frequency selectivity of the experienced channel improves. This is particularly true at higher SNRs, since it becomes easier for the optimal receiver to resolve ICI and ISI. Fig. 6(b) illustrates the tightness of the UBs on $C_{bicm}'$ and $C_{cm}'$ in a $2 \times 2$ MIMO-GFDM system, where the channel has $64$ delay paths. Compared to the UB on $C_{cm}'$, the UB on $C_{bicm}'$ is looser. According to Section III, both UBs are derived by means of perfect interference cancellation. Since the BICM receiver ignores the correlation of code bits conditional on the received samples and CSI, the genie knowledge of interference plays a more critical role in enhancing $C_{bicm}'$ than $C_{cm}'$. Fig. 6(c) depicts the UBs of $C_{bicm}'$ and $C_{cm}'$ for a $4 \times 4$ MIMO system, where the trellis-based approach becomes too complex to evaluate $C_{cm}'$ and $C_{bicm}'$ exactly. Both of them show the dependence of frequency selectivity on the information rates, which is identical to that in Fig. 6(a).

Next, we compare both waveforms under the use of practical channel coding schemes. Assume one codeword per OFDM and GFDM block. The decoding performance is evaluated in the context of BICM decoding and BICM-ID, respectively.

1) BICM Decoding: Using the turbo code with polynomial generator $\{1, 15/13\}$, Fig. 7(a) shows the the coded bit error rates (BERs) achieved by $10$ turbo iterations and at two different code rates, i.e., $1/2$ and $5/6$.

In Fig. 7(a), we observe OFDM using SD delivers the best decoding performance. The performance gap between OFDM-SD and OFDM-EP enlarges as the code rate increases. Since turbo codes are capacity-achieving, this observation can be linked to their BICO capacities depicted in Fig. 7(b). The BICO capacity achieved by using SD can approach to the BICM capacity of OFDM. On the contrary, the performance of EP is degraded with the SNR ranging from $11 \, \text{dB}$ to $25 \, \text{dB}$ because of inaccurate Gaussian approximation. Therefore, it becomes suboptimal for achieving higher information rates.

2For every $10$ information bits input to the two identical and parallel-concatenated recursive convolutional codes, we select all systematic bits plus the first parity bit from the first component code and only keep the last parity bit from the second one.

<table>
<thead>
<tr>
<th></th>
<th>Nr. of subsymb. $(M)$</th>
<th>subcarrier spacing</th>
<th>DFT size $(K)$</th>
<th>active subcarriers $(K_{cm})$</th>
<th>pulse shaping flt.</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFDM</td>
<td>5</td>
<td>15·5 kHz</td>
<td>64</td>
<td>30</td>
<td>RC with $\alpha = 0.5$</td>
</tr>
<tr>
<td>OFDM</td>
<td>1</td>
<td>15 kHz</td>
<td>64·5</td>
<td>30·5</td>
<td>rect.</td>
</tr>
</tbody>
</table>

Table I: Configurations of GFDM and OFDM
Figure 5: Efficiency of the proposed Hyb-EP algorithm (after convergence), where the multipath fading channel has a uniform power delay profile

Figure 6: GFDM vs. OFDM in terms of BICM and CM capacities. Note that the curves related to OFDM are invariant under different numbers of delay taps in the considered channel model.
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2015.2482479, IEEE Transactions on Wireless Communications

However, from the perspective of hardware implementation, EP can still be a competitive option for MIMO detection. According to [23], one critical implementation issue of SD is its channel-dependent throughput. Good channel conditions are needed for achieving a throughput, comparable to that of LMMSE-based detection algorithms with iterative interference cancellation. Furthermore, the complexity of SD grows exponentially with the number of transmit antennas, while EP has polynomial complexity. This makes it particularly attractive when we extend to large-scale MIMO systems [12], [13].

For GFDM, SD is impractical due to ICI and ISI. The proposed EP-based algorithm is a suboptimal solution with manageable complexity. Due to ICI and ISI, GFDM requires two more EP iterations than OFDM. Fig. 7(a) depicts the BERs achieved by combining EP with different scheduling schemes. Hyb-EP and Seq-EP achieve similar BERs, while Par-EP is evidently worse. Since Hyb-EP supports a high level of parallelism, it is more suitable than Seq-EP for latency constrained applications. In Fig. 7(b), we can observe GFDM-Hyb-EP starts to outperform OFDM-EP at SNRs beyond 20 dB. At higher SNRs, it becomes easier to resolve interference. As a result, the benefit of exploiting frequency selectivity appears for GFDM at higher code rates, e.g., 5/6 in Fig. 7(a).

In [6], the LMMSE equalizer as well as the matched filter with successive interference cancellation (MF-SIC) were adopted for equalizing the GFDM signal in single antenna systems. Here, they are extended for MIMO-GFDM cases and labeled as LMMSE and MF-SIC in Fig. 7(a). In particular, the application of LMMSE equalization for jointly tackling ICI, IAI and ISI requires us to invert a matrix with dimension of $K_{on}MN_{c}$, which is impractical. Extending MF-SIC [6] for MIMO-GFDM, we first perform LMMSE FD equalization to resolve IAI and then apply MF-SIC [6] with four iterations to tackle the remaining ISI and ICI in the $N_{t}$ GFDM blocks. Since it treats IAI and ICI/ISI separately, the achieved decoding performance is worse than that of LMMSE and Hyb-EP.

Fig. 7(c) shows that the convergence behavior of turbo decoding improves as the number of Hyb-EP iterations increases. It also reveals more than one option to achieve a target BER. For instance, targeting the BER of $10^{-6}$ at 14 dB, we can choose either 3 Hyb-EP iterations plus 8 turbo iterations or 4 Hyb-EP iterations plus 5 turbo iterations. Since turbo decoding is known to be time-consuming, the latter may be preferred for a reduced overall decoding time.

2) BICM-ID: Using a rate-1/2 convolutional code with polynomials $[133, 171]_{8}$, the BCJR decoder [16] feeds back the extrinsic information of the code bits to the detection unit involved in the EP-based algorithm. The resulting BICM-ID receiver consists of three functional units, i.e., the equalization, detection and decoding unit. They form a doubly iterative receiver. The inner iteration is between the equalizer and detector and they together form the outer iteration loop with the decoder.

In this scenario, the comparison is made against the performance of ML decoding. Since ML decoding aims to find the codeword that maximizes the likelihood function, it achieves the minimum frame error rate (FER). Fig. 8 depicts LBs on the FER of ML decoding. They are obtained with the aid of the adopted BICM-ID receiver. Namely, the likelihood of the codeword recovered by the suboptimal receiver is compared to that of the transmitted codeword. If the former is larger, the ML decoder would yield a decoding failure. Otherwise, we assume the ML decoder can find the correct codeword. Under this assumption, we effectively overestimate the performance of the ML decoder. Therefore, the obtained FER is a LB. The tightness of such LB depends on the efficiency of the adopted suboptimal receiver in achieving the ML decoding performance. Using a near-optimum receiver, the gap is expected to be small.

When the SNR is larger than 10.5 dB, Fig. 8 shows the EP-based algorithm approaches the ML decoding performance in both systems. It is noted that the performance of OFDM is insensitive to the number of delay taps. Because of interleaving, correlated code bits are carried by subcarriers with sufficient spacing. For such rich multipath channels, its performance approaches to that of an i.i.d. fading channel case. On the contrary, GFDM permits ICI and ISI and the information of code bits spreads over multiple frequency bins. Exploiting frequency diversity, GFDM outperforms OFDM and the gain increases with the number of delay taps.

Figure 7: Coded BER of BICM decoding in a 4 × 4 MIMO 64-path Rayleigh fading channel with the uniform power delay profile.
Figure 8: Coded FERs of BICM-ID in a $4 \times 4$ MIMO multipath Rayleigh fading channel with the uniform power delay profile.

Figure 9: Impact of imperfect CSI on the decoding performance, where the system setup is identical to that for generating Fig. 8(a).

C. Impact of Imperfect Channel Knowledge

This part studies the impact of imperfect CSI on the performance of the proposed algorithm. We exemplarily adopt a preamble-based LMMSE channel estimator to estimate CSI for subsequent data blocks under the same channel realization and SNR. The adopted preamble is made from $N_t$ orthogonal pilot vectors that are alternately and periodically inserted in FD. Due to imperfect CSI, Fig. 9(a) reveals performance degradation. The SNR loss can be analyzed by reformulating the channel input-output relation in (13)

$$R[\nu] = \sum_{k \in K(\nu)} \sum_{m=0}^{M-1} H_{k,m}[\nu] d_{k,m} + \sum_{k \in K(\nu)} \sum_{m=0}^{M-1} (H_{k,m}[\nu] - \tilde{H}_{k,m}[\nu]) d_{k,m} + W[\nu],$$

where $\tilde{H}_{k,m}[\nu]$ is the estimated channel matrix. Interpreting channel estimation error as an additional noise term, we can compute the effective SNR by analogy with the approach in [28]. Fig. 9(b) shows a good match to the decoding performance in Fig. 9(a).

VII. CONCLUSION

Applications for 5G networks will demand outstanding performance and flexibility from the physical layer. Achieving higher capacity and better BER performance under MIMO multipath fading channels is mandatory for future wireless systems. An advanced receiver with manageable complexity and latency can strengthen the benefits of using the flexible non-orthogonal waveform GFDM in 5G networks. This work has modeled the input-output relation of a MIMO-GFDM system in FD. On this basis, we have analyzed constellation constrained GFDM from an information theoretic perspective.
By permitting the presence of interference, the results have demonstrated that GFDM can be more efficient than OFDM in achieving higher data rates in frequency selective fading channels. On the other hand, the presence of interference challenges the design of optimum receivers. We therefore have applied the EP algorithm for achieving near-optimum decoding performance. The resulting iterative solution consists of an LMMSE based equalization unit and a soft detection unit. On top of the receiver structure, we have subsequently proposed a scheduling scheme. It allows for a high-level form of parallel computation, desirable for latency constrained applications. The overall computational complexity is linear in the number of active subcarriers and quadratic in the number of subbands. By means of simulation, GFDM adopting the EP-based iterative receiver has finally achieved better decoding performance than OFDM using the optimal ML decoder in a rich multipath environment.

ACKNOWLEDGEMENTS

The computations were performed on a computing cluster at the Center for Information Services and High Performance Computing (ZIH) at TU Dresden. The authors thank the FP7 project ICT-318555 “5GNOW” for partially funding the work. This work was also supported in part by CNPq-Brasil and Finer/Initiell/Radiocommunication Reference Center under Grant No. 01.14.0231.00 hosted by Inatel.

APPENDIX A

EP FOR APPROXIMATE BAYESIAN INFERENCE

This section presents a brief introduction of EP. Here, we assume readers are familiar with basic concepts of FG and BP. A comprehensive introduction of them can be found in [29].

Analogous to BP, EP can be visualized as a message passing algorithm that operates on a FG. There are two rules for message passing, one for each direction (see Fig. A.1). Assume a vector \( x \) of variables and its subvector \( x_v \) is the argument of a factor function \( f_v(x_v) \). The set \( K(v) \) contains the indices of entries in \( x \) that are also entries of \( x_v \). Then, the message from the factor node \( f_v \) representing \( f_v(x_v) \) to the variable node \( x_k \) with \( k \in K(v) \) is calculated as

\[
m_{f_v \rightarrow x_k}(x_k) = \frac{\text{proj}_G \left\{ \frac{1}{Z} \text{argmax} \left\{ f_v(x_v) \prod_{k' \in K(v)} m_{x_{k'} \rightarrow f_v}(x_k') \right\} \right\}}{m_{x_k \rightarrow f_v}(x_k)},
\]

where \( Z \) is a normalization constant and \( G \) here represents an exponential family. If ignoring the projection step (or the argument of the projection is already an element in \( Q \)), the message update rule in (A.1) is completely identical to that of BP. The key use of projection here is to ensure the messages can remain manageable forms during the procedure of message passing. The update rule for messages propagating in the reverse direction is completely identical to BP

\[
m_{x_k \rightarrow f_v}(x_k) = \prod_{\nu' \in V(k), \nu' \neq \nu} m_{f_{\nu'} \rightarrow x_k}(x_k),
\]

where the set \( V(k) \) contains the indices of factor nodes that are connected to \( x_k \). The product of messages from both directions yields a belief of \( x_k \)

\[
q(x_k) = m_{f_v \rightarrow x_k}(x_k)m_{x_k \rightarrow f_v}(x_k).
\]

It is a member of the exponential family \( Q \). With proper normalization, exponential families are closed under multiplication. Therefore, it is convenient to presume the message \( m_{f_v \rightarrow x_k}(x_k) \) and \( m_{x_k \rightarrow f_v}(x_k) \) after normalization lie in \( Q \). In many applications, it is enough to work with unnormalized distributions. For more details of EP, we refer readers to [30].

Consider a special case, in which \( Q \) is chosen to be the family \( G \) of proper Gaussians. The message update equation given in (A.1) becomes

\[
m_{f_v \rightarrow x_k}(x_k) = \frac{1}{m_{x_k \rightarrow f_v}(x_k)} \mathcal{CN}(x_k; \mu_{x_k + f_v}, \sigma^2_{x_k + f_v}),
\]

where the Gaussian function is the outcome of projection

\[
\mathcal{CN}(x_k; \mu_{x_k + f_v}, \sigma^2_{x_k + f_v}) \triangleq \text{proj}_G \left\{ \frac{1}{Z} \text{argmax} \left\{ f_v(x_v) \prod_{k' \in K(v)} m_{x_{k'} \rightarrow f_v}(x_k') \right\} \right\}.
\]

Here, we use \( \leftrightarrow \) in the notion of the mean \( \mu_{x_k + f_v} \) and variance \( \sigma^2_{x_k + f_v} \) since the Gaussian function parameterized by them is the product of the messages from the direction \( f_v \rightarrow x_k \) and \( x_k \rightarrow f_v \), see (A.4). The message \( m_{f_v \rightarrow x_k}(x_k) \) and \( m_{x_k \rightarrow f_v}(x_k) \) are parameterized as

\[
\begin{align*}
m_{f_v \rightarrow x_k}(x_k) &\propto \mathcal{CN}(x_k; \mu_{x_k + f_v}, \sigma^2_{x_k + f_v}) \\
m_{x_k \rightarrow f_v}(x_k) &\propto \mathcal{CN}(x_k; \mu_{x_k + f_v}, \sigma^2_{x_k + f_v}).
\end{align*}
\]

By doing so, the message update equations in (A.2) and (A.4) are converted to the update equations for the mean and variance associated to each message. Specifically, the argument of \( \text{proj}_G \{ \cdot \} \) in (A.5) is a distribution function depending on \( \{ \mu_{x_k + f_v}, \sigma^2_{x_k + f_v} \} \). Denoting it as \( p(x_k) \), the outcome of projection is

\[
\begin{align*}
\mu_{x_k + f_v} &= \int_{x_k} x_k p(x_k) dx_k \\
\sigma^2_{x_k + f_v} &= \int_{x_k} |x_k - \mu_{x_k + f_v}|^2 p(x_k) dx_k.
\end{align*}
\]

In terms of \( \mu_{x_k + f_v} \) and \( \sigma^2_{x_k + f_v} \), the update equation in (A.4) boils down to

\[
\begin{align*}
\mu_{f_v \rightarrow x_k} &= \mu_{x_k + f_v} - \frac{\mu_{x_k + f_v} - \mu_{x_k - f_v}}{\sigma^2_{x_k + f_v}} \\
\sigma^2_{f_v \rightarrow x_k} &= \frac{\sigma^2_{x_k + f_v} - \sigma^2_{x_k - f_v}}{\sigma^2_{x_k - f_v}}.
\end{align*}
\]

By analogy, under the choice of \( Q = G \), the update equation in (A.2) becomes

\[
\begin{align*}
\sigma^2_{x_k \rightarrow f_v} &= \left( \sum_{\nu' \in V(k), \nu' \neq \nu} \sigma^2_{f_{\nu'} \rightarrow x_k} \right)^{-1} \\
\mu_{x_k \rightarrow f_v} &= \frac{\sigma^2_{x_k \rightarrow f_v}}{\sum_{\nu' \in V(k), \nu' \neq \nu} \sigma^2_{f_{\nu'} \rightarrow x_k}}.
\end{align*}
\]

Apart from projection into the Gaussian family \( G \), another one that is used in this paper is to project a multivariate...
distribution $p(x)$ into the set $\mathcal{F}$ of fully factorized distributions, i.e., $\mathcal{F} = \{ q(x) : q(x) = \prod_k q(x_k) \}$. The outcome is the product of the marginals of $p(x)$, i.e.,

$$\prod_{\mathcal{F}} \{ p(x) \} = \prod_k p(x_k). \quad (A.10)$$

**REFERENCES**


Dan Zhang received the B.Sc. degree in electrical engineering from Zhejiang University, China in 2004, and both the M.Sc. and the Dr.-Ing. degree in Electrical Engineering and Information Technology from RWTH Aachen University, Aachen, Germany in 2007 and 2013, respectively. Her research interests are in the area of statistical signal processing, communication theory and optimization theory. During her PhD studies, she focused on iterative receiver designs. Since September 2014 she works as a post-doc researcher at Vodafone Chair Mobile Communications Systems at Technical University Dresden. She has worked on various national and international research projects with publication in journals, conference proceedings and workshops. Her current research focus is on the transceiver design of multicarrier waveforms with the objective of providing a new air interface for 5G networks.

Luciano Leonel Mendes received the BSc. and MSc. in electrical engineering from Inatel, Brazil in 2001 and 2003, respectively. In 2007 he received the Doctor in electrical engineering degree from Unicamp, Brazil. Since 2001 he is a professor at Inatel, Brazil, where he has acted as technical manager of the hardware development laboratory from 2006 to 2012. He has coordinated the Master Program at Inatel and several research projects funded by FAPEMIG, FINEP and BNDES. He was a post-doc visiting researcher, sponsored by CNPq-Brasil, at Vodafone Chair Mobile Communications Systems at the Technical University Dresden from 2013 until 2015. His main area of research is wireless communication and currently he is working on multicarrier modulations for 5G networks and future mobile communications systems.

Maximilian Matthé received the Dipl.-Ing degree in electrical engineering from Technical University Dresden (TU Dresden), Dresden, Germany, in 2013. He is currently pursuing the Ph.D. in the Vodafone Chair Mobile Communication Systems at TU Dresden. During his studies, he focused on mobile communications systems and communication theory. He performed his internship at National Instruments. In his Diploma Thesis he concentrated on waveform design for flexible multicarrier transmission systems. His research focuses on the design and evaluation of MIMO architectures for future cellular networks.

Ivan Simões Gaspar studied electronic technician (1998) at ETE ‘FMC’ and received his B.S.E.E. and M.Sc. degrees with emphasis in telecommunications from Inatel in 2003 and 2008. From 2003 to 2011 held a coaching position as a technical supervisor and product manager in the R&D department of Hitachi Kokusai Linear S/A, responsible for the design of digital TV transmitters, terrestrial and satellite microwave links and multiplexer equipment operating in single frequency networks. From 2008 to 2011 he was also a lecturer at Inatel. Since February 2012 he has been a research associate and doctorate candidate at the Vodafone Chair / Technical University Dresden, working on new concepts, algorithms and performance evaluation of innovative modulation schemes in the 5GNOW project and in the RF Lead User Program with National Instruments.

Nicola Michailow received his Dipl.-Ing. degree in electrical engineering with focus on wireless communications and information theory from Techni- cal University Dresden (TU Dresden), Germany in 2010. From 2008 to 2009 he worked in the R&D department of Asahi Kasei Corporation, Japan, developing signal processing algorithms for sensor data analysis. Since 2010, he has been research associate at the Vodafone Chair at TU Dresden, receiving the Dr.-Ing. degree in 2015. His professional interests revolve around non-orthogonal waveforms for future 5G cellular systems, new applications like the Internet of Things and the Tactile Internet and moreover the FPGA-based prototype implementation of flexible multicarrier modulation schemes. During his time at the Vodafone Chair, he contributed to the FP7 projects 5GNOW, CREW, QOSMOS, EX- ALTED and was part of the RF Lead User Program with National Instruments.

Gerhard R. Fettweis (F’09) received the Ph.D. degree from RWTH Aachen, Germany, in 1990. Thereafter he was at IBM Research and then at TCSI Inc., USA. Since 1994, he has been Vodafone Chair Professor at Technical University Dresden, Dresden, Germany, with his main research interest on wireless transmission and chip design. He is an Honorary Doctorate at Technical University Tampere. As a repeat entrepreneur he has co-founded 11 startups so far. He has setup funded projects in size of close to EUR 1/2 billion, notably he runs the German Science Foundation’s CRC HAECE and COE cFAED. He is actively involved in organizing IEEE conferences, and has been TPC Chair of ICC 2009 and TTM 2012, General Chair of VTC Spring 2013, and DATE 2014.