



**HAL**  
open science

# A New Hybrid Message Passing Algorithm for Joint User Activity Detection, Channel Estimation and Data Decoding in Grant-Free OFDM-IDMA

Fakher Sagheer, Frédéric Lehmann, Antoine Berthet

► **To cite this version:**

Fakher Sagheer, Frédéric Lehmann, Antoine Berthet. A New Hybrid Message Passing Algorithm for Joint User Activity Detection, Channel Estimation and Data Decoding in Grant-Free OFDM-IDMA. IEEE Transactions on Vehicular Technology, In press, 10.1109/TVT.2024.3377889 . hal-04632993

**HAL Id: hal-04632993**

**<https://hal.science/hal-04632993>**

Submitted on 3 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A New Hybrid Message Passing Algorithm for Joint User Activity Detection, Channel Estimation and Data Decoding in Grant-Free OFDM-IDMA

Fakher Sagheer\*, Frédéric Lehmann\*, and Antoine O. Berthet†

## Abstract

Grant-free access has the potential to reduce considerably the latency with respect to conventional contention-based access. However, in highly dynamic wireless networks, this paradigm comes with a number of new challenges, such as performing user activity detection and channel estimation directly upon packet arrival on top of symbol detection and decoding, due to limited signal overhead. In this paper, we consider an efficient hybrid message-passing algorithm, combining belief propagation with expectation propagation, to jointly perform all aforementioned tasks. A new expectation propagation message approximation rule is introduced for the sake of low-complexity Gaussian message-passing, without compromising the performance. Motivated by the need for non-orthogonal multiple access over frequency-selective channels, we validate the proposed scheme over a multi-antenna OFDM-interleave-division multiple access scenario. Numerical performance evaluations demonstrate the efficiency of the proposed method compared to existing ones in the same context.

## Index Terms

\*F. Sagheer and F. Lehmann are with SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris, 91120 Palaiseau, France. F. Sagheer was supported by the LabEx DigiCosme under Project ANR-11-LABEX-0045-DIGICOSME.

†A. O. Berthet is with Laboratoire des signaux et systèmes (L2S), Université Paris-Saclay, CNRS, CentraleSupélec, 91190 Gif-sur-Yvette, France.

Corresponding author: F. Sagheer, e-mail: fakher.sagheer@telecom-sudparis.eu.

Manuscript received received May 26, 2023. Revised January 30, 2024.

Grant-free NOMA, OFDM, IDMA, user activity detection, factor graphs, belief propagation, expectation propagation, Gaussian approximation.

## I. INTRODUCTION

Ultrareliable low-latency communication (URLLC) is a new service introduced in 5G radio for delay-sensitive applications such as the industrial Internet of Things (IoT), autonomous driving, remote surgery, smart grids and online entertainment industry, to name a few [1]. The intermittent traffic generated by the corresponding devices calls for a new paradigm called grant-free access, which eliminates most of the signaling overhead induced by traditional handshaking protocols [2] in order to reduce the end-to-end delay to a matter of millisecond.

Like any change of paradigm, grant-free access comes with its own challenges. First, in the absence of explicit resource allocation, the receiver does not know a priori which users are transmitting and must therefore perform user activity detection (UAD) in addition to channel estimation (CE), multi-user detection (MUD), symbol demodulation (DEM) and decoding (DEC). Both separate [3], [4] and joint [5]–[8] UAD and MUD have been investigated and exhibit satisfactory performances even in massive access scenarios. However, these artificial intelligence (AI) or compressed sensing (CS) techniques rely either on large training datasets or on the unrealistic assumption of perfect channel state information (CSI) from a potentially very large number of users (often including those who are muted). This leads to a second challenge, which consists in CE for grant-free access. Preamble-based joint CE, UAD and MUD has been proposed in [11]–[15], but large preambles may defeat the very purpose of grant-free access, which is to limit signaling overhead. Methods using limited or no preamble exist [16], [17], but only in the specific context of massive MIMO. Motivated by the ability of non-orthogonal multiple access (NOMA) [18] to superimpose several signals over the same resources with controlled interference, many grant-free access schemes employ NOMA in the physical layer, rather than competing random access (RA)-based techniques [19], [20].

In this paper, we will focus on solving the multiple access problem at the receiver side inherent to grant-free NOMA using message-passing algorithms due to their flexibility for performing inference in probabilistic graphical models (see [21] for a unifying view of belief propagation (BP), expectation-propagation (EP), mean-field (MF) and variational Bayesian (VB) inference). Previous work considered separate CS-based UAD and discrete BP for perfect-CSI MUD [22],

and joint UAD and CE using either hybrid BP/MF [23] or hybrid BP/approximate message passing (AMP) [24] processing before separate MUD. The few attempts made for joint CE, MUD, UAD and DEC can be categorized depending on their underlying user activity and symbol models. First, BP-EP-VB in [29] relies on user activity being modeled by continuous-valued precision parameters of the channel coefficients, while modulated symbols are considered as discrete-valued. Secondly, in hybrid BP-EP [26] user activity is modeled by a binary random variable, while modulated symbols are considered as continuous-valued. Thirdly, BP [25] and auxiliary variable hybrid BP-EP-MF [27], [28] consider both the user activity variables and the modulated symbols as discrete-valued. A consequence of the mixed discrete-continuous probabilistic model inherent to the presence of discrete-valued user activity variables and/or modulated symbols, is that some messages are treated as probability mass functions (p.m.f.s) and others as probability density functions (p.d.f.s). Consequently, marginalization w.r.t. the user activity variables involves cumbersome discrete summations in [25], [28], while in [26] an extra expectation maximization (EM) procedure is needed to reconcile individual message-passing stages for UAD, CE and MUD. Here, we propose to solve this issue by introducing a unified treatment of all messages as Gaussian p.d.f.s, including for those incoming and outgoing from the user activity variables. In this way, any marginalization operation for message computation boils down to integration (instead of a computationally intensive mixed summation-integral operation) without extra effort and without leaving the message-passing framework. Moreover, we validate the proposed grant-free NOMA receiver over OFDM-IDMA, which is of interest in its own right. Indeed, interleave division multiple access (IDMA) [31] is easily combined with off-the-shelf building blocks such as standard modulation formats, OFDM, multi-antenna reception and channel coding [32], [33]. Besides being the simplest NOMA design ensuring interleaving-based user separation (i.e. without the need for involved codebook and/or spreading sequence construction) [2], to the best of our knowledge only [34] considers grant-free OFDM-IDMA, where preamble-based UAD is followed by conventional MUD under perfect CSI. On the other hand, message-passing for joint CE, MUD and DEC for MIMO-OFDM-IDMA is evaluated in [35] using BP with Gaussian approximation (GaBP) [36], [37], but under a setting incompatible with grant-free access that assumes perfect user activity knowledge. Hence there is a need for a thorough investigation of message-passing for joint UAD, CE, MUD and DEC for OFDM-IDMA dispensing with bandwidth-inefficient preambles. We now summarize our main contributions:

- The underlying rationale behind the design of the new hybrid GaBP/EP message-passing algorithm is novel in the sense that all messages incoming and outgoing from hidden variables (whether discrete or continuous-valued) are now treated as Gaussians (except in the DEC, where this would be irrelevant). The sole remaining issue of partitioning hidden variables between those inferred via BP (i.e. needing an extra Gaussian approximation) and those inferred via EP (i.e. with inbuilt projections over Gaussians), is motivated by the achievable performances at high signal-to-noise ratio.
- In our context, a generic theoretical problem arises from the fact that a noise-free signal is mostly a product of hidden variables (e.g. a data symbol times a channel gain or a UAD variable). Thus, marginalizing out a given hidden variable according to the corresponding Gaussian message, results in a mean and covariance depending on the remaining variables that are not explained away for the message outgoing from an observation node. We offer a principled solution to projecting such messages onto the set of Gaussian p.d.f.'s in Appendix A.
- From a practical point of view, explaining away all discrete hidden variables via Gaussian p.d.f.'s gives an opportunity for substantial complexity reduction by dispensing with prohibitively complex multiple summations.
- From an application point of view, virtually all message-passing algorithms work under the simplifying assumption of spatially uncorrelated Rayleigh fading models, which may no longer be relevant in (sub-)terahertz bandwidths considered for future 6G standards.
- Scattered pilot-based CE, UAD and DEC are performed on a per OFDM-block basis, thus enabling reliable short data burst transmission for sporadic IoT-like traffic over frequency-selective channels having coherence time as short as one OFDM block.

This paper is organized as follows: First, Sec. II describes our system model for grant-free transmissions, with OFDM-IDMA as the physical layer. In particular, the frequency selective channel and the user activity variable being independently distributed on a per-OFDM block basis, the model is suitable for bursty traffic. Using a factor graph representation of the system model, the proposed hybrid GaBP/EP algorithm for joint UAD, CE, MUD and DEC is introduced in Sec. III along with relevant implementation characteristics in Sec. IV. Finally, numerical assessment of the proposed method is provided in Sec. V.

**Notation.** Bold letters indicate vectors and matrices while  $\mathbf{0}_{m \times n}$  (resp.  $\mathbf{I}_m$ ) is the  $m \times n$  all-zero (resp. the  $m \times m$  identity) matrix and  $\text{diag}(\mathbf{a})$  is the diagonal matrix, whose diagonal entries are stored in vector  $\mathbf{a}$  and whose off-diagonal entries are zero.  $\mathcal{CN}(\mathbf{x}; \mathbf{m}, \Sigma)$  denotes a complex Gaussian distribution of the variable  $\mathbf{x}$ , with mean  $\mathbf{m}$  and covariance matrix  $\Sigma$ .  $\mathcal{B}(p)$  denotes a Bernoulli distribution with success probability  $p$ .

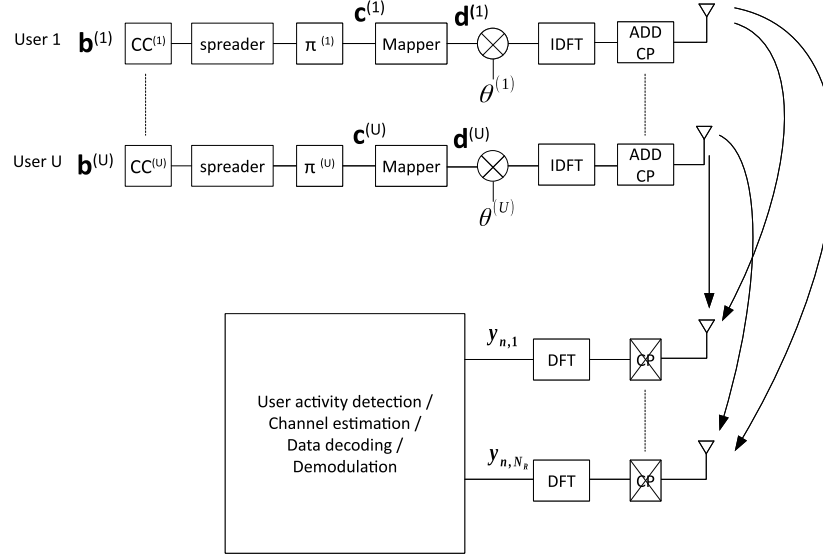


Fig. 1. SIMO-OFDM-IDMA system model.

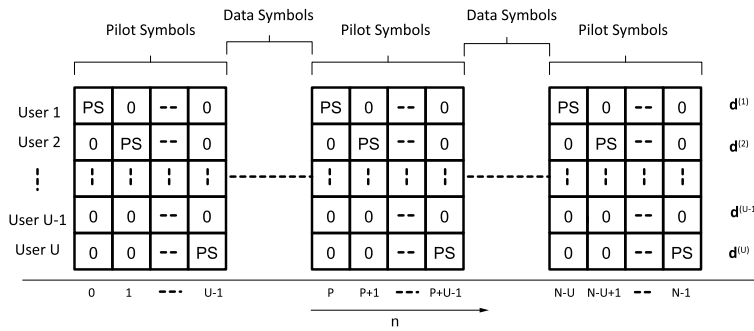


Fig. 2. Insertion of pilot symbols between data symbol with repetition period of  $P$ . PS (resp. 0) denotes non-zero (resp. zero) pilot symbol.

## II. SYSTEM MODEL

We focus on the grant-free NOMA uplink scenario described in Fig. 1. At the transmitter side (see Sec. II-A), IDMA-based NOMA allows uncoordinated multiple-access, while OFDM is used to combat multipath Rayleigh fading (see Sec. II-B). For the receiver side (see Sec. II-C), we describe an equivalent discrete-time complex baseband model along with the corresponding factor graph (see Sec. II-D) for the sake of joint UAD, CE, MUD, DEM and DEC.

### A. Transmitter

We consider a wireless access network with at most  $U$  single-antenna users. Grant-free IDMA is taken into account by modeling the activity of the  $u$ -th user with an i.i.d. (independently, and identically distributed) binary valued random variable  $\theta^{(u)} \sim \mathcal{B}(p_a^{(u)})$ . When the  $u$ -th user is active, it generates a sequence of uniformly distributed i.i.d. (u.i.i.d.) information bits  $\mathbf{b}^{(u)} \in \{0, 1\}^{N_b}$ , that is encoded to the binary sequence  $\mathbf{c}^{(u)} \in \{0, 1\}^{N_c}$  by e.g., a convolutional code  $CC^{(u)}$ , followed by a spreader (repetition code) and a user-specific interleaver  $\pi^{(u)}$ . This can be written as  $\mathbf{c}^{(u)} = \mathcal{C}^{(u)}(\mathbf{b}^{(u)})$ , where the one-to-one function  $\mathcal{C}^{(u)}(\cdot)$  denotes the combined effect of encoding and interleaving. By modulating the symbol labels  $\mathbf{c}_n^{(u)} = [c_{n,1}^{(u)}, c_{n,2}^{(u)}, \dots, c_{n,Q}^{(u)}]^T$  taken from  $\mathbf{c}^{(u)}$ , the  $n$ -th complex modulated symbol is obtained as  $d_n^{(u)} = \chi(\mathbf{c}_n^{(u)})$ , where  $\chi(\cdot)$  denotes a one-to-one  $Q$ -ary modulation operator enforcing  $E[|d_n^{(u)}|^2] = 1$ . Moreover, we define  $\chi_q^{-1}(d_n^{(u)})$  as the  $q$ -th label of  $d_n^{(u)}$  (i.e.  $c_{n,q}^{(u)}$  according to our notations). Inserting a known pilot sequence with a repetition period of  $P$  over the subcarriers indexed by  $\mathcal{P}^{(u)}$  as shown in Fig. 2, results in a length- $N$  vector of modulated symbol,  $\mathbf{d}^{(u)} = [d_0^{(u)}, \dots, d_{N-1}^{(u)}]^T$ . In order to combat multipath fading, OFDM transmission is implemented by passing  $\theta^{(u)}\mathbf{d}^{(u)}$  to an  $N$ -point Inverse Discrete Fourier Transform (IDFT) [32]. We also define  $\mathbf{B} = [\mathbf{b}^{(1)}, \mathbf{b}^{(2)}, \dots, \mathbf{b}^{(U)}]$ ,  $\mathbf{C} = [\mathbf{c}^{(1)}, \mathbf{c}^{(2)}, \dots, \mathbf{c}^{(U)}]$ , and  $\mathbf{D} = [\mathbf{d}^{(1)}, \mathbf{d}^{(2)}, \dots, \mathbf{d}^{(U)}]$  as the matrices of information bits, coded bits, modulated symbols. Subcarrier orthogonality is assumed to be maintained by inserting a long enough cyclic prefix (CP) to absorb the intersymbol interference caused by the channel's multipath delay spread and asynchronism.

### B. Channel model

A single-input multiple-output (SIMO) block fading multipath channel model affecting the  $u$ -th user is assumed, with hyperparameters  $E_s^{(u)}$  and  $\rho^{(u)}$  modeling the current user's average

receive energy and antenna correlation coefficient, respectively. The corresponding baseband channel impulse response (CIR) has the form

$$\mathbf{h}^{(u)}(\tau) = \sqrt{E_s^{(u)}} \sum_{l=0}^{L-1} \sqrt{p_l} \mathbf{\Gamma}^{(u)1/2} \mathbf{g}_l^{(u)} \delta(\tau - lT_s), \quad (1)$$

where  $p_l$  is the average power at the  $l$ -th delay,  $T_s$  is the sampling period, the  $\mathbf{g}_l^{(u)}$ 's are i.i.d. random vectors with distribution  $\mathcal{CN}(0, \mathbf{I}_{N_R})$ , and the antenna correlation matrix is modeled by  $\mathbf{\Gamma}^{(u)} = [\rho^{(u)|i-j|}]_{1 \leq i, j \leq N_R}$  [38]. For a standard exponentially decaying power delay profile,  $p_l = A e^{-lT_s/\bar{\sigma}_\tau}$ , for  $l = 0, \dots, L-1$ , where  $\bar{\sigma}_\tau$  is the rms delay spread and  $A$  is a normalization constant such that  $\sum_{l=0}^{L-1} p_l = 1$ . The corresponding channel frequency response (CFR) affecting the discrete frequency  $n/(NT_s)$  after  $N$ -point discrete Fourier transform (DFT) has the form

$$\mathbf{x}_n^{(u)} = \sqrt{E_s^{(u)}} \sum_{l=0}^{L-1} \sqrt{p_l} \mathbf{\Gamma}^{(u)1/2} \mathbf{g}_l^{(u)} e^{-j2\pi nl/N}, \quad (2)$$

so that  $E[(\mathbf{x}_n^{(u)} - \mathbf{x}_{n-1}^{(u)})(\mathbf{x}_n^{(u)} - \mathbf{x}_{n-1}^{(u)})^H] = \sigma^2 E_s^{(u)} \mathbf{\Gamma}^{(u)}$ , where  $\sigma = \sqrt{2 \sum_{l=0}^{L-1} p_l (1 - \cos(2\pi l/N))}$  is now the frequency correlation coefficient between consecutive discrete frequencies. For simplicity, the receiver postulates a surrogate frequency-domain Gauss-Markov approximate model for the CFR as [39]

$$\mathbf{x}_n^{(u)} = \mathbf{x}_{n-1}^{(u)} + \mathbf{\Delta}_n^{(u)}, \quad (3)$$

where  $\mathbf{\Delta}_n^{(u)} \sim \mathcal{CN}(\mathbf{\Delta}_n^{(u)}; \mathbf{0}_{N_R \times 1}, \zeta \sigma^2 E_s^{(u)} \mathbf{\Gamma}^{(u)})$  is an i.i.d. process noise with tuning parameter  $\zeta$  accounting for the approximation error.

### C. Received signal

After CP removal and  $N$ -point DFT, the baseband equivalent signal  $\mathbf{y}_n = [y_{n,1}, \dots, y_{n,N_R}]^T$  over the  $N_R$  receive antennas at subcarrier  $n$  has the form

$$\mathbf{y}_n = \sum_{u=1}^U \theta^{(u)} d_n^{(u)} \mathbf{x}_n^{(u)} + \mathbf{w}_n, \quad (4)$$

where  $\mathbf{w}_n \sim \mathcal{CN}(\mathbf{0}_{N_R \times 1}, \mathbf{R})$  is a i.i.d. Gaussian noise with covariance matrix  $\mathbf{R} = N_0 \mathbf{I}_{N_R}$ . In the sequel, we let  $\mathbf{x}^{(u)} = [\mathbf{x}_0^{(u)T}, \mathbf{x}_1^{(u)T}, \dots, \mathbf{x}_{N-1}^{(u)T}]^T$  denote the  $u$ -th user CFRs over all subcarrier indices and  $\mathbf{X} = [\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(U)}]$  denote the CFR matrix. Similarly, we let  $\boldsymbol{\theta} = [\theta^{(1)}, \theta^{(1)}, \dots, \theta^{(U)}]^T$  (resp.  $\mathbf{y} = [\mathbf{y}_0^T, \mathbf{y}_1^T, \dots, \mathbf{y}_{N-1}^T]^T$ ) denote the vector of activity variables for all users (resp. the vector of observations over all subcarriers) during the current OFDM block.



#### D. Factor graph representation of the system model

We use the elegant and convenient formalism of factor graphs to visualize conditional independencies among subsets of random variables [21]. Let us turn our attention to the factor graph corresponding to the joint *a posteriori* distribution of the hidden variables in our system model. Using Bayes' rule and the fact that hidden variables of different users are independent (see Sec. II-A-II-B), we can express the *a posteriori* distribution of all hidden variables as

$$p(\mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{X}, \boldsymbol{\theta} | \mathbf{y}) \propto p(\mathbf{y} | \mathbf{D}, \mathbf{X}, \boldsymbol{\theta}) \cdot \prod_{u=1}^U \{p(\mathbf{b}^{(u)})p(\mathbf{c}^{(u)} | \mathbf{b}^{(u)})p(\mathbf{d}^{(u)} | \mathbf{c}^{(u)})p(\mathbf{x}^{(u)})P(\theta^{(u)})\}, \quad (5)$$

where the last line comes from the reasonable assumption of independence between the user activity variables, the channel CFRs and the data bits for a given user. Now, further simplifications hold true in (5) under subcarrier orthogonality (see Sec. II-A)

$$p(\mathbf{y} | \mathbf{D}, \mathbf{X}, \boldsymbol{\theta}) = \prod_{n=0}^{N-1} p(\mathbf{y}_n | \mathbf{d}_n, \mathbf{X}_n, \boldsymbol{\theta}), \quad (6)$$

where  $\mathbf{d}_n = [d_n^{(1)}, d_n^{(2)}, \dots, d_n^{(U)}]^T$  and  $\mathbf{X}_n = [\mathbf{x}_n^{(1)}, \mathbf{x}_n^{(2)}, \dots, \mathbf{x}_n^{(U)}]$ . Also since information bits are u.i.i.d,  $p(\mathbf{b}^{(u)}) = (\frac{1}{2})^{N_b}$ , and since  $\mathcal{C}^{(u)}(\cdot)$  and  $\chi(\cdot)$  are deterministic functions

$$\begin{aligned} p(\mathbf{c}^{(u)} | \mathbf{b}^{(u)}) &\propto \delta(\mathbf{c}^{(u)} - \mathcal{C}^{(u)}(\mathbf{b}^{(u)})) \\ p(\mathbf{d}^{(u)} | \mathbf{c}^{(u)}) &\propto \prod_{n=0}^{N-1} \delta(d_n^{(u)} - \chi(\mathbf{c}_n^{(u)})), \end{aligned} \quad (7)$$

where  $\delta(\cdot)$  stands for the Kronecker delta function. Finally, since  $\mathbf{x}^{(u)}$  is modeled as a Markovian process (see Sec. II-B)

$$p(\mathbf{x}^{(u)}) = p(\mathbf{x}_0^{(u)}) \prod_{n=1}^{N-1} p(\mathbf{x}_n^{(u)} | \mathbf{x}_{n-1}^{(u)}). \quad (8)$$

The fraction of the factor graph corresponding to the tasks devoted to inferring the hidden variables of the  $u$ -th user (i.e. DEC, DEM, CE, UAD) is represented in Fig 3, where we have used the following shorthand notation for function nodes

$$\begin{aligned} \chi_n^{(u)} &= \delta(d_n^{(u)} - \chi(\mathbf{c}_n^{(u)})) \\ f_n^{(u)} &= p(\mathbf{x}_n^{(u)} | \mathbf{x}_{n-1}^{(u)}) \\ g_n &= p(\mathbf{y}_n | \mathbf{d}_n, \mathbf{X}_n, \boldsymbol{\theta}). \end{aligned} \quad (9)$$

Since all users share the same observations, it is understood that  $U$  such subgraphs are stacked in parallel planes.

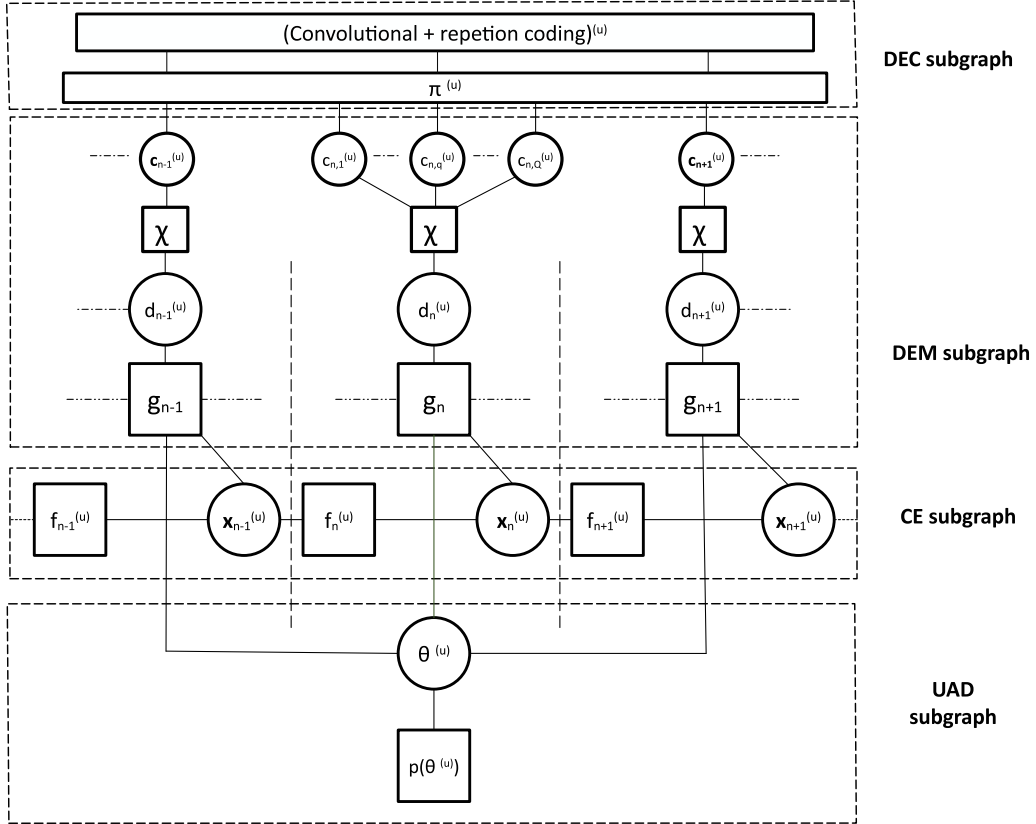


Fig. 3. Fraction of the factor graph corresponding to the  $u$ -th user hidden variables in grant-free coded OFDM-IDMA. The subgraph in each dashed box corresponds to a particular inference task (i.e. DEC, DEM, CE, UAD).

### III. PROPOSED HYBRID GABP/EP MESSAGE-PASSING ALGORITHM

The loopy factor graph representation of our problem is convenient to perform approximate inference of the hidden variables using a divide-and-conquer approach. We let  $\mu_{a \rightarrow b}(\cdot)$  denote the message sent by node  $a$  to node  $b$  in the factor graph. Unlike concurrent approaches [25], [29], we propose an hybrid GaBP/EP message-passing algorithm in the spirit of [42] for all messages on the UAD, CE and MUD subgraphs to obtain a low-complexity receiver. It follows that marginalizing out all variables reduces to simple Gaussian integrals, thus circumventing the computationally intensive summation operations defined over discrete alphabets. EP [41] being a natural framework to project messages into the set of Gaussian p.d.f.'s  $\mathcal{G}$ , a new principled way to do so is introduced in Appendix A. However, we select GaBP to compute the message incoming and outgoing from a user activity variables, since we found that its EP counterpart is

an uninformative message when the corresponding user is inactive (thus rendering inactive users impossible to classify as such). Accordingly, in the sequel we assume that messages incoming to the observation factor node  $g_n$  have the form

$$\begin{aligned}\mu_{\theta^{(u)} \rightarrow g_n}^{GaBP}(\theta^{(u)}) &\propto \mathcal{CN}(\theta^{(u)}; m_{\theta^{(u)} \rightarrow g_n}, \sigma_{\theta^{(u)} \rightarrow g_n}^2) \\ \mu_{d_n^{(u)} \rightarrow g_n}^{EP}(d_n^{(u)}) &\propto \mathcal{CN}(d_n^{(u)}; m_{d_n^{(u)} \rightarrow g_n}, \sigma_{d_n^{(u)} \rightarrow g_n}^2) \\ \mu_{\mathbf{x}_n^{(u)} \rightarrow g_n}^{EP}(\mathbf{x}_n^{(u)}) &\propto \mathcal{CN}(\mathbf{x}_n^{(u)}; \mathbf{m}_{\mathbf{x}_n^{(u)} \rightarrow g_n}, \Sigma_{\mathbf{x}_n^{(u)} \rightarrow g_n}),\end{aligned}\tag{10}$$

whose expression will be derived as (30), (11) and (25), respectively.

## A. EP messages

### 1) DEM subgraph:

a) *EP Message from  $d_n^{(u)}$  to  $g_n$* : After the DEC stage we are ready to compute the message back to  $g_n$ , that is readily available from [46, Eq. (30)-(33)]

$$\mu_{d_n^{(u)} \rightarrow g_n}^{EP}(d_n^{(u)}) \propto \frac{\text{proj}_{\mathcal{G}} \left( \frac{1}{Z} \mu_{g_n \rightarrow d_n^{(u)}}^{EP}(d_n^{(u)}) \prod_{q=1}^Q \mu_{c_{n,q}^{(u)} \rightarrow \chi_n^{(u)}}^{BP}(\chi_q^{-1}(d_n^{(u)})) \right)}{\mu_{g_n \rightarrow d_n^{(u)}}^{EP}(d_n^{(u)})},\tag{11}$$

where  $\mu_{c_{n,q}^{(u)} \rightarrow \chi_n^{(u)}}^{BP}(\cdot)$  is the DEC output in p.m.f. form computed in Sec. III-B1. Using the standard KL-based projection, the numerator of (11) is a Gaussian density, whose mean  $m_{d_n^{(u)}}$  and variance  $\sigma_{d_n^{(u)}}^2$  are straightforward to compute via moment-matching. Consequently,  $\mu_{d_n^{(u)} \rightarrow g_n}^{EP}(d_n^{(u)}) \propto \mathcal{CN}(d_n^{(u)}; m_{d_n^{(u)} \rightarrow g_n}, \sigma_{d_n^{(u)} \rightarrow g_n}^2)$ , where

$$\begin{aligned}\sigma_{d_n^{(u)} \rightarrow g_n}^2 &^{-1} = \sigma_{d_n^{(u)}}^2 &^{-1} - \sigma_{g_n \rightarrow d_n^{(u)}}^2 &^{-1} \\ \frac{m_{d_n^{(u)} \rightarrow g_n}}{\sigma_{d_n^{(u)} \rightarrow g_n}^2} &= \frac{m_{d_n^{(u)}}}{\sigma_{d_n^{(u)}}^2} - \frac{m_{g_n \rightarrow d_n^{(u)}}}{\sigma_{g_n \rightarrow d_n^{(u)}}^2}.\end{aligned}\tag{12}$$

b) *Message from  $g_n$  to  $d_n^{(u)}$* : Using the EP rule at  $d_n^{(u)}$ ,  $\mu_{d_n^{(u)} \rightarrow \chi_n^{(u)}}^{EP}(\cdot)$  has the form

$$\mu_{g_n \rightarrow d_n^{(u)}}^{EP}(d_n^{(u)}) \propto \mathcal{CN}(d_n^{(u)}; m_{g_n \rightarrow d_n^{(u)}}, \sigma_{g_n \rightarrow d_n^{(u)}}^2),\tag{13}$$

whose parameters are obtained by applying the EP rule at  $g_n$  as

$$\mu_{g_n \rightarrow d_n^{(u)}}^{EP}(d_n^{(u)}) \propto \frac{\text{proj}_{\mathcal{G}} \left( \frac{1}{Z} \mu_{d_n^{(u)} \rightarrow g_n}^{EP}(d_n^{(u)}) \tilde{f}(\mathbf{y}_n | d_n^{(u)}) \right)}{\mu_{d_n^{(u)} \rightarrow g_n}^{EP}(d_n^{(u)})},\tag{14}$$

where  $\tilde{f}(\mathbf{y}_n | d_n^{(u)})$  can be viewed as the likelihood function of  $d_n^{(u)}$  averaged over all other hidden variables at fixed  $\mathbf{y}_n$  (its expression, obtained via Thm. A.1, is postponed to Appendix B).

Now, combining the second line of (10) with (53), using the Kalman correction formula [43, p. 40] the numerator of (14) becomes

$$\text{proj}_{\mathcal{G}} \left( \frac{1}{Z} \mu_{d_n^{(u)} \rightarrow g_n}^{EP} (d_n^{(u)}) \tilde{f}(\mathbf{y}_n | d_n^{(u)}) \right) \approx \mathcal{CN}(d_n^{(u)}; \tilde{m}_{d_n^{(u)}}, \tilde{\sigma}_{d_n^{(u)}}^2), \quad (15)$$

where

$$\begin{aligned} \tilde{\mathbf{K}}_{d_n^{(u)}} &= \sigma_{d_n^{(u)} \rightarrow g_n}^2 \mathbf{h}_{d_n^{(u)}}^H \left( \sigma_{d_n^{(u)} \rightarrow g_n}^2 \mathbf{h}_{d_n^{(u)}} \mathbf{h}_{d_n^{(u)}}^H + \Sigma_{d_n^{(u)}} \right)^{-1} \\ \tilde{m}_{d_n^{(u)}} &= m_{d_n^{(u)} \rightarrow g_n} + \tilde{\mathbf{K}}_{d_n^{(u)}} \left( \mathbf{y}_n - \mathbf{h}_{d_n^{(u)}} m_{d_n^{(u)} \rightarrow g_n} - \mathbf{I}_{d_n^{(u)}} \right) \\ \tilde{\sigma}_{d_n^{(u)}}^2 &= \left( 1 - \tilde{\mathbf{K}}_{d_n^{(u)}} \mathbf{h}_{d_n^{(u)}} \right) \sigma_{d_n^{(u)} \rightarrow g_n}^2. \end{aligned} \quad (16)$$

Consequently,

$$\begin{aligned} \mu_{g_n \rightarrow d_n^{(u)}}^{EP} (d_n^{(u)}) &\approx \mathcal{CN}(d_n^{(u)}; m_{g_n \rightarrow d_n^{(u)}}, \sigma_{g_n \rightarrow d_n^{(u)}}^2), \quad \text{where} \\ \sigma_{g_n \rightarrow d_n^{(u)}}^2 &= \tilde{\sigma}_{d_n^{(u)}}^2 - \sigma_{d_n^{(u)} \rightarrow g_n}^2 \\ \frac{m_{g_n \rightarrow d_n^{(u)}}}{\sigma_{g_n \rightarrow d_n^{(u)}}^2} &= \frac{\tilde{m}_{d_n^{(u)}}}{\tilde{\sigma}_{d_n^{(u)}}^2} - \frac{m_{d_n^{(u)} \rightarrow g_n}}{\sigma_{d_n^{(u)} \rightarrow g_n}^2}. \end{aligned} \quad (17)$$

c) *Message from  $\chi_n^{(u)}$  to  $c_{n,q}^{(u)}$* : Similarly, the messages from DEM towards DEC, is readily available from [46, Eq. (38)]

$$\begin{aligned} \mu_{\chi_n^{(u)} \rightarrow c_{n,q}^{(u)}}^{EP} (c_{n,q}^{(u)} = 0) &\propto \frac{\sum_{d_n^{(u)}: \chi_q^{-1}(d_n^{(u)})=0} \mu_{g_n \rightarrow d_n^{(u)}}^{EP} (d_n^{(u)}) \prod_{q=1}^Q \mu_{c_{n,q}^{(u)} \rightarrow \chi_n^{(u)}}^{BP} (\chi_q^{-1}(d_n^{(u)}))}{\mu_{c_{n,q}^{(u)} \rightarrow \chi_n^{(u)}}^{BP} (c_{n,q}^{(u)} = 0)}, \\ \mu_{\chi_n^{(u)} \rightarrow c_{n,q}^{(u)}}^{EP} (c_{n,q}^{(u)} = 1) &\propto \frac{\sum_{d_n^{(u)}: \chi_q^{-1}(d_n^{(u)})=1} \mu_{g_n \rightarrow d_n^{(u)}}^{EP} (d_n^{(u)}) \prod_{q=1}^Q \mu_{c_{n,q}^{(u)} \rightarrow \chi_n^{(u)}}^{EP} (\chi_q^{-1}(d_n^{(u)}))}{\mu_{c_{n,q}^{(u)} \rightarrow \chi_n^{(u)}}^{EP} (c_{n,q}^{(u)} = 1)}. \end{aligned} \quad (18)$$

2) *CE subgraph*:

a) *Message from  $g_n$  to  $\mathbf{x}_n^{(u)}$* : Applying the EP rule at  $g_n$  results in

$$\mu_{g_n \rightarrow \mathbf{x}_n^{(u)}}^{EP} (\mathbf{x}_n^{(u)}) \propto \frac{\text{proj}_{\mathcal{G}} \left( \frac{1}{Z} \mu_{\mathbf{x}_n^{(u)} \rightarrow g_n}^{EP} (\mathbf{x}_n^{(u)}) \tilde{f}(\mathbf{y}_n | \mathbf{x}_n^{(u)}) \right)}{\mu_{\mathbf{x}_n^{(u)} \rightarrow g_n}^{EP} (\mathbf{x}_n^{(u)})}, \quad (19)$$

where  $\tilde{f}(\mathbf{y}_n|\mathbf{x}_n^{(u)})$  can be viewed as the likelihood function of  $\mathbf{x}_n^{(u)}$  averaged over all other hidden variables at fixed  $\mathbf{y}_n$ . Based on Thm. A.1 we obtain a Gaussian approximation of  $\tilde{f}(\mathbf{y}_n|\mathbf{x}_n^{(u)})$  (applying exactly the methodology employed to derive  $\tilde{f}(\mathbf{y}_n|d_n^{(u)})$  in (50)-(55)) as

$$\tilde{f}(\mathbf{y}_n|\mathbf{x}_n^{(u)}) \approx \mathcal{CN}(\mathbf{x}_n^{(u)}; \mathbf{m}_{\mathbf{x}_n^{(u)}}(\mathbf{x}_n^{(u)}), \Sigma_{\mathbf{x}_n^{(u)}}), \quad (20)$$

whose mean and covariance are given in Appendix B.

Now, combining the third line of (10) with (20), using the Kalman correction formula [43, p. 40] the numerator of (19) becomes

$$\text{proj}_{\mathcal{G}} \left( \frac{1}{Z} \mu_{\mathbf{x}_n^{(u)} \rightarrow g_n}^{EP}(\mathbf{x}_n^{(u)}) \tilde{f}(\mathbf{y}_n|\mathbf{x}_n^{(u)}) \right) \approx \mathcal{CN}(\mathbf{x}_n^{(u)}; \tilde{\mathbf{m}}_{\mathbf{x}_n^{(u)}}, \tilde{\Sigma}_{\mathbf{x}_n^{(u)}}), \quad (21)$$

where

$$\begin{aligned} \tilde{\mathbf{K}}_{\mathbf{x}_n^{(u)}} &= \Sigma_{\mathbf{x}_n^{(u)} \rightarrow g_n} \mathbf{h}_{\mathbf{x}_n^{(u)}}^H \left( \mathbf{h}_{\mathbf{x}_n^{(u)}} \Sigma_{\mathbf{x}_n^{(u)} \rightarrow g_n} \mathbf{h}_{\mathbf{x}_n^{(u)}}^H + \Sigma_{\mathbf{x}_n^{(u)}} \right)^{-1} \\ \tilde{\mathbf{m}}_{\mathbf{x}_n^{(u)}} &= \mathbf{m}_{\mathbf{x}_n^{(u)} \rightarrow g_n} + \tilde{\mathbf{K}}_{\mathbf{x}_n^{(u)}} \left( \mathbf{y}_n - \mathbf{h}_{\mathbf{x}_n^{(u)}} \mathbf{m}_{\mathbf{x}_n^{(u)} \rightarrow g_n} - \mathbf{I}_{d_n^{(u)}} \right) \\ \tilde{\Sigma}_{\mathbf{x}_n^{(u)}} &= \left( \mathbf{I}_{N_R} - \tilde{\mathbf{K}}_{\mathbf{x}_n^{(u)}} \mathbf{h}_{\mathbf{x}_n^{(u)}} \right) \Sigma_{\mathbf{x}_n^{(u)} \rightarrow g_n}. \end{aligned} \quad (22)$$

Consequently,

$$\begin{aligned} \mu_{g_n \rightarrow \mathbf{x}_n^{(u)}}^{EP}(\mathbf{x}_n^{(u)}) &\approx \mathcal{CN}(\mathbf{x}_n^{(u)}; \mathbf{m}_{g_n \rightarrow \mathbf{x}_n^{(u)}}, \Sigma_{g_n \rightarrow \mathbf{x}_n^{(u)}}), \quad \text{where} \\ \Sigma_{g_n \rightarrow \mathbf{x}_n^{(u)}}^{-1} &= \tilde{\Sigma}_{\mathbf{x}_n^{(u)}}^{-1} - \Sigma_{\mathbf{x}_n^{(u)} \rightarrow g_n}^{-1} \\ \Sigma_{g_n \rightarrow \mathbf{x}_n^{(u)}}^{-1} \mathbf{m}_{g_n \rightarrow \mathbf{x}_n^{(u)}} &= \tilde{\Sigma}_{\mathbf{x}_n^{(u)}}^{-1} \tilde{\mathbf{m}}_{\mathbf{x}_n^{(u)}} - \Sigma_{\mathbf{x}_n^{(u)} \rightarrow g_n}^{-1} \mathbf{m}_{\mathbf{x}_n^{(u)} \rightarrow g_n}. \end{aligned} \quad (23)$$

*b) EP Message-passing inside the CE subgraph:* Forward-backward EP message-passing inside the CE subgraph boils down to ordinary Gaussian message-passing of the form

$$\begin{aligned} \mu_{f_n^{(u)} \rightarrow \mathbf{x}_n^{(u)}}^{EP}(\mathbf{x}_n^{(u)}) &\propto \mathcal{N}_C(\mathbf{x}_n^{(u)}; \mathbf{m}_{f_n^{(u)} \rightarrow \mathbf{x}_n^{(u)}}, \Sigma_{f_n^{(u)} \rightarrow \mathbf{x}_n^{(u)}}) \\ \mu_{f_{n+1}^{(u)} \rightarrow \mathbf{x}_n^{(u)}}^{EP}(\mathbf{x}_n^{(u)}) &\propto \mathcal{N}_C(\mathbf{x}_n^{(u)}; \mathbf{m}_{f_{n+1}^{(u)} \rightarrow \mathbf{x}_n^{(u)}}, \Sigma_{f_{n+1}^{(u)} \rightarrow \mathbf{x}_n^{(u)}}), \end{aligned} \quad (24)$$

whose mean and covariance are updated similarly to Kalman filtering and smoothing (we omit the details here, the interested reader is referred to [40, Fig. 15], [44]).

c) *Message from  $\mathbf{x}_n^{(u)}$  to  $g_n$* : We are now ready to compute the message sent back to  $g_n$ . Applying the EP rule at  $\mathbf{x}_n^{(u)}$ , we obtain

$$\mu_{\mathbf{x}_n^{(u)} \rightarrow g_n}^{EP}(\mathbf{x}_n^{(u)}) \propto \mathcal{N}_C(\mathbf{x}_n^{(u)} : \mathbf{m}_{\mathbf{x}_n^{(u)} \rightarrow g_n}, \Sigma_{\mathbf{x}_n^{(u)} \rightarrow g_n}), \quad (25)$$

whose mean and covariance are computed as

$$\begin{aligned} \Sigma_{\mathbf{x}_n^{(u)} \rightarrow g_n}^{-1} &= \Sigma_{f_n^{(u)} \rightarrow \mathbf{x}_n^{(u)}}^{-1} + \Sigma_{f_{n+1}^{(u)} \rightarrow \mathbf{x}_n^{(u)}}^{-1} \\ \Sigma_{\mathbf{x}_n^{(u)} \rightarrow g_n}^{-1} \mathbf{m}_{\mathbf{x}_n^{(u)} \rightarrow g_n} &= \Sigma_{f_n^{(u)} \rightarrow \mathbf{x}_n^{(u)}}^{-1} \mathbf{m}_{f_n^{(u)} \rightarrow \mathbf{x}_n^{(u)}} + \Sigma_{f_{n+1}^{(u)} \rightarrow \mathbf{x}_n^{(u)}}^{-1} \mathbf{m}_{f_{n+1}^{(u)} \rightarrow \mathbf{x}_n^{(u)}}. \end{aligned} \quad (26)$$

## B. BP messages

1) *DEC subgraph*: On the DEC subgraph, since all variables are binary it is natural in EP to project all messages onto the family of binomial p.m.f.s, so that EP boils down to BP [46]. Consequently, ordinary BP decoding is in order for the repetition code and the convolutional code  $CC^{(u)}$  (see [40] for the update rules).

### 2) UAD subgraph:

a) *Message from  $g_n$  to  $\theta^{(u)}$* : This time, as mentioned in the introductory part of this section, we use the BP rule at  $g_n$  resulting in the binary-valued message

$$\mu_{g_n \rightarrow \theta^{(u)}}^{BP}(\theta^{(u)}) \propto \tilde{f}(\mathbf{y}_n | \theta^{(u)}) \quad (27)$$

where  $\tilde{f}(\mathbf{y}_n | \theta^{(u)})$  is the continuous Gaussian mixture given by (58), that can be viewed as the likelihood function of  $\theta^{(u)}$  averaged over all other hidden variables at fixed  $\mathbf{y}_n$ .

Using standard moment-matching [43, p. 106-108], we obtain the Gaussian approximation

$$\tilde{f}(\mathbf{y}_n | \theta^{(u)}) \approx \mathcal{CN}(\mathbf{y}_n : \mathbf{m}_{\theta^{(u)}, n}(\theta^{(u)}), \Sigma_{\theta^{(u)}, n}(\theta^{(u)})), \quad (28)$$

whose mean given by (59) (resp. covariance given by (60)) accounts for the average  $u$ -th user useful signal and multi-access interference conditional on  $\theta^{(u)}$  (resp. accounts for the noise and residual uncertainty on all hidden variables, except  $\theta^{(u)}$ ).

b) *Message from  $\theta^{(u)}$  to  $g_n$* : Now, applying the BP rule at  $\theta^{(u)}$

$$\mu_{\theta^{(u)} \rightarrow g_n}^{BP}(\theta^{(u)}) \propto P(\theta^{(u)}) \prod_{\substack{n'=0: \\ n' \neq n}}^{N-1} \mu_{g_{n'} \rightarrow \theta^{(u)}}^{BP}(\theta^{(u)}). \quad (29)$$

Finally, since this message is desired in GaBP form, we obtain the corresponding Gaussian message using moment-matching as

$$\mu_{\theta^{(u)} \rightarrow g_n}^{GaBP}(\theta^{(u)}) \propto \mathcal{CN}(\theta^{(u)}; m_{\theta^{(u)} \rightarrow g_n}, \sigma_{\theta^{(u)} \rightarrow g_n}^2), \quad (30)$$

where

$$m_{\theta^{(u)} \rightarrow g_n} = \frac{p_a^{(u)} \prod_{\substack{n'=0 \\ n' \neq n}}^{N-1} \frac{\mu_{g_{n'} \rightarrow \theta^{(u)}}^{BP}(\theta^{(u)} = 1)}{\mu_{g_{n'} \rightarrow \theta^{(u)}}^{BP}(\theta^{(u)} = 0)}}{1 - p_a^{(u)} + p_a^{(u)} \prod_{\substack{n'=0 \\ n' \neq n}}^{N-1} \frac{\mu_{g_{n'} \rightarrow \theta^{(u)}}^{BP}(\theta^{(u)} = 1)}{\mu_{g_{n'} \rightarrow \theta^{(u)}}^{BP}(\theta^{(u)} = 0)}} \quad (31)$$

and

$$\sigma_{\theta^{(u)} \rightarrow g_n}^2 = m_{\theta^{(u)} \rightarrow g_n} (1 - m_{\theta^{(u)} \rightarrow g_n}). \quad (32)$$

Also, the belief of  $\theta^{(u)}$

$$P(\theta^{(u)} | \mathbf{y}_{0:N-1}) \propto P(\theta^{(u)}) \prod_{n=0}^{N-1} \mu_{g_n \rightarrow \theta^{(u)}}^{BP}(\theta^{(u)}) \quad (33)$$

can be used to perform  $u$ -th user maximum *a posteriori* activity hard detection as

$$\hat{\theta}^{(u)} = \begin{cases} 1 & \text{if } P(\theta^{(u)} = 1 | \mathbf{y}_{0:N-1}) > P(\theta^{(u)} = 0 | \mathbf{y}_{0:N-1}) \\ 0 & \text{otherwise.} \end{cases} \quad (34)$$

#### IV. OVERALL RECEIVER IMPLEMENTATION

##### A. Hyperparameter estimation

Note that computing the covariance matrices (57) and initializing the CE recursions (24) to the CFR priors, need the knowledge of  $\{E_s^{(u)}, \rho^{(u)}\}_{u=1}^U$ . While estimation of hyperparameters could be done in a unified way by including them into the factor graph, we adopt a *separate low-cost* estimation strategy with moderate suboptimality w.r.t. perfect hyperparameter knowledge (See. Sec. V). Similarly to [51], we build a simple pilot-only correlation-based estimator, that will be shown to meet the more complex maximum-likelihood method at signal-to-noise ratios of practical interest (See. Sec. V).

$$\hat{E}_s^{(u)} = \frac{1}{|\mathcal{P}^{(u)}| N_R} \sum_{n \in \mathcal{P}^{(u)}} \sum_{j=1}^{N_R} \frac{|y_{n,j}|^2 - N_0}{|d_n^{(u)}|^2} \quad (35)$$

$$\hat{\rho}^{(u)} = \frac{1}{\hat{E}_s^{(u)} |\mathcal{P}^{(u)}| (N_R - 1)} \sum_{n \in \mathcal{P}^{(u)}} \sum_{j=1}^{N_R-1} \frac{\text{Re}\{y_{n,j} y_{n,j+1}^*\}}{|d_n^{(u)}|^2}. \quad (36)$$

At the first iteration of the proposed receiver, if for the  $u$ -th user  $\hat{E}_s^{(u)} < N_0$  (i.e. below the noise floor) the  $u$ -th user is pre-classified as inactive and we let

$$\begin{cases} \hat{E}_s^{(u)} = \max_{u'=1, \dots, U} \hat{E}_s^{(u')} \\ \hat{\rho}^{(u)} = \max_{u'=1, \dots, U} \hat{\rho}^{(u')}. \end{cases} \quad (37)$$

so as to maximize the uncertainty measured by the aforementioned hyperparameter-dependent covariance matrices. During subsequent iterations, we apply the same procedure, except that the  $u$ -th user is pre-classified as inactive, when at the previous iteration  $\hat{\theta}^{(u)} = 0$  in (34).

### B. Initialization

Before starting the first iteration, initialization is needed for  $\mu_{c_{n,q}^{(u)} \rightarrow \chi_n^{(u)}}(c_{n,q}^{(u)})$  and all messages incoming to  $g_n$  in (10). We initialize those messages using the prior information available from Sec. II. Thus,  $\mu_{c_{n,q}^{(u)} \rightarrow \chi_n^{(u)}}(c_{n,q}^{(u)})$  is initialized to a uniform p.m.f. for all coded digits and all users. On a pilot subcarrier (resp. information subcarrier)  $\mu_{d_n^{(u)} \rightarrow g_n}^{EP}(d_n^{(u)})$  is initialized to Gaussian with mean equal to the pilot symbol and zero variance (resp. a Gaussian with zero mean and unit variance). Similarly,  $\mu_{\mathbf{x}_n^{(u)} \rightarrow g_n}^{EP}(\mathbf{x}_n^{(u)})$  is initialized to a complex Gaussian with zero mean and covariance matrix  $E_s^{(u)} \mathbf{\Gamma}^{(u)}$ . However, it was found that the proposed method works best with the initialization  $\mu_{\theta^{(u)} \rightarrow g_n}^{GaBP}(\theta^{(u)}) = \mathcal{NC}(\theta^{(u)}; 1, 0)$ , which corresponds to assuming the existence of all users before any processing.

### C. Message-passing schedule

Over a graph such as the one in Fig. 3, a wide variety of message-passing schedules are possible. We choose to process all  $U$  user subgraphs one at a time with the following serial schedule (see Fig. 4):

- 1) For the current user index  $u$ , we reset  $\mu_{\theta^{(u)} \rightarrow g_n}^{GaBP}(\theta^{(u)}) = \mathcal{NC}(\theta^{(u)}; 1, 0)$ , which improves the probability of missed detection ( $P_{md}$ ) and the probability of false alarm ( $P_{fa}$ ) considerably
- 2) CE subgraph processing executing the steps in Sec. III-A2 in that order
- 3) DEM subgraph processing executing the steps in Sec. III-A1 in that order
- 4) DEC subgraph processing (see Sec. III-B1)



5) UAD subgraph processing executing the steps in Sec. III-B2 in that order.

The reason for beginning with CE is that all subsequent stages will be heavily dependent on the quality of the channel estimates. Also, since DEC enhances the reliability of the symbol estimates compared to the DEM stage, it is preferable to perform UAD after DEC instead of after DEM. Finally, since the graph is loopy, this schedule needs to be iterated  $N_{it}$  times until convergence.

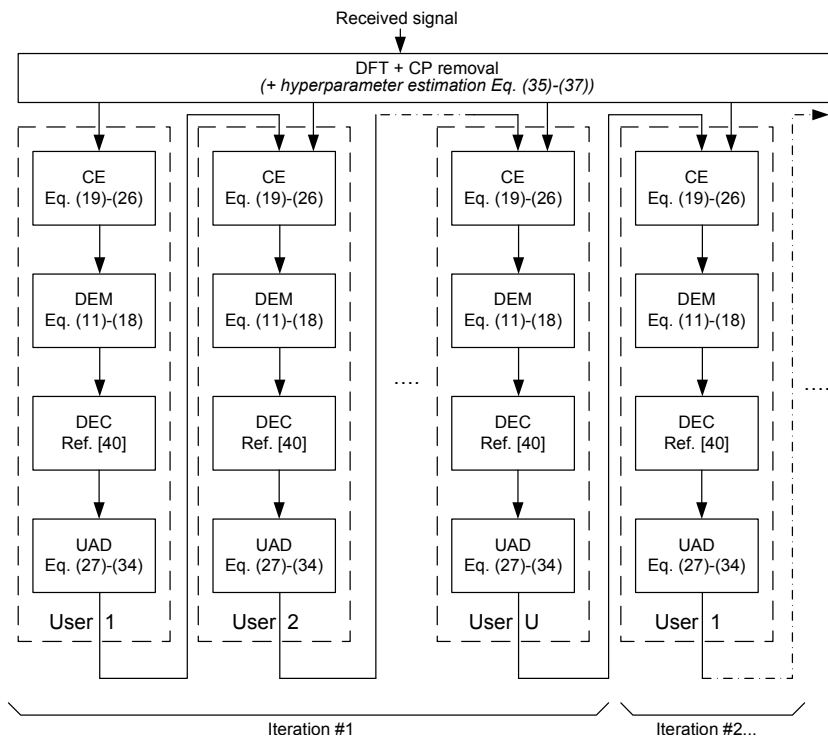


Fig. 4. Receiver flow chart.

#### D. Numerical stability

Overflow/underflow on bit or symbol probabilities is avoided in the log domain by using log-likelihood ratios (LLRs). Also, Gaussian message-passing with EP involves a division of two Gaussian densities, that is well defined only when the corresponding variance/covariance is positive-definite. In the rare instances where this is not the case, we adopt the solution advocated in [46], which is to replace the division of Gaussians by the numerator only.

Aside from the aforementioned usual problems, the proposed algorithm has also a few numerical issues on its own. First, when the iteration index grows large,  $\mu_{g_n \rightarrow d_n^{(u)}}^{EP}(d_n^{(u)})$  will have

a tendency to have a mean close to the correct symbol value with variance approaching zero. In such a situation, the mean  $m_{d_n^{(u)}}$  and variance  $\sigma_{d_n^{(u)}}^2$  of the argument of the projection operator in (11) are undefined due to numerical overflow. The solution here is to assign to  $m_{d_n^{(u)}}$  the mean of  $\mu_{g_n \rightarrow d_n^{(u)}}^{EP}(d_n^{(u)})$  and set  $\sigma_{d_n^{(u)}}^2$  to zero. Secondly, when  $\|m_{\theta^{(u)} \rightarrow g_n} \mathbf{m}_{\mathbf{x}_n^{(u)} \rightarrow g_n}\|_2^2 / N_R \leq 10^{-15}$ , which typically happens when the  $u$ -th user is inactive,  $\tilde{f}(\mathbf{y}_n | d_n^{(u)})$  can be assimilated to a constant so that (14) boils down to a constant. Since we need such an uninformative message to be represented as a Gaussian instead, we set it to a zero-mean complex Gaussian with variance  $10^{12}$ . In the same way, when  $|m_{\theta^{(u)} \rightarrow g_n} m_{d_n^{(u)} \rightarrow g_n}|^2 \leq 10^{-15}$ , which typically happens when the  $u$ -th user is inactive (and of course if subcarrier index  $n$  corresponds to a zero-valued pilot symbol for the  $u$ -th user),  $\tilde{f}(\mathbf{y}_n | \mathbf{x}_n^{(u)})$  can be assimilated to a constant so that (19) boils down to a constant. Since we need such an uninformative message to be represented as a Gaussian instead, we set it to a zero-mean complex Gaussian with variance  $10^{12} \mathbf{I}_{N_R}$ .

### E. Complexity analysis

The computational complexity per-user per-subcarrier per-iteration of

- each mean vector evaluation in (54), (56) and (59) is  $\mathcal{O}(N_R)$ ;
- each covariance matrix evaluation in (55), (57) and (60) is  $\mathcal{O}(N_R^2)$ ;
- CE subgraph processing in Sec. III-A2 is  $\mathcal{O}(N_R^3)$ ; due to matrix inversion in (26)
- DEM subgraph processing in Sec. III-A1 is  $\mathcal{O}(N_R^3)$  due to matrix inversion in (16);
- UAD subgraph processing in Sec. III-B2 is  $\mathcal{O}(2N_R^3)$  due to matrix inversion in evaluating (28) for each value of the existence variable in  $\{0, 1\}$ .

Consequently, we obtain the desirable property that the per-subcarrier per-iteration computational complexity of the proposed hybrid GaBP/EP message-passing receiver grows only linearly with  $U$ , the maximum number of users. Also, comparing to standard BP in [25], all steps except for CE and UAD subgraph processing, experience a complexity reduction by a factor of  $Q$ , which is of particular interest for higher order modulation formats. This is due to EP being employed to represent all messages incoming and outgoing from that data symbol variables as a Gaussian density instead of a discrete p.m.f., thus avoiding  $Q$ -fold discrete summations when marginalizing out data symbols.

### F. Benchmark algorithm

We propose a reduced-complexity benchmark algorithm inspired from [47]- [48], that is similar to the proposed method in all respects, except that hard UAD is performed, instead of *a posteriori* p.m.f.-based soft UAD. This method is based on the normalized correlation between the received signal cleansed from the estimated multi-access interference (MAI) and the re-estimated  $u$ -th user useful signal, as given by the following equation

$$R^{(u)} = \frac{|\sum_{n=0}^{N-1} (\hat{d}_n^{(u)} \mathbf{m}_{\mathbf{x}_n^{(u)} \rightarrow g_n})^H \mathbf{z}_n|}{\sqrt{\sum_{n=0}^{N-1} \mathbf{z}_n^H \mathbf{z}_n}}, \quad (38)$$

where  $\mathbf{z}_n$  is the received signal after cancelling the estimated MAI

$$\mathbf{z}_n = \mathbf{y}_n - \sum_{u' \neq u} \hat{\theta}^{(u')} \hat{d}_n^{(u')} \mathbf{m}_{\mathbf{x}_n^{(u')} \rightarrow g_n} \quad (39)$$

and  $\hat{d}_n^{(u)}$  is the hard decision on the  $u$ -th user's symbol over subcarrier  $n$  reconstructed from the DEC output. Here hard UAD is obtained by performing hypothesis testing

$$\hat{\theta}^{(u)} = \begin{cases} 1 & \text{if } R^{(u)} > \lambda_t \\ 0 & \text{otherwise,} \end{cases} \quad (40)$$

where  $\lambda_t$  is the threshold needed to obtain a constant  $P_{fa} \approx 10^{-5}$ .

## V. SIMULATION RESULTS

All simulations use the grant-free OFDM-IDMA setup described in Sec. V-A. In section V-B, the performances of the proposed hybrid GaBP/EP algorithm in Sec. III are compared with three algorithms: a more computationally intensive version of the algorithm in the form of BP [25], the reduced-complexity standard correlator-based benchmark algorithm in Sec. IV-F [47]- [48]), and finally, a competing version of hybrid EP/BP based on scalar auxiliary variables [27], [28]. Mostly overlooked issues in grant-free access regarding the robustness w.r.t. antenna correlation and unknown hyperparameters, are thoroughly investigated in Sec. V-C and Sec. V-D, respectively. In the absence of traffic statistics, a reasonable choice for  $p_a^{(u)}$  is  $1/2$  for  $u = 1, \dots, U$ .

Parameter	Value
Channel coding	$(5/7)_8$ recursive convolutional code
Spreader	rate-1/4 repetition code
Modulation	16-QAM
Total number of users ( $U$ )	16
Number of active users	12
OFDM subcarriers ( $N$ )	1024
Pilot spacing ( $P$ )	24
CP size	$N/8$ samples
Channel power delay profile	exponential decay constant=3 taps
Tx antennas	1
Rx antennas ( $N_R$ )	4
Rx modeling error parameter ( $\zeta$ )	15

TABLE I

SYSTEM MODEL PARAMETERS.

user index	$\theta^{(u)}$	$\rho^{(u)}$	$E_s^{(u)}/N_0$ (dB)
$u \in \{1, 2, 3, , 4\}$	0	0	$E_s/N_0$ (dB)
$u \in \{5, \dots, 16\}$	1	0	$E_s/N_0$ (dB)

TABLE II

EQUAL RECEIVE ENERGY SCENARIO.

### A. Setup

The system model parameters in Sec. II-A-II-B are summarized in Tab. I. In the equal receive energy scenario given by Tab. II, all active users have the same reference signal-to-noise ratio (SNR)  $E_s/N_0$  (dB). In the unequal receive energy scenario given by Tab III, half the active users suffer from a 6 dB penalty w.r.t. the reference SNR  $E_s/N_0$  (dB).

user index	$\theta^{(u)}$	$\rho^{(u)}$	$E_s^{(u)}/N_0$ (dB)
$u \in \{1, 2, 3, 4\}$	0	0.4	$E_s/N_0$ (dB)
$u \in \{7, 8, 10, 11, 13, 15\}$	1	0.4	$E_s/N_0$ (dB)
$u \in \{5, 6, 9, 12, 14, 16\}$	1	0.4	$E_s/N_0$ (dB) - 6 dB

TABLE III  
UNEQUAL RECEIVE ENERGY SCENARIO.

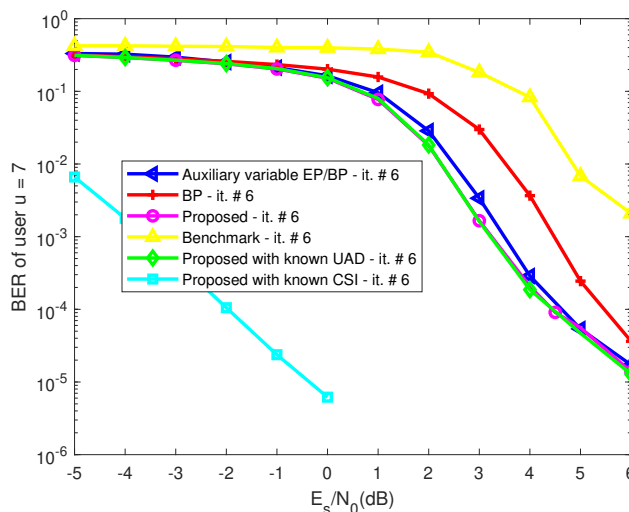


Fig. 5. BER under equal energy and known hyperparameters.

### B. Comparison with existing methods

We compare the proposed method with existing alternatives, under standard assumptions of equal energy reception (see Tab. II) and known hyperparameters.

Fig. 5 (resp. Fig. 6) shows the bit error rate (BER) (resp. the CFR estimation mean squared error (MSE)) for an active user indexed by  $u = 7$ . After convergence at iteration 6, the proposed method outperforms both the less computationally intensive benchmark algorithm but also the more computationally intensive BP [25]. This confirms that the simplifications made possible by hybrid GaBP/EP lead to an interesting performance vs. complexity tradeoff. It is remarkable that the proposed method suffers no performance loss w.r.t. the known UAD lower bound. Results (although not shown to lack of space) indicate that the single-user lower bound with known CSI

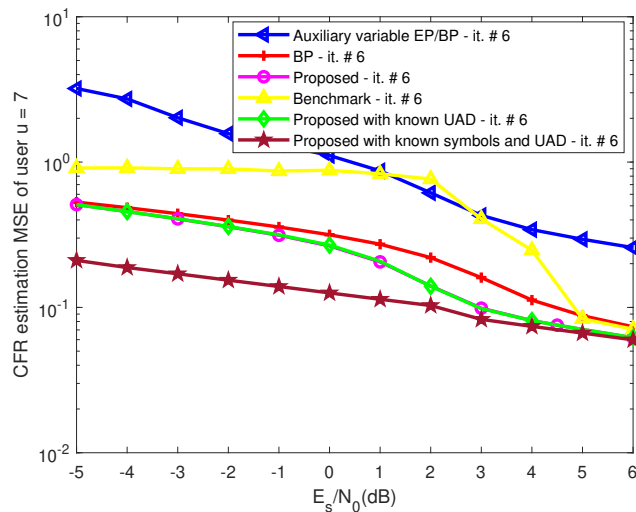


Fig. 6. CFR estimation MSE under equal energy and known hyperparameters.

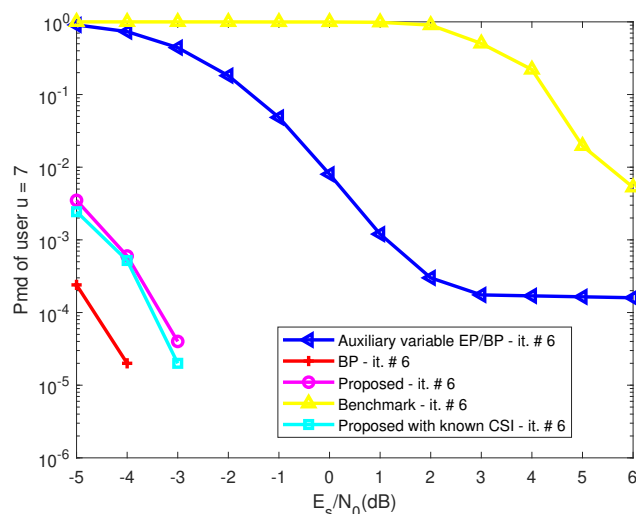


Fig. 7.  $P_{md}$  under equal energy and known hyperparameters.

has the same results as the proposed method with known CSI. Consequently, the 6 dB loss w.r.t. to known CSI can be interpreted as the price to be paid to obtain fixed-complexity message update rules in the form of Gaussian approximations under higher order 16-QAM modulation. Note that the large gap in channel parameter estimation MSE of the auxiliary variable hybrid BP/EP w.r.t. the proposed method, translates only into a negligible BER gap, as observed previously in [52]. The performance of UAD is evaluated in terms of probability of missed detection (resp. false alarm),  $P_{md}$  (resp.  $P_{fa}$ ) for an active (resp. inactive) user, as depicted in Fig. 7 (resp. Fig. 8).

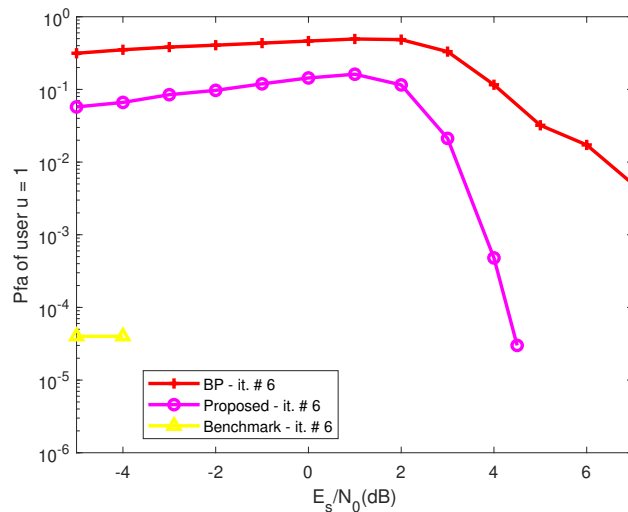


Fig. 8.  $P_{fa}$  under equal energy and known hyperparameters.

We observe that the benchmark algorithm reaches the prescribed  $P_{fa} \sim 10^{-5}$  at the expense of high  $P_{md}$  over the entire SNR range. On the other hand, while BP exhibits a moderate 1 dB power efficiency gain over the proposed method, its  $P_{fa}$  is uniformly worse over the entire SNR range. While the auxiliary variable hybrid BP/EP has negligible  $P_{fa}$  over the entire SNR range, it also exhibits an error floor on the  $P_{md}$ . This would be harmful to our grant-free setting, where asking non-cooperative users for retransmission of lost packets is undesirable. We conclude that the  $P_{md}$  vs.  $P_{fa}$  tradeoff of the proposed hybrid GaBP/EP approach is better than for the BP, benchmark and auxiliary variable hybrid BP/EP counterparts.

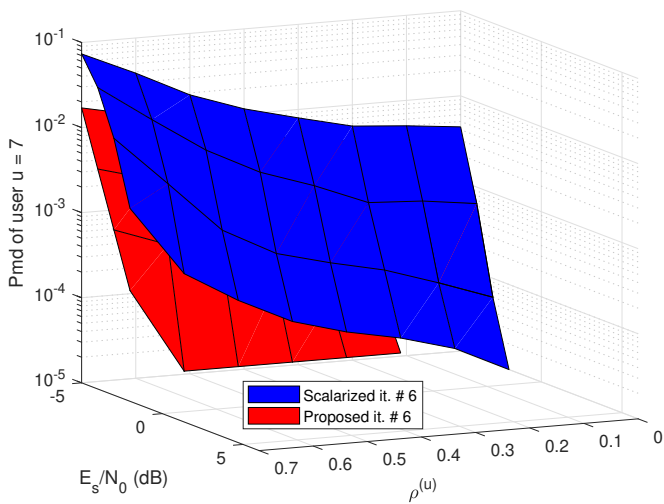


Fig. 11.  $P_{md}$  under equal energy and known hyperparameters.

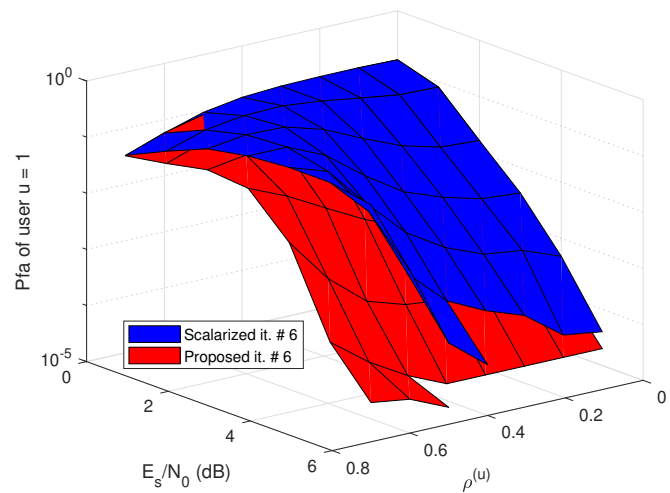


Fig. 12.  $P_{fa}$  under equal energy and known hyperparameters.

### C. Robustness w.r.t. antenna correlation

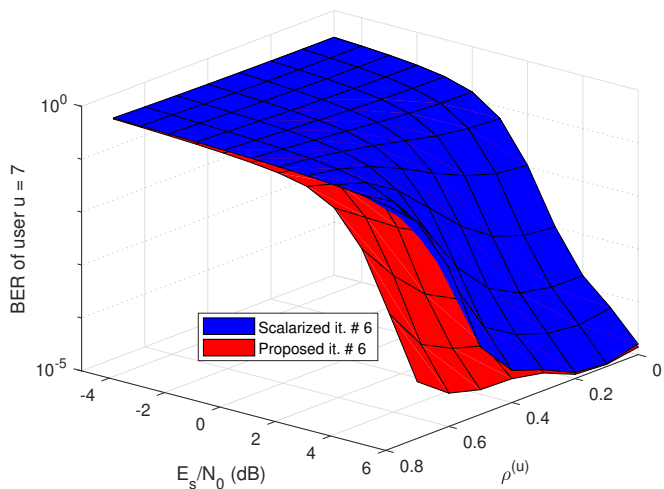


Fig. 9. BER under equal energy and known hyperparameters.

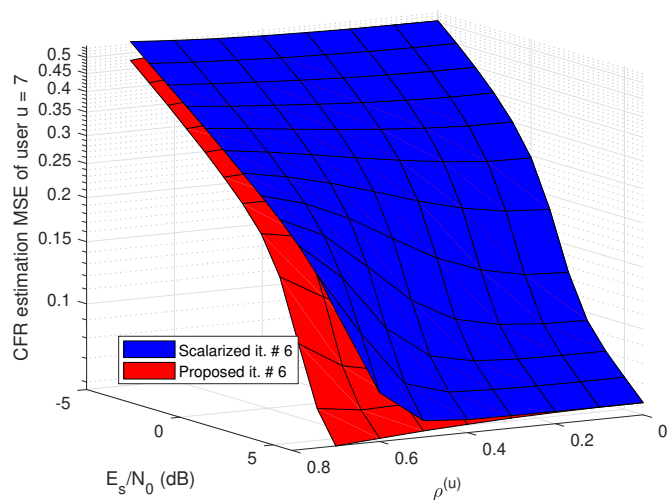


Fig. 10. CFR estimation MSE under equal energy and known hyperparameters.

In a preliminary conference paper [30], a simplified scalarized version of the proposed algorithm was derived under zero-antenna correlation. Again for simplicity we restrict ourselves to standard assumptions of equal energy reception (see Tab. II) and known hyperparameters. Fig 9-12 show all aforementioned performance metrics, by varying the antenna correlation coefficient



$\rho^{(u)}, u = 1, \dots, U$ . We compare the proposed method taking explicitly antenna correlation into account, with [30] working under the false hypothesis of zero-antenna correlation. The proposed algorithm is uniformly better over the entire range of SNR and  $\rho^{(u)}$ , which makes the proposed algorithm more general. However, since the scalarized version replaces the terms in  $N_R^3$  in Sec. IV-E by  $N_R$  in the complexity analysis [30], the complexity increase of the proposed method makes good engineering sense only for significant performance advantages w.r.t. the scalarized version, i.e. for applications such that  $\rho^{(u)} \geq 0.4$ . The auxiliary variable hybrid BP/EP exhibits a similar behavior, which is easily explained by the fact that it also relies on a channel gain independence assumption that is violated under antenna correlation.

#### *D. Robustness w.r.t. unknown hyperparameters*

Let us turn our attention to the case of unequal energy reception (see Tab. III) and study the ability of the proposed receiver to estimate unknown hyperparameters on the fly (since user activity can change from one OFDM block to the next) using the scheme in Sec. IV-A. Fig. 13 (resp. Fig. 14) illustrates the normalized estimation MSE for symbol energy estimation (resp. the estimation MSE for the antenna correlation coefficient). Note that the 6 dB loss in terms of transmit power for the low energy w.r.t. the reference energy user is reflected in the corresponding performance curves of the proposed hyperparameter estimation scheme. Also at high SNR, for energy (resp. antenna correlation) estimation, the suboptimality is negligible (resp. mild) when compared to joint maximum likelihood (ML) hyperparameter estimation which needs a complex optimization scheme taking into account boundary conditions on the antenna correlation coefficient that we implemented using BLEIC [50].

Fig. 15-17 (resp. Fig. 19-20) show all aforementioned performance metrics for a low energy (resp. a reference energy) user as well as  $P_{fa}$  for an inactive user in Fig. 18, for the proposed receiver with hyperparameter estimation. Comparing with perfect hyperparameter knowledge, the BER (resp.  $P_{md}$  and  $P_{fa}$ ) suffers from a moderate penalty of less than 1 dB (resp. less than 2 dB). Reception under unequal and unknown hyperparameters is typical of grant-free access, since the receiver and the users would be non-cooperative. Our findings clearly show that the proposed method is robust in such a challenging yet realistic setting, provided that a negligible complexity increase (due to the simple hyperparameter estimation scheme in Sec. IV.A.) and mild performance losses are tolerated.

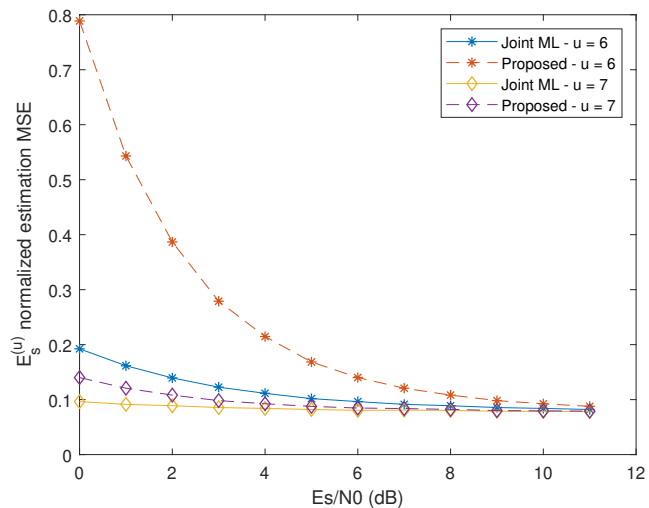


Fig. 13. Normalized estimation MSE of the symbol energy under unequal energy and unknown hyperparameters - low energy user  $u = 6$  and reference energy user  $u = 7$ .

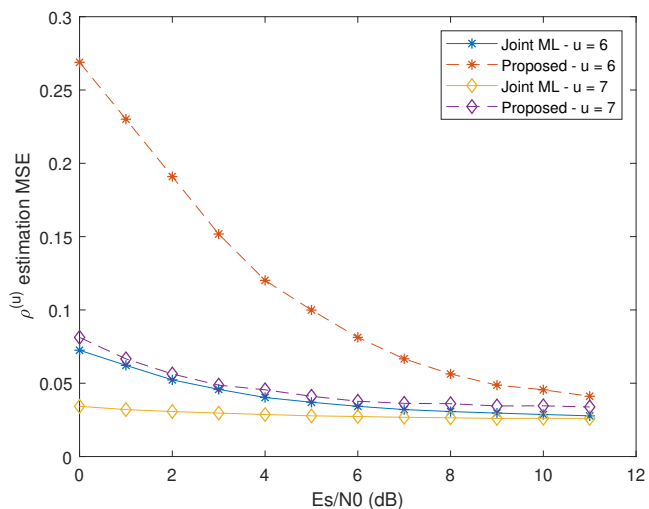


Fig. 14. Antenna correlation estimation MSE under unequal energy and unknown hyperparameters - low energy user  $u = 6$  and reference energy user  $u = 7$ .

## VI. CONCLUSION

In this paper, we have investigated a grant-free access system relying on multi-antenna OFDM-IDMA transmissions. An efficient hybrid GaBP/EP message-passing algorithm has been developed for joint UAD, CE, MUD and DEC, that compares favorably with existing methods in terms of performance vs complexity trade-off.

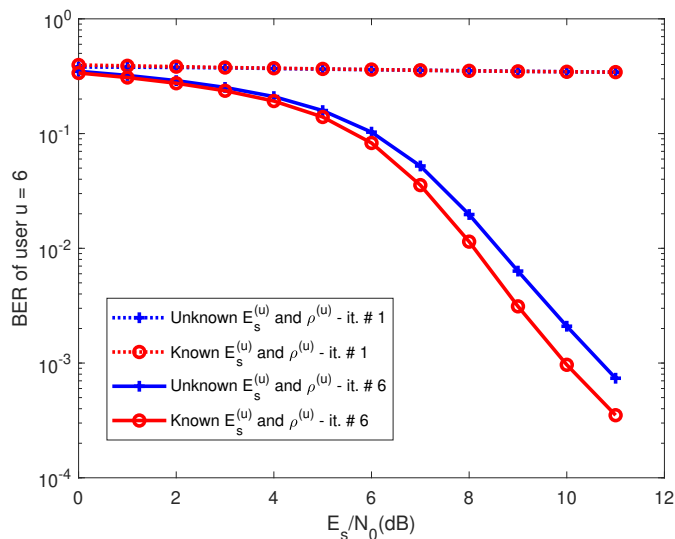


Fig. 15. BER under unknown hyperparameters - low energy user  $u = 6$ .

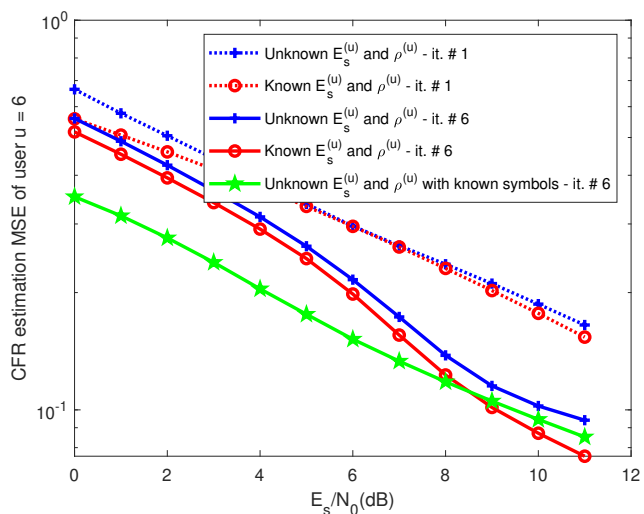


Fig. 16. CFR estimation MSE under unequal energy and unknown hyperparameters - low energy user  $u = 6$ .

Future work will consider how to adapt the present system model to grant-free massive access [11]. For instance, increasing the maximum number of supported users can be achieved by letting each user occupy only a subcarrier-chunk [28]. The application to other NOMA schemes [2] and more involved dynamical traffic with explicit modeling of the temporal correlation in the user activity state [5], [24], [26] would be of interest. Extensions of the proposed message-passing scheme to joint sensing and communication [53] or reconfigurable intelligent

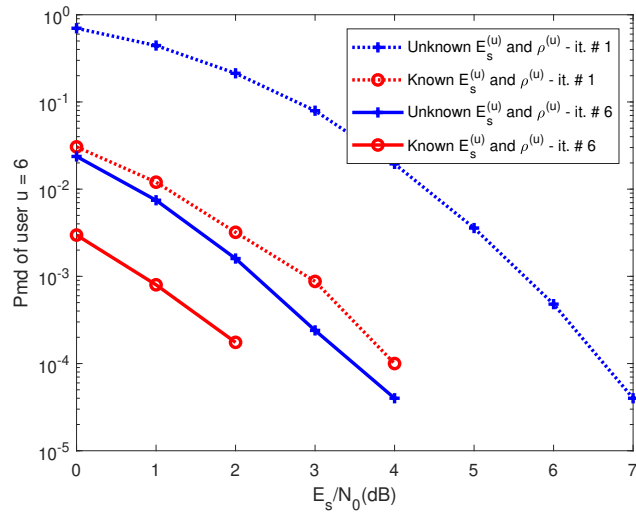


Fig. 17.  $P_{md}$  under unequal energy and unknown hyperparameters - low energy user  $u = 6$ .

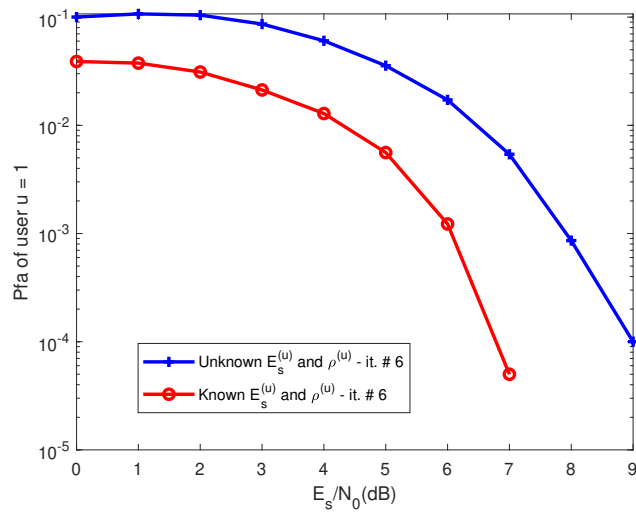


Fig. 18.  $P_{fa}$  under unequal energy and unknown hyperparameters - zero-energy user  $u = 1$ .

surfaces (RISs) technologies [54], [55] could also have an impact on future 6G networks. The applicability of other inference strategies with low implementation cost such as approximate message-passing (AMP) will also be investigated [9].

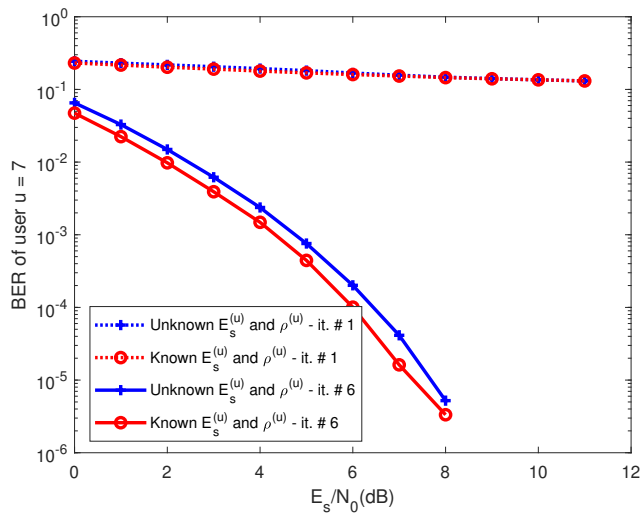


Fig. 19. BER under unequal energy and unknown hyperparameters - reference energy user  $u = 7$ .

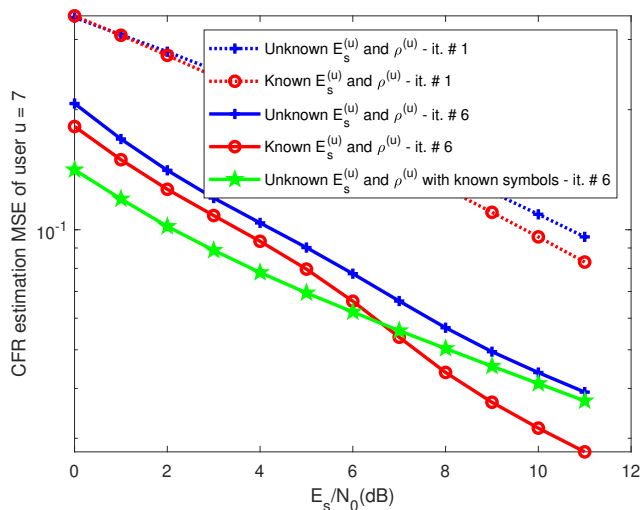


Fig. 20. CFR estimation MSE under unequal energy and unknown hyperparameters - reference energy user  $u = 7$ .

## APPENDIX A

### NEW PRINCIPLED CONDITIONAL GAUSSIAN CONTINUOUS MIXTURE REDUCTION

Let  $\boldsymbol{\pi}$  be a parameter vector with prior distribution  $p(\boldsymbol{\pi})$ . Consider a conditional Gaussian continuous mixture of the form [45, p. 239]

$$p(\mathbf{z}|\boldsymbol{\pi}) = \int_{\boldsymbol{\theta}} \boldsymbol{\omega}(\boldsymbol{\theta}) \mathcal{CN}(\mathbf{z}; \mathbf{m}(\boldsymbol{\theta}|\boldsymbol{\pi}), \boldsymbol{\Sigma}(\boldsymbol{\theta}|\boldsymbol{\pi})) d\boldsymbol{\theta}, \quad (41)$$

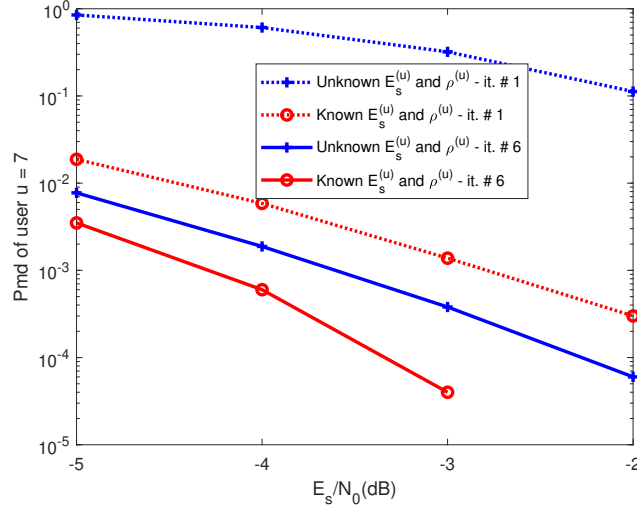


Fig. 21.  $P_{md}$  under unequal energy and unknown hyperparameters - reference energy user  $u = 7$ .

where  $\mathbf{z} \in \mathbb{C}^d$  ( $d$  is a strictly positive integer) and  $\int_{\boldsymbol{\theta}} \omega(\boldsymbol{\theta}) d\boldsymbol{\theta} = 1$ .

*Theorem A.1:* The Gaussian density of the form  $q(\mathbf{z}|\boldsymbol{\pi}) = \mathcal{CN}(\mathbf{z}; \mathbf{m}(\boldsymbol{\pi}), \boldsymbol{\Sigma})$  minimizing  $E_{p(\boldsymbol{\pi})}[KL(p||q)]$  (where  $KL(\cdot)$  denotes the Kullback-Leibler divergence) satisfies

$$\begin{aligned} \mathbf{m}(\boldsymbol{\pi}) &= \int_{\boldsymbol{\theta}} \omega(\boldsymbol{\theta}) \mathbf{m}(\boldsymbol{\theta}|\boldsymbol{\pi}) d\boldsymbol{\theta} \\ \boldsymbol{\Sigma} &= \int_{\boldsymbol{\pi}} \left[ \int_{\boldsymbol{\theta}} \omega(\boldsymbol{\theta}) ((\mathbf{m}(\boldsymbol{\pi}) - \mathbf{m}(\boldsymbol{\theta}|\boldsymbol{\pi}))(\mathbf{m}(\boldsymbol{\pi}) - \mathbf{m}(\boldsymbol{\theta}|\boldsymbol{\pi}))^H + \boldsymbol{\Sigma}(\boldsymbol{\theta}|\boldsymbol{\pi})) d\boldsymbol{\theta} \right] \cdot p(\boldsymbol{\pi}) d\boldsymbol{\pi}. \end{aligned} \quad (42)$$

*Proof:* Let us develop the criterion to be minimized as

$$\begin{aligned} E_{p(\boldsymbol{\pi})}[KL(p||q)] &= \int_{\boldsymbol{\pi}} \left[ \int_{\mathbf{z}} p(\mathbf{z}|\boldsymbol{\pi}) \ln \frac{p(\mathbf{z}|\boldsymbol{\pi})}{q(\mathbf{z}|\boldsymbol{\pi})} d\mathbf{z} \right] p(\boldsymbol{\pi}) d\boldsymbol{\pi} \\ &= - \int_{\boldsymbol{\pi}} \left[ \int_{\mathbf{z}} p(\mathbf{z}|\boldsymbol{\pi}) \ln q(\mathbf{z}|\boldsymbol{\pi}) d\mathbf{z} \right] p(\boldsymbol{\pi}) d\boldsymbol{\pi} + \int_{\boldsymbol{\pi}} \left[ \int_{\mathbf{z}} p(\mathbf{z}|\boldsymbol{\pi}) \ln p(\mathbf{z}|\boldsymbol{\pi}) d\mathbf{z} \right] p(\boldsymbol{\pi}) d\boldsymbol{\pi} \\ &= \int_{\boldsymbol{\pi}} \left[ \int_{\mathbf{z}} -p(\mathbf{z}|\boldsymbol{\pi}) \ln q(\mathbf{z}|\boldsymbol{\pi}) d\mathbf{z} \right] p(\boldsymbol{\pi}) d\boldsymbol{\pi} + C. \end{aligned} \quad (43)$$

Indeed, since the minimization is performed w.r.t.  $q(\cdot|\cdot)$ , the second double integral in (Eq. (43)) can be considered as a constant independent of  $\boldsymbol{\pi}$  that we call  $C$ . Also, replacing  $-\ln q(\mathbf{z}|\boldsymbol{\pi})$

by its expression, we obtain

$$\begin{aligned} -\ln q(\mathbf{z}|\boldsymbol{\pi}) &= d \ln \pi + \ln |\boldsymbol{\Sigma}| + (\mathbf{z} - \mathbf{m}(\boldsymbol{\pi}))^H \boldsymbol{\Sigma}^{-1} (\mathbf{z} - \mathbf{m}(\boldsymbol{\pi})) \\ &= d \ln \pi + \ln |\boldsymbol{\Sigma}| + \mathbf{z}^H \boldsymbol{\Sigma}^{-1} \mathbf{z} - \mathbf{m}(\boldsymbol{\pi})^H \boldsymbol{\Sigma}^{-1} \mathbf{z} - \mathbf{z}^H \boldsymbol{\Sigma}^{-1} \mathbf{m}(\boldsymbol{\pi}) + \mathbf{m}(\boldsymbol{\pi})^H \boldsymbol{\Sigma}^{-1} \mathbf{m}(\boldsymbol{\pi}). \end{aligned} \quad (44)$$

Using the property  $\text{trace}(\mathbf{AB}) = \text{trace}(\mathbf{BA})$  for any matrices of suitable dimensions  $\mathbf{A}$  and  $\mathbf{B}$ , we obtain an alternative expression as

$$-\ln q(\mathbf{z}|\boldsymbol{\pi}) = d \ln \pi + \ln |\boldsymbol{\Sigma}| + \text{trace}(\boldsymbol{\Sigma}^{-1} (\mathbf{z} - \mathbf{m}(\boldsymbol{\pi})) (\mathbf{z} - \mathbf{m}(\boldsymbol{\pi}))^H). \quad (45)$$

Injecting (44) into the last line of (43), since  $p(\boldsymbol{\pi}) \geq 0 \quad \forall \boldsymbol{\pi}$ , optimization w.r.t.  $\mathbf{m}(\boldsymbol{\pi})$  is equivalent to minimizing the inner integral inside the brackets, which results in solving

$$\int_{\mathbf{z}} \frac{\partial}{\partial \mathbf{m}(\boldsymbol{\pi})} \left[ -\mathbf{m}(\boldsymbol{\pi})^H \boldsymbol{\Sigma}^{-1} \mathbf{z} - \mathbf{z}^H \boldsymbol{\Sigma}^{-1} \mathbf{m}(\boldsymbol{\pi}) + \mathbf{m}(\boldsymbol{\pi})^H \boldsymbol{\Sigma}^{-1} \mathbf{m}(\boldsymbol{\pi}) \right] p(\mathbf{z}|\boldsymbol{\pi}) d\mathbf{z} = \mathbf{0}, \quad (46)$$

whose unique solution is (see [49, Tab. 20.4])

$$\mathbf{m}(\boldsymbol{\pi}) = \int_{\mathbf{z}} \mathbf{z} p(\mathbf{z}|\boldsymbol{\pi}) d\mathbf{z}, \quad (47)$$

which in turn admits the closed form given in (Eq. 42) for the continuous Gaussian mixture  $p(\mathbf{z}|\boldsymbol{\pi})$  [43, p. 106-108].

Similarly, injecting (45) into the last line of (43), results in solving

$$\int_{\boldsymbol{\pi}} \left[ \int_{\mathbf{z}} \frac{\partial}{\partial \boldsymbol{\Sigma}} \left( \text{trace}(\boldsymbol{\Sigma}^{-1} (\mathbf{z} - \mathbf{m}(\boldsymbol{\pi})) (\mathbf{z} - \mathbf{m}(\boldsymbol{\pi}))^H) + \ln |\boldsymbol{\Sigma}| \right) p(\mathbf{z}|\boldsymbol{\pi}) d\mathbf{z} \right] p(\boldsymbol{\pi}) d\boldsymbol{\pi} = \mathbf{0}, \quad (48)$$

whose unique solution is (see [49, Tab. 20.4 and Tab. 20.5])

$$\boldsymbol{\Sigma} = \int_{\boldsymbol{\pi}} \left[ \int_{\mathbf{z}} (\mathbf{z} - \mathbf{m}(\boldsymbol{\pi})) (\mathbf{z} - \mathbf{m}(\boldsymbol{\pi}))^H p(\mathbf{z}|\boldsymbol{\pi}) d\mathbf{z} \right] p(\boldsymbol{\pi}) d\boldsymbol{\pi}, \quad (49)$$

which in turn admits the closed form given in (Eq. 42) for the continuous Gaussian mixture  $p(\mathbf{z}|\boldsymbol{\pi})$  [43, p. 106-108]. ■

## APPENDIX B

### LIKELIHOOD OF $d_n^{(u)}$ AND $\mathbf{x}_n^{(u)}$ NEEDED IN EP USING THM. A.1

The likelihood of  $d_n^{(u)}$  can be written as

$$\tilde{f}(\mathbf{y}_n | d_n^{(u)}) = \int_{\mathcal{C}^U} \int_{\mathcal{C}^{U-1}} \int_{\mathcal{C}^{U N_R}} g_n \prod_{u'=1}^U \mu_{\theta^{(u')} \rightarrow g_n}^{Ga-BP}(\theta^{(u')}) \prod_{u' \neq u} \mu_{d_n^{(u')} \rightarrow g_n}^{EP}(d_n^{(u')}) \prod_{u'=1}^U \mu_{\mathbf{x}_n^{(u')} \rightarrow g_n}^{EP}(\mathbf{x}_n^{(u')}) \prod_{u'=1}^U d\theta^{(u')} \prod_{u' \neq u} dd_n^{(u')} \prod_{u'=1}^U d\mathbf{x}_n^{(u')} \quad (50)$$

Using the observation model (4) and replacing the messages incoming to  $g_n$  by (10), we obtain

$$\begin{aligned} \tilde{f}(\mathbf{y}_n | d_n^{(u)}) &= \int_{\mathcal{C}^U} \int_{\mathcal{C}^{U-1}} \int_{\mathcal{C}^{UN_R}} \mathcal{CN} \left( \mathbf{y}_n; \sum_{u'=1}^U \theta^{(u')} d_n^{(u')} \mathbf{x}_n^{(u')}, \mathbf{R} \right) \prod_{u'=1}^U \mathcal{CN}(\theta^{(u')}; m_{\theta^{(u')} \rightarrow g_n}, \sigma_{\theta^{(u')} \rightarrow g_n}^2) \\ &\cdot \prod_{u' \neq u} \mathcal{CN}(d_n^{(u')}; m_{d_n^{(u')} \rightarrow g_n}, \sigma_{d_n^{(u')} \rightarrow g_n}^2) \prod_{u'=1}^U \mathcal{CN}(\mathbf{x}_n^{(u')}; \mathbf{m}_{\mathbf{x}_n^{(u')} \rightarrow g_n}, \Sigma_{\mathbf{x}_n^{(u')} \rightarrow g_n}) \prod_{u'=1}^U d\theta^{(u')} \prod_{u' \neq u} dd_n^{(u')} \prod_{u'=1}^U d\mathbf{x}_n^{(u')}, \end{aligned} \quad (51)$$

which, using [43, p. 38], simplifies to

$$\begin{aligned} \tilde{f}(\mathbf{y}_n | d_n^{(u)}) &= \int_{\mathcal{C}^U} \int_{\mathcal{C}^{UN_R}} \mathcal{CN} \left( \mathbf{y}_n; \theta^{(u)} d_n^{(u)} \mathbf{x}_n^{(u)} + \sum_{u' \neq u} \theta^{(u')} m_{d_n^{(u')} \rightarrow g_n} \mathbf{x}_n^{(u')}, \sum_{u' \neq u} |\theta^{(u')}|^2 \sigma_{d_n^{(u')} \rightarrow g_n}^2 \mathbf{x}_n^{(u')} \mathbf{x}_n^{(u')H} + \mathbf{R} \right) \\ &\cdot \prod_{u'=1}^U \mathcal{CN}(\theta^{(u')}; m_{\theta^{(u')} \rightarrow g_n}, \sigma_{\theta^{(u')} \rightarrow g_n}^2) \prod_{u'=1}^U \mathcal{CN}(\mathbf{x}_n^{(u')}; \mathbf{m}_{\mathbf{x}_n^{(u')} \rightarrow g_n}, \Sigma_{\mathbf{x}_n^{(u')} \rightarrow g_n}) \prod_{u'=1}^U d\theta^{(u')} \prod_{u'=1}^U d\mathbf{x}_n^{(u')}. \end{aligned} \quad (52)$$

This conditional Gaussian continuous mixture can be approximated by the following Gaussian, applying Thm. A.1 (setting  $\boldsymbol{\pi} = d_n^{(u)}$  and  $\boldsymbol{\theta} = \{\theta^{(u')}, \mathbf{x}_n^{(u')}\}_{u'=1}^U$ )

$$\tilde{f}(\mathbf{y}_n | d_n^{(u)}) \approx \mathcal{CN}(\mathbf{y}_n; \mathbf{m}_{d_n^{(u)}}(d_n^{(u)}), \Sigma_{d_n^{(u)}}), \quad (53)$$

where

$$\mathbf{m}_{d_n^{(u)}}(d_n^{(u)}) = \mathbf{h}_{d_n^{(u)}} d_n^{(u)} + \mathbf{I}_{d_n^{(u)}}, \quad \mathbf{h}_{d_n^{(u)}} = m_{\theta^{(u)} \rightarrow g_n} \mathbf{m}_{\mathbf{x}_n^{(u)} \rightarrow g_n}, \quad \mathbf{I}_{d_n^{(u)}} = \sum_{u' \neq u} m_{\theta^{(u')} \rightarrow g_n} m_{d_n^{(u')} \rightarrow g_n} \mathbf{m}_{\mathbf{x}_n^{(u')} \rightarrow g_n}, \quad (54)$$

in which the first (resp. second) term accounts for the  $u$ -th user's useful signal conditional on  $d_n^{(u)}$  (resp. the average multi-access interference affecting the  $u$ -th user) over the  $n$ -th subcarrier.

Moreover,

$$\begin{aligned} \Sigma_{d_n^{(u)}} &= \left[ |m_{\theta^{(u)} \rightarrow g_n}|^2 \Sigma_{\mathbf{x}_n^{(u)} \rightarrow g_n} + \sigma_{\theta^{(u)} \rightarrow g_n}^2 (\mathbf{m}_{\mathbf{x}_n^{(u)} \rightarrow g_n} \mathbf{m}_{\mathbf{x}_n^{(u)} \rightarrow g_n}^H + \Sigma_{\mathbf{x}_n^{(u)} \rightarrow g_n}) \right] \\ &+ \sum_{u' \neq u} \sigma_{d_n^{(u')} \rightarrow g_n}^2 (|m_{\theta^{(u')} \rightarrow g_n}|^2 + \sigma_{\theta^{(u')} \rightarrow g_n}^2) (\mathbf{m}_{\mathbf{x}_n^{(u')} \rightarrow g_n} \mathbf{m}_{\mathbf{x}_n^{(u')} \rightarrow g_n}^H + \Sigma_{\mathbf{x}_n^{(u')} \rightarrow g_n}) \\ &+ \sum_{u' \neq u} |m_{d_n^{(u')} \rightarrow g_n}|^2 \left[ |m_{\theta^{(u')} \rightarrow g_n}|^2 \Sigma_{\mathbf{x}_n^{(u')} \rightarrow g_n} + \sigma_{\theta^{(u')} \rightarrow g_n}^2 (\mathbf{m}_{\mathbf{x}_n^{(u')} \rightarrow g_n} \mathbf{m}_{\mathbf{x}_n^{(u')} \rightarrow g_n}^H + \Sigma_{\mathbf{x}_n^{(u')} \rightarrow g_n}) \right] + \mathbf{R}, \end{aligned} \quad (55)$$

where the first term (resp. two next terms) accounts for the residual uncertainty on the user activity variable, data symbol and CFR affecting the  $u$ -th user's signal over the  $n$ -th subcarrier (resp. the residual uncertainty of the hidden variables concerning all other users). Also,  $\mathbf{R}$  accounts for the channel noise covariance.



Similarly, applying Thm. A.1 (setting  $\boldsymbol{\pi} = \mathbf{x}_n^{(u)}$  and  $\boldsymbol{\theta} = \{d_n^{(u')}, \theta^{(u')}\}_{u'=1}^U$ ), the likelihood of  $\mathbf{x}_n^{(u)}$  can be approximated in the form (20) with parameters

$$\mathbf{m}_{\mathbf{x}_n^{(u)}}(\mathbf{x}_n^{(u)}) = \mathbf{h}_{\mathbf{x}_n^{(u)}} \mathbf{x}_n^{(u)} + \mathbf{I}_{d_n^{(u)}}, \quad \mathbf{h}_{\mathbf{x}_n^{(u)}} = m_{\theta^{(u)} \rightarrow g_n} m_{d_n^{(u)} \rightarrow g_n}. \quad (56)$$

$$\begin{aligned} \Sigma_{\mathbf{x}_n^{(u)}} &= E_s^{(u)} \boldsymbol{\Gamma}^{(u)} \left[ |m_{\theta^{(u)} \rightarrow g_n}|^2 \sigma_{d_n^{(u)} \rightarrow g_n}^2 + \sigma_{\theta^{(u)} \rightarrow g_n}^2 (|m_{d_n^{(u)} \rightarrow g_n}|^2 + \sigma_{d_n^{(u)} \rightarrow g_n}^2) \right] \\ &+ \sum_{u' \neq u} \Sigma_{\mathbf{x}_n^{(u')} \rightarrow g_n} (|m_{\theta^{(u')} \rightarrow g_n}|^2 + \sigma_{\theta^{(u')} \rightarrow g_n}^2) (|m_{d_n^{(u')} \rightarrow g_n}|^2 + \sigma_{d_n^{(u')} \rightarrow g_n}^2) \\ &+ \sum_{u' \neq u} \left[ |m_{\theta^{(u')} \rightarrow g_n}|^2 \sigma_{d_n^{(u')} \rightarrow g_n}^2 + \sigma_{\theta^{(u')} \rightarrow g_n}^2 (|m_{d_n^{(u')} \rightarrow g_n}|^2 + \sigma_{d_n^{(u')} \rightarrow g_n}^2) \right] \mathbf{m}_{\mathbf{x}_n^{(u')} \rightarrow g_n} \mathbf{m}_{\mathbf{x}_n^{(u')} \rightarrow g_n}^H + \mathbf{R}. \end{aligned} \quad (57)$$

## APPENDIX C

### LIKELIHOOD OF $\theta^{(u)}$ NEEDED IN BP

The likelihood of  $\theta^{(u)}$  has the form of a continuous Gaussian mixture

$$\tilde{f}(\mathbf{y}_n | \theta^{(u)}) = \int_{\mathcal{C}^{U-1}} \int_{\mathcal{C}^U} \int_{\mathcal{C}^{UN_r}} g_n \prod_{u' \neq u} \mu_{\theta^{(u')} \rightarrow g_n}^{Ga-BP}(\theta^{(u')}) \prod_{u'=1}^U \mu_{d_n^{(u')} \rightarrow g_n}^{EP}(d_n^{(u')}) \prod_{u'=1}^U \mu_{\mathbf{x}_n^{(u')} \rightarrow g_n}^{EP}(\mathbf{x}_n^{(u')}) \prod_{u' \neq u} d\theta^{(u')} \prod_{u'=1}^U dd_n^{(u')} \prod_{u'=1}^U d\mathbf{x}_n^{(u')} \quad (58)$$

Using ordinary moment-matching [43, p. 106-108], we obtain a Gaussian approximation of the form (28), whose parameters are given below

$$\mathbf{m}_{\theta^{(u)}, n}(\theta^{(u)}) = \theta^{(u)} m_{d_n^{(u)} \rightarrow g_n} \mathbf{m}_{\mathbf{x}_n^{(u)} \rightarrow g_n} + \mathbf{I}_{d_n^{(u)}}. \quad (59)$$

$$\begin{aligned} \Sigma_{\theta^{(u)}, n}(\theta^{(u)}) &= |\theta^{(u)}|^2 \left[ |m_{d_n^{(u)} \rightarrow g_n}|^2 \Sigma_{\mathbf{x}_n^{(u)} \rightarrow g_n} + \sigma_{d_n^{(u)} \rightarrow g_n}^2 (\mathbf{m}_{\mathbf{x}_n^{(u)} \rightarrow g_n} \mathbf{m}_{\mathbf{x}_n^{(u)} \rightarrow g_n}^H + \Sigma_{\mathbf{x}_n^{(u)} \rightarrow g_n}) \right] \\ &+ \sum_{u' \neq u} \sigma_{\theta^{(u')} \rightarrow g_n}^2 (|m_{d_n^{(u')} \rightarrow g_n}|^2 + \sigma_{d_n^{(u')} \rightarrow g_n}^2) (\mathbf{m}_{\mathbf{x}_n^{(u')} \rightarrow g_n} \mathbf{m}_{\mathbf{x}_n^{(u')} \rightarrow g_n}^H + \Sigma_{\mathbf{x}_n^{(u')} \rightarrow g_n}) \\ &+ \sum_{u' \neq u} |m_{\theta^{(u')} \rightarrow g_n}|^2 \left[ |m_{d_n^{(u')} \rightarrow g_n}|^2 \Sigma_{\mathbf{x}_n^{(u')} \rightarrow g_n} + \sigma_{d_n^{(u')} \rightarrow g_n}^2 (\mathbf{m}_{\mathbf{x}_n^{(u')} \rightarrow g_n} \mathbf{m}_{\mathbf{x}_n^{(u')} \rightarrow g_n}^H + \Sigma_{\mathbf{x}_n^{(u')} \rightarrow g_n}) \right] + \mathbf{R}. \end{aligned} \quad (60)$$

Note that both the mean and the covariance depend on  $\theta^{(u)}$ . This approximation being more precise than the one in EP via Thm. A.1 (where only the mean would depend on  $\theta^{(u)}$ , while the covariance would be a constant) happens to be crucial for proper UAD applying (34).

## REFERENCES

- [1] M. Bennis, M. Debbah and H. V. Poor, "Ultrareliable and Low-Latency Wireless Communication: Tail, Risk, and Scale," in *Proceedings of the IEEE*, vol. 106, no. 10, pp. 1834-1853, Oct. 2018.
- [2] M. B. Shahab, R. Abbas, M. Shirvanimoghaddam and S. J. Johnson, "Grant-Free Non-Orthogonal Multiple Access for IoT: A Survey," in *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1805-1838, thirdquarter 2020.
- [3] C. Bockelmann, "Iterative Soft Interference Cancellation for Sparse BPSK Signals," in *IEEE Communications Letters*, vol. 19, no. 5, pp. 855-858, May 2015.
- [4] W. Kim, Y. Ahn and B. Shim, "Deep Neural Network-Based Active User Detection for Grant-Free NOMA Systems," in *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2143-2155, April 2020.
- [5] H. F. Schepker, C. Bockelmann and A. Dekorsy, "Efficient Detectors for Joint Compressed Sensing Detection and Channel Decoding," in *IEEE Transactions on Communications*, vol. 63, no. 6, pp. 2249-2260, June 2015.
- [6] B. Wang, L. Dai, T. Mir and Z. Wang, "Joint User Activity and Data Detection Based on Structured Compressive Sensing for NOMA," in *IEEE Communications Letters*, vol. 20, no. 7, pp. 1473-1476, July 2016.
- [7] B. Wang, L. Dai, Y. Zhang, T. Mir and J. Li, "Dynamic Compressive Sensing-Based Multi-User Detection for Uplink Grant-Free NOMA," in *IEEE Communications Letters*, vol. 20, no. 11, pp. 2320-2323, Nov. 2016.
- [8] Y. Mei, Z. Gao, Y. Wu, W. Chen, J. Zhang, D.W.K. Ng and M. Di Renzo., "Compressive Sensing-Based Joint Activity and Data Detection for Grant-Free Massive IoT Access," in *IEEE Transactions on Wireless Communications*, vol. 21, no. 3, pp. 1851-1869, March 2022.
- [9] C. Wei, H. Liu, Z. Zhang, J. Dang and L. Wu, "Approximate Message Passing-Based Joint User Activity and Data Detection for NOMA," in *IEEE Communications Letters*, vol. 21, no. 3, pp. 640-643, March 2017.
- [10] J. Zhang, Y. Pan and J. Xu, "Compressive Sensing for Joint User Activity and Data Detection in Grant-Free NOMA," in *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 857-860, June 2019.
- [11] K. Senel and E. G. Larsson, "Grant-Free Massive MTC-Enabled Massive MIMO: A Compressive Sensing Approach," in *IEEE Transactions on Communications*, vol. 66, no. 12, pp. 6164-6175, Dec. 2018.
- [12] A. Bayesteh, E. Yi, H. Nikopour and H. Baligh, "Blind detection of SCMA for uplink grant-free multiple-access," 2014 11th International Symposium on Wireless Communications Systems (ISWCS), 2014, pp. 853-857.
- [13] Q. Zou, H. Zhang, D. Cai and H. Yang, "A Low-Complexity Joint User Activity, Channel and Data Estimation for Grant-Free Massive MIMO Systems," in *IEEE Signal Processing Letters*, vol. 27, pp. 1290-1294, 2020.
- [14] A. T. Abebe and C. G. Kang, "Joint Channel Estimation and MUD for Scalable Grant-Free Random Access," in *IEEE Communications Letters*, vol. 23, no. 12, pp. 2229-2233, Dec. 2019.
- [15] J. Ding and J. Choi, "Comparison of Preamble Structures for Grant-Free Random Access in Massive MIMO Systems," in *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 166-170, Feb. 2020.
- [16] T. Ding, X. Yuan and S. C. Liew, "Sparsity Learning-Based Multiuser Detection in Grant-Free Massive-Device Multiple Access," in *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3569-3582, July 2019.
- [17] Y. Han, B. D. Rao and J. Lee, "Massive Uncoordinated Access With Massive MIMO: A Dictionary Learning Approach," in *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 1320-1332, Feb. 2020.
- [18] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen and L. Hanzo, "A Survey of Non-Orthogonal Multiple Access for 5G," in *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2294-2323, thirdquarter 2018.
- [19] J. Goseling, M. Gastpar and J. H. Weber, "Random Access With Physical-Layer Network Coding," in *IEEE Transactions on Information Theory*, vol. 61, no. 7, pp. 3670-3681, July 2015,

- [20] E. Paolini, C. Stefanovic, G. Liva and P. Popovski, "Coded random access: applying codes on graphs to design random access protocols," in *IEEE Communications Magazine*, vol. 53, no. 6, pp. 144-150, June 2015.
- [21] M. J. Wainwright and M. I. Jordan, "Graphical Models, Exponential Families, and Variational Inference", *Foundations and Trends in Machine Learning: Vol. 1: No. 1-2*, pp 1-305, 2008.
- [22] S. Sharma, K. Deka and Y. Hong, "User Activity Detection-Based Large SCMA System for Uplink Grant-Free Access," 2019 *IEEE International Conference on Communications Workshops (ICC Workshops)*, 2019, pp. 1-6.
- [23] Y. Zhang, Q. Guo, Z. Wang, J. Xi and N. Wu, "Block Sparse Bayesian Learning Based Joint User Activity Detection and Channel Estimation for Grant-Free NOMA Systems," in *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9631-9640, Oct. 2018.
- [24] W. Zhu, M. Tao, X. Yuan and Y. Guan, "Message Passing-Based Joint User Activity Detection and Channel Estimation for Temporally-Correlated Massive Access," in *IEEE Transactions on Communications*, vol. 71, no. 6, pp. 3576-3591, June 2023.
- [25] F. Lehmann, "Joint User Activity Detection, Channel Estimation, and Decoding for Multiuser/Multiantenna OFDM Systems," in *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8263-8275, Sept. 2018
- [26] W. Yuan, N. Wu, Q. Guo, D. W. K. Ng, J. Yuan and L. Hanzo, "Iterative Joint Channel Estimation, User Activity Tracking, and Data Detection for FTN-NOMA Systems Supporting Random Access," in *IEEE Transactions on Communications*, vol. 68, no. 5, pp. 2963-2977, May 2020.
- [27] Z. Yuan, C. Zhang, Z. Wang, Q. Guo and J. Xi, "An Auxiliary Variable-Aided Hybrid Message Passing Approach to Joint Channel Estimation and Decoding for MIMO-OFDM," in *IEEE Signal Processing Letters*, vol. 24, no. 1, pp. 12-16, Jan. 2017.
- [28] Y. Zhang, Z. Yuan, Q. Guo, Z. Wang, J. Xi and Y. Li, "Bayesian Receiver Design for Grant-Free NOMA With Message Passing Based Structured Signal Estimation," in *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8643-8656, Aug. 2020.
- [29] F. Wei, W. Chen, Y. Wu, J. Ma and T. A. Tsiftsis, "Message-Passing Receiver Design for Joint Channel Estimation and Data Decoding in Uplink Grant-Free SCMA Systems," in *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 167-181, Jan. 2019.
- [30] F. Sagheer, F. Lehmann and A. O. Berthet, "Low-Complexity Dynamic Channel Estimation in Multi-Antenna Grant-Free NOMA," 2022 *IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, Helsinki, Finland, 2022, pp. 1-7.
- [31] L. Liu, J. Tong and L. Ping, "Analysis and optimization of CDMA systems with chip-level interleavers," in *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 1, pp. 141-150, Jan. 2006.
- [32] L. Ping, Q. Guo and J. Tong, "The OFDM-IDMA approach to wireless communication systems," in *IEEE Wireless Communications*, vol. 14, no. 3, pp. 18-24, June 2007.
- [33] P. Hammarberg, F. Rusek and O. Edfors, "Channel Estimation Algorithms for OFDM-IDMA: Complexity and Performance," in *IEEE Transactions on Wireless Communications*, vol. 11, no. 5, pp. 1722-1732, May 2012
- [34] S. Kim, H. Kim, H. Noh, Y. Kim and D. Hong, "Novel Transceiver Architecture for an Asynchronous Grant-Free IDMA System," in *IEEE Transactions on Wireless Communications*, vol. 18, no. 9, pp. 4491-4504, Sept. 2019.
- [35] C. Novak, G. Matz and F. Hlawatsch, "IDMA for the Multiuser MIMO-OFDM Uplink: A Factor Graph Framework for Joint Data Detection and Channel Estimation," in *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 4051-4066, Aug.15, 2013.
- [36] J. Hu, H.-A. Loeliger, J. Dauwels and F. Kschischang, "A general computation rule for lossy summaries/messages with

- examples from equalization,” *Proc. 44th Allerton Conf. on Communication, Control and Computing*, Monticello, Illinois, Sept. 2006.
- [37] D. Bickson and D. Dolev, O. Shental, P.H. Siegel and J.K. Wolf, “Gaussian belief propagation based multiuser detection,” *Proc. ISIT 08*, pp. 1878-1882, Toronto, Canada, Jul. 2008.
- [38] N. Czink, B. Bandemer, C. Oestges, T. Zemen and A. Paulraj, ”Analytical Multi-User MIMO Channel Modeling: Subspace Alignment Matters,” in *IEEE Transactions on Wireless Communications*, vol. 11, no. 1, pp. 367-377, January 2012.
- [39] C. Knievel, P. A. Hoeher, A. Tyrrell and G. Auer, ”Multi-Dimensional Graph-Based Soft Iterative Receiver for MIMO-OFDM,” in *IEEE Transactions on Communications*, vol. 60, no. 6, pp. 1599-1609, June 2012.
- [40] F. R. Kschischang, B. J. Frey and H.-A. Loeliger, ”Factor graphs and the sum-product algorithm,” in *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498-519, Feb 2001.
- [41] T. Minka, ”Divergence Measures and Message Passing,” Microsoft Research Cambridge, MSR-TR-2005-173, pp. 1-17, January 2005.
- [42] P. Sun, C. Zhang, Z. Wang, C. N. Manchón and B. H. Fleury, ”Iterative Receiver Design for ISI Channels Using Combined Belief- and Expectation-Propagation,” in *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1733-1737, Oct. 2015.
- [43] H. Tanizaki, *Nonlinear filters: estimation and applications*, Berlin, Germany: Springer, 1996.
- [44] D.C. Fraser and J.E. Potter, “The optimum linear smoother as a combination of two optimum linear filters,” *IEEE Trans. Automat. Contr.*, pp. 387-390, vol. 14, no. 4, Aug. 1969.
- [45] J.J. Shink, *Probability, Random Variables, and Random Processes: Theory and Signal Processing Applications*, Hoboken, N.J.: John Wiley & Sons, 2013.
- [46] M. Senst and G. Ascheid, ”How the Framework of Expectation Propagation Yields an Iterative IC-LMMSE MIMO Receiver,” 2011 IEEE Global Telecommunications Conference - GLOBECOM 2011, 2011, pp. 1-6.
- [47] M. Ruan, M.C. Reed, and Z. Shi, “Successive multiuser detection and interference cancelation for contention based OFDMA ranging channel,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 481-487, Feb. 2010.
- [48] Q. Wang and G. Ren, “Iterative maximum likelihood detection for initial ranging process in 802.16 OFDMA systems,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2778-2787, May 2015.
- [49] P. Vaidyanathan, S. Phoong and Y. Lin, *Signal Processing and Optimization for Transceiver Systems*, Cambridge: Cambridge University Press, 2010.
- [50] S. Bochkhanov. Bound and linear equality/inequality constrained optimization. (Online; accessed in April 2023): <http://www.alglib.net/optimization/boundandlinearlyconstrained.php>.
- [51] P. Händel, A. Eriksson and T. Wigren, “Performance analysis of a correlation based single tone frequency estimator,” *Elsevier Signal Processing*, vol. 44, iss. 2, pp. 223-231, June 1995.
- [52] J. Dauwels and H.-A. Loeliger, “Phase estimation by message passing,” 2004 IEEE International Conference on Communications (ICC), Paris, France, 2004, pp. 523-527 Vol.1.
- [53] X. Tong, Z. Zhang, J. Wang, C. Huang and M. Debbah, “Joint Multi-User Communication and Sensing Exploiting Both Signal and Environment Sparsity,” in *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1409-1422, Nov. 2021.
- [54] L. Wei, C. Huang, G. C. Alexandropoulos, C. Yuen, Z. Zhang and M. Debbah, “Channel Estimation for RIS-Empowered Multi-User MISO Wireless Communications,” in *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 4144-4157, June 2021.
- [55] L. Wei et al., “Joint Channel Estimation and Signal Recovery for RIS-Empowered Multiuser Communications,” in *IEEE Transactions on Communications*, vol. 70, no. 7, pp. 4640-4655, July 2022.