

RESEARCH

Open Access

Sparse code multiple access on the generalized frequency division multiplexing



Guilherme P. Aquino^{*†}  and Luciano L. Mendes[†]

Abstract

Recent advances in the communication systems culminated in a new class of multiple access schemes, named non-orthogonal multiple access (NOMA), where the main goal is to increase the spectrum efficiency by overlapping data from different users in a single time-frequency resource used by the physical layer. NOMA receivers can resolve the interference among data symbols from different users, increasing the overall system spectrum efficiency without introducing symbol error rate (SER) performance loss, which makes this class of multiple access techniques interesting for future mobile communication systems. This paper analyzes one promising NOMA technique, called sparse code multiple access (SCMA), where C users can share $U < C$ time-frequency resources from the physical layer. Initially, the SCMA and orthogonal frequency division multiplexing (OFDM) integration is considered, defining a benchmark for the overall SER performance for the multiple access technique. Furthermore, this paper proposes the SCMA and generalized frequency division multiplexing (GFDM) integration. Since GFDM is a highly flexible non-orthogonal waveform that can mimic several other waveforms as corner cases, it is an interesting candidate for future wireless communication systems. This paper proposes two approaches for combining SCMA and GFDM. The first one combines a soft equalizer, called block expectation propagation (BEP), and a multi-user detection (MUD) scheme based on the sum-product algorithm (SPA). This approach achieves the best SER performance, but with the significant increment of the complexity at the receiver. In the second approach, BEP is integrated with a simplified MUD, which is an original contribution of this paper, aiming for reducing the receiver's complexity at the cost of SER performance loss. The solutions proposed in this paper show that SCMA-GFDM can be an interesting solution for future mobile networks.

Keywords: NOMA, SCMA, GFDM, BEP, SPA, Simplified MUD

1 Introduction

New key technologies have been recently investigated in order to achieve the requirements of the future mobile networks. The recently proposed fifth generation of mobile networks (5G) has contrasting requirements, such as low latency, massive connectivity, and high spectral efficiency [1]. The high demand for mobile broadband and the advent of Internet of Things (IoT) [2] have imposed several challenges to the 5G radio access networks (RANs). One important requirement for 5G networks is the higher spectrum efficiency when compared

with previous generations of mobile communication networks [3]. New waveforms and innovative multiple access schemes are being pointed out as the main enablers for improving the spectrum efficiency and capacity of the spectrum efficiency and capacity of the RANs [4].

Recently, OFDM [5] has been adopted by the 3rd Generation Partnership Project (3GPP) as the 5G air interface for the 5G new radio (NR) [6]. However, OFDM has several restrictions [7], and it lacks the necessary flexibility to fulfill the requirements of the future wireless systems. The discussion around sixth generation of mobile networks (6G) networks [8] has already started, and non-orthogonal waveforms present the flexibility to cover the application scenarios foreseen for the next mobile networks.

*Correspondence: guilhermeaquino@inatel.br

[†]Guilherme P. Aquino and Luciano L. Mendes contributed equally to this work. Radiocommunication Reference Center, Instituto Nacional de Telecomunicações, Santa Rita do Sapucaí, Brazil

Among the several candidate waveforms available in the literature [9–11], GFDM is an interesting option due to its flexibility and overall performance in mobile applications [7]. GFDM is a multicarrier waveform that employs K individually filtered subcarriers to reduce the out-of-band emission (OOBE). Each subcarrier carries M data symbols, which means that each GFDM symbol transmits $N = KM$ symbols. A critically sampled prototype filter is circularly shifted in time and frequency to carry all data symbols, resulting in a N sample-long GFDM block. A single cyclic prefix (CP) is added to protect the entire GFDM block against the channel delay profile, reducing the waveform overhead [12]. GFDM is a flexible waveform that can be tailored for different applications, and it can cover several other waveforms as corner cases, such as single carrier frequency domain equalization (SCFDE), OFDM, and filterbank multicarrier (FBMC) [13]. GFDM provides a degree of flexibility that is unseen in any waveform. Besides the flexibility, the interaction among the GFDM subsymbols and subcarriers adds diversity that can help overcoming the doubly dispersive channels when non-linear detectors are employed on the receiver side [14]. The benefits introduced by GFDM make this waveform an interesting candidate for future mobile networks, and therefore, it will be considered as the baseline time-frequency waveform for the physical layer in this paper.

Besides the waveform, the multiple access scheme also plays an important role in the overall spectrum efficiency of the mobile network. The conventional multiple access schemes used in 5G networks, called orthogonal multiple access (OMA), clearly do not address the challenges foreseen for the 6G networks [15]. In all OMA schemes, such as frequency division multiple access (FDMA) and time division multiple access (TDMA), there is no mutual interference among the users, simplifying the receiver's structure. However, these schemes cannot reach the multi-user channel capacity of the radio interface [16]. Therefore, in scenarios with massive connectivity, the OMA schemes will demand wider bandwidth to accommodate the data of all users, contradicting the premise of spectral efficiency for beyond 5G (B5G) and 6G networks.

NOMA schemes take a completely different approach compared to OMA by leaving the orthogonality principle and allowing the interaction among the user's data. The most recent multiple access techniques allow for increasing the number of users sharing the available resources without increasing the bandwidth. This principle is called overloading, and it can increment the spectrum efficiency by allowing more users to share the same time-frequency resources [17]. In the overloaded RAN, the data symbols of different users can be superimposed in the time-frequency resources, enhancing the capacity and spectral efficiency of the system. NOMA schemes introduce a controllable mutual interference among the users. This

interference must be processed by the receiver to avoid bit error rate (BER) performance losses. Therefore, the cost for the higher spectrum efficiency provided by the NOMA schemes is the receiver complexity to deal with the interference [18].

Several NOMA techniques have been recently presented in the literature. The main ones are pattern division multiple access (PDMA) [19], interleaved division multiple access (IDMA) [20], multi-user shared access (MUSA) [21], lattice partition multiple access (LPMA) [22], and SCMA [23]. The PDMA scheme combines code, power, and spatial domains to create non-orthogonal patterns that are designed to maximize the diversity and minimize the collisions of multiple users in the same communication resource [19]. The PDMA receiver can use message passing algorithm (MPA) [24] or successive interference cancellation (SIC) algorithms to remove the interference and harvest the diversity [19]. IDMA uses a concept of chips interleaving. This procedure is equivalent to spreading the information across the available time-frequency resources, leading to a better BER performance even in an overloaded scenario [20]. MUSA can be viewed as an overloaded code division multiple access (CDMA) scheme. MUSA provides a grant-free multiple access for massive connections of low-complexity devices [25]. The MUSA multiplexer uses low-correlation spreading sequences, while the MUSA receiver employs a SIC algorithm to mitigate the overloading multi-user interference [21]. The LPMA scheme multiplexes the data from different users in the same resources by assigning different lattice codes [22]. Each user can mitigate the multi-user interference using a modulo lattice operation with hybrid parallel/SIC algorithm [22].

Among the relevant NOMA schemes, SCMA deserves special attention because it outperforms the aforementioned schemes in terms of BER performance [18, 26]. SCMA is based on different sparse codewords that are part of different sets, named codebooks. Each codebook represents an access layer, which is assigned for a user in the radio interface. The code sparsity reduces the collision among the layers for a given time-frequency resource. This feature allows the design of a near-optimal MPA detector with acceptable complexity [26]. Codebooks can be designed using multi-dimensional constellations, leading to a shaping gain [27]. This gain is responsible for the excellent SCMA performance in terms of BER [28].

The SCMA features and performance are attracting the attention of the scientific community. Several contributions have been presented in the last years. In [29], the authors enable SCMA to be used in downlink wireless access. In [27, 28], the authors deal with the SCMA codebook design. The performance comparison among different NOMA techniques is presented in [16, 18], where SCMA is highlighted by its good performance. In [30],

the authors introduce a low-complexity SCMA receiver based on MPA. In [31], the authors propose a new hybrid automatic repeat request (HARQ) combined with SCMA. Finally, in [32], the authors present a real-time SCMA transceiver prototype that can provide up to 300% overloading in field tests. All these contributions show that SCMA is a strong candidate for modern multi-user wireless systems.

In all abovementioned references, it is assumed that the physical layer has orthogonal time-frequency resources, such as OFDM. Recently, [33] combined SCMA with GFDM and analyzed the SER performance of the resulting scheme under different channel models, showing the advantages of combining NOMA with non-orthogonal waveforms.

The main aim of this paper is to propose two new receiver approaches to further improve the SCMA-GFDM SER performance under time-varying and frequency-selective channels. The first approach employs BEP channel equalizer [34–36] with a SPA MUD. This approach increments the system SER performance, but the receiver’s complexity grows significantly. In order to reduce the complexity, this paper introduces a second approach where the BEP equalizer is combined with a simplified MUD (SMUD). The SMUD severely reduces the receiver’s complexity at the cost of a small degradation in the SER performance. The results presented in this paper show that SCMA-GFDM presents high spectrum efficiency and good SER performance without significantly increasing the overall complexity of the receiver.

The remainder of this paper is organized as follows. Section 2 brings the main concepts of the SCMA-OFDM scheme. Section 3 presents the GFDM basic principles and its integration with SCMA and introduces the two receiver approaches that are the main contributions of this paper. Section 4 shows the SER performance analysis of the proposed approaches. Section 5 brings the final conclusions of this paper.

2 Background on SCMA-OFDM

OFDM [37] is a well-known waveform, extensively used in wireless system due to its low complexity and robustness against frequency-selective channels. On conventional OFDM, each OFDM symbol carries K quadrature amplitude modulation (QAM) symbols using K subcarriers, leading to a unitary overload factor. By combining OFDM and SCMA, the overall system capacity can be significantly improved.

An SCMA encoder can be defined as a mapper where $Q = \log_2(J)$ bits, where J is the size of a J -QAM constellation [23], are represented by a predefined U -dimensional complex codeword. Each U -dimensional complex codeword of the codebook set is a sparse vector with $Q < U$ non-zeroed entries. One SCMA layer is composed of a

codebook employed by a user to encode its data to be allocated in a set of U OFDM subcarriers. Figure 1 shows one SCMA layer for $Q = 2$ bits, QAM with $J = 4$ and $U = 4$. The codebook is a $U \times J$ matrix where each column is a codeword to represent one sequence of Q bits. This procedure is equivalent to spread the QAM into U subcarriers, i.e., the information that would be transmitted over a single subcarrier now is spread into U different subcarriers.

The SCMA scheme contains $C > U$ different layers [26], each one with a specific codebook designed according to [29]. In [23], the authors present a procedure to define the maximum overloading factor in an SCMA system. The number of SCMA layers is given by:

$$C = \binom{U}{Q}. \tag{1}$$

Because of the sparsity of the code, the number of non-zeroed elements of a given codeword that can collide with codewords from different codebooks is given by:

$$g = \binom{U-1}{Q-1} = \frac{CQ}{U}. \tag{2}$$

Therefore, overloading factor is given by:

$$\lambda = \frac{C}{U} = \frac{g}{Q}. \tag{3}$$

Figure 2 shows a multi-user SCMA encoder for the downlink channel in mobile networks, assuming $Q = 2$ and $U = 4$. Here, there are $C = 6$ different codebooks, each one employed to send data to a specific user, leading to an overloading factor of 1.5. All the codebooks used in this paper are presented in the Appendix. The bits to be sent to the users are mapped into a sequence x_{uc} , which are added together, leading to:

$$s_u = \sum_{c \in \xi(u)} x_{uc}, \tag{4}$$

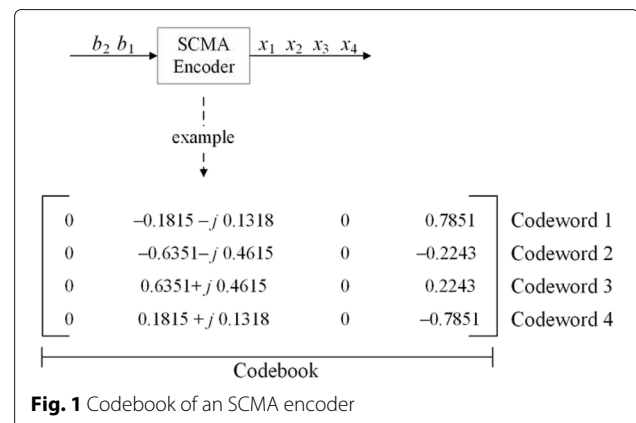
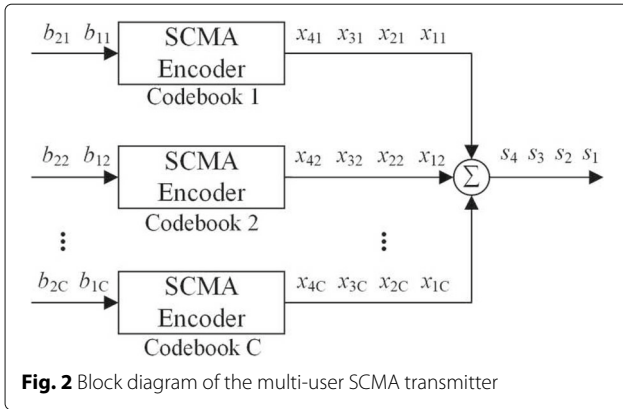


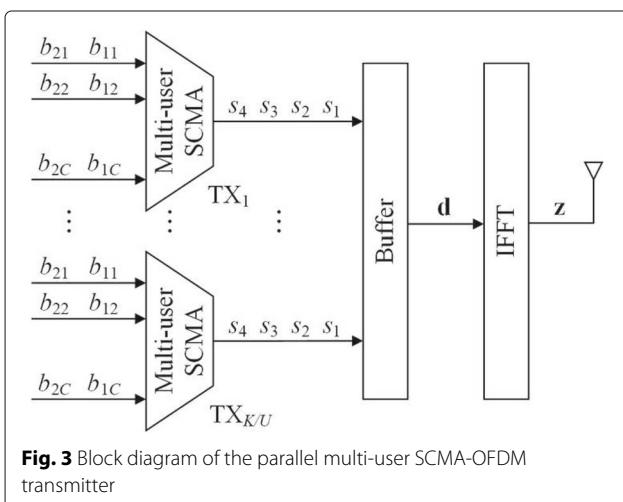
Fig. 1 Codebook of an SCMA encoder



that is the transmit sequence to be mapped in U OFDM subcarriers. Notice that $\xi(u)$ is a set formed by all layers that collide in the u th physical layer time-frequency resource. It is worth mentioning that the set of resources that are used by the same layer is denoted by $\zeta(c)$.

Wireless communication system that employs OFDM as the air interface usually adopts a high number of subcarriers, which means that $K \gg U$. In this paper, it is considered that K/U parallel SCMA encoders will be used, each one with C layers. All SCMA encoders use the same codebooks, and the orthogonality among the different SCMA transmit sequences is provided by the OFDM structure, where no intersymbol interference (ISI) is introduced if the cyclic prefix is larger than the channel delay profile and intercarrier interference (ICI) is avoided by the orthogonality among the subcarriers. Figure 3 depicts the block diagram of this SCMA-OFDM system.

The i th SCMA sequence, $\mathbf{s}_i = [s_1^i, s_2^i, s_3^i, s_4^i]^T$, will use U subcarriers of the OFDM symbol. All sequences are stacked into a $K \times 1$ vector $\mathbf{d} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_{K/U}]^T$. This vector is applied to an inverse fast Fourier transform (iFFT) algorithm, resulting in the transmit SCMA-OFDM



symbol, given by:

$$\mathbf{z} = \mathbf{F}_K^H \mathbf{d}, \quad (5)$$

where \mathbf{F}_K is a $K \times K$ Fourier matrix and $(\cdot)^H$ is the Hermitian operator. The overall payload of this system is given by:

$$\Lambda = K \frac{C}{U} \log_2(J) = K\lambda Q. \quad (6)$$

A CP is added to the OFDM symbol on the transmit side and removed on the receive side. Assuming that the CP is larger than the channel delay profile, the received signal after the CP removal is given by:

$$\mathbf{y} = \mathbf{H}\mathbf{z} + \mathbf{n} = \mathbf{H}\mathbf{F}_K^H \mathbf{d} + \mathbf{n}, \quad (7)$$

where \mathbf{n} is the additive white Gaussian noise (AWGN) vector with variance σ^2 and \mathbf{H} is a circulant channel matrix, based on the channel impulse, response given by:

$$\mathbf{H} = \begin{bmatrix} h_0 & 0 & 0 & \dots & h_2 & h_1 \\ h_1 & h_0 & 0 & \dots & h_3 & h_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ h_{L_{\text{ch}}-1} & h_{L_{\text{ch}}-2} & h_{L_{\text{ch}}-3} & \dots & 0 & 0 \\ 0 & h_{L_{\text{ch}}-1} & h_{L_{\text{ch}}-2} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & h_1 & h_0 \end{bmatrix}. \quad (8)$$

The OFDM subcarriers of the received signal can be decoupled by the fast Fourier transform (FFT), leading to:

$$\mathbf{Y} = \mathbf{F}_K \mathbf{y} = \mathbf{F}_K \mathbf{H} \mathbf{F}_K^H \mathbf{d} + \mathbf{F}_K \mathbf{n} = \mathcal{H} \mathbf{d} + \tilde{\mathbf{n}}, \quad (9)$$

where $\mathcal{H} = \mathbf{F}_K \mathbf{H} \mathbf{F}_K^H$ is a diagonal matrix containing the channel frequency response and $\tilde{\mathbf{n}}$ is the AWGN vector in the frequency domain. This means that each sample of the received signal in the frequency domain is given by:

$$Y_k = \mathcal{H}_{k,k} d_k + \tilde{n}_k. \quad (10)$$

The result presented in (10) shows that the SCMA decoder only needs to deal with the multi-user interference. Each set of U subcarriers must be applied to an SCMA MUD detector implemented using SPA. In this paper, the SPA has been chosen as the MUD algorithm to decouple the data from all SCMA layers because it is the most employed technique in the literature [23, 26–29, 31, 32]. The main reason for the success of the SPA as an SCMA MUD is its performance, which approaches the one achieved by the near-optimal detector [24].

It is worth to mention that SPA is a MPA-based decoder designed using a factor graph representation [18, 24]. The factor graph is a bi-partitioned graph containing layer nodes and resource nodes. The interaction among the layer and resource nodes is represented by a connection among them [24]. Figure 4 shows the SCMA-OFDM factor graph. Notice that each node Y_k has only one direct connection with one d_k , showing the orthogonality among

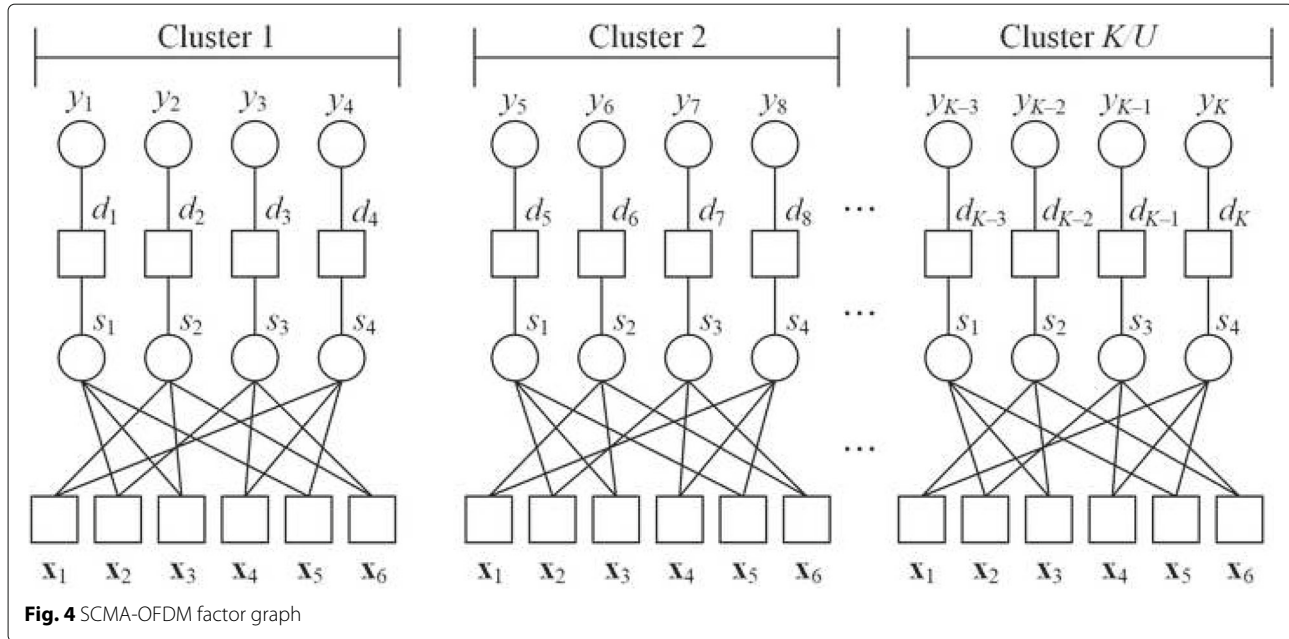


Fig. 4 SCMA-OFDM factor graph

the OFDM subcarriers. The vector \mathbf{d} is formed by the concatenation of the vectors s_u . Therefore, each cluster of U samples from d_k has a direct connection with the corresponding samples of the vector s_u . The relationship between s_u and x_{uc} is represented by the SCMA factor graph [23], as depicted in Fig. 4, where the resource nodes of a specific cluster represent the k th complex sample of the received signal s_u and the layer nodes represent the received codeword \mathbf{x}_c of the c th user.

Once the SCMA factor graph is known and assuming that the channel state information (CSI) is available at the receive side, the iterative SPA can be employed to extract the data from each user. For each cluster, the message passed from a resource node s_u to a layer node \mathbf{x}_c at the t th iteration is given by:

$$\begin{aligned}
 & m_{s_u \rightarrow \mathbf{x}_c}^{(t)}(\mathbf{x}_c) \\
 &= \sum_{\mathbf{x}_i | i \in \xi(u) \setminus c} \left\{ \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2} \|y_u - \mathcal{H}_{u,u} s_u\|^2\right) \right\} \\
 & \quad \prod_{i \in \xi(u) \setminus c} m_{\mathbf{x}_i \rightarrow s_u}^{(t-1)}(\mathbf{x}_i),
 \end{aligned} \tag{11}$$

where $\xi(u) \setminus c$ represents the set of all possible elements in $\xi(u)$ except c . The message passed from layer node \mathbf{x}_c to the resource node s_u in the t th iteration is given by:

$$m_{\mathbf{x}_c \rightarrow s_u}^{(t)}(\mathbf{x}_c) = \prod_{i \in \zeta(c) \setminus u} m_{s_i \rightarrow \mathbf{x}_c}^{(t-1)}(\mathbf{x}_c), \tag{12}$$

where $\zeta(c) \setminus u$ represents the set of all possible elements in $\zeta(c)$ except u . After τ iterations, the marginal probability distribution of each layer node is given by:

$$p(\mathbf{x}_c) = \prod_{u \in \zeta(c)} m_{s_u \rightarrow \mathbf{x}_c}^{(\tau)}(\mathbf{x}_c). \tag{13}$$

If the hard decision approach is employed, the codeword with highest probability of being transmitted \mathbf{x}_c is used to define the received J -QAM symbol for the c th user of each cluster.

According to [23, 38, 39], the SPA complexity is mainly defined by the size of the constellation alphabet in a given resource node and by the number of connections linking this resource node to the layer nodes. In other words, SPA complexity can be defined as $\mathcal{O}(D^\theta)$, where $D = J^g$ is the size of the constellation alphabet in a given resource node and θ is the number of connections at the resource node. Therefore, the complexity associated with each resource node in the SCMA-OFDM scheme is $\mathcal{O}(J^g)$. Considering all U resource nodes in a cluster, the number of iterations τ employed at the SPA, and K/U clusters, the overall complexity of the SPA-based SCMA-OFDM receiver is $\mathcal{O}(\tau K J^g)$.

3 Proposed SCMA-GFDM Integration

GFDM is a modern waveform designed to overcome several limitations presented by OFDM. GFDM employs circular baseband filtering for each one of the K subcarriers, where each one carries M QAM symbol. This means that GFDM has a time-frequency frame structure composed by K subcarriers and M subsymbols that carries $N = KM$ QAM data symbols. GFDM uses one prototype filter that is circularly shifted in time and in frequency. This approach adds an extra degree of freedom to GFDM when compared with OFDM, which is limited to K subcarriers. GFDM uses one CP to protect the entire block, reducing

the overhead in the physical layer when compared with OFDM. The GFDM filtering reduces the OOB, allowing this waveform to be used in cognitive radio networks [40]. All versions of the prototype filter can be organized in the transmit matrix \mathbf{A} , allowing the GFDM symbol to be defined as [41]:

$$\mathbf{z} = \mathbf{A}\mathbf{d}, \quad (14)$$

where \mathbf{d} is the data vector containing $N = KM$ QAM symbols.

After the CP removal, the received GFDM signal is given by:

$$\mathbf{y} = \mathbf{H}\mathbf{z} + \mathbf{n} = \mathbf{H}\mathbf{A}\mathbf{d} + \mathbf{n} = \tilde{\mathbf{H}}\mathbf{d} + \mathbf{n}, \quad (15)$$

where $\tilde{\mathbf{H}} = \mathbf{H}\mathbf{A}$ is the effective channel matrix.

In a GFDM system, the prototype filter can be freely chosen, allowing the system to be non-orthogonal. In this case, ICI, ICI, or both might rise. The GFDM receiver must resolve the self-introduced interference to avoid SER performance loss. Linear equalizers, such as matched filter (MF), zero-forcing (ZF), and minimum mean square error (MMSE), can be used to recover the data symbols with controlled SER performance loss. The MF, ZF, and MMSE detectors are respectively defined as:

$$\begin{aligned} \mathbf{B}_{\text{MF}} &= \mathbf{A}^H \mathbf{H}^{-1} \\ \mathbf{B}_{\text{ZF}} &= \tilde{\mathbf{H}}^{-1} = \mathbf{A}^{-1} \mathbf{H}^{-1} \\ \mathbf{B}_{\text{MMSE}} &= \left(\tilde{\mathbf{H}}^H \tilde{\mathbf{H}} + \frac{\sigma^2}{E} \mathbf{I}_N \right)^{-1} \tilde{\mathbf{H}}^H, \end{aligned} \quad (16)$$

where E is the average energy of the GFDM signal. The received data vector can be estimated as:

$$\mathbf{y}' = \mathbf{B}\mathbf{z}, \quad (17)$$

where \mathbf{B} can be any linear detector presented in (16). MF maximizes the signal-to-noise ratio (SNR), but it is unable to resolve the self-introduced interference. ZF removes the interference but enhances the noise. MMSE is a trade-off between MF and ZF, acting as a MF at low SNR and as ZF at high SNR. For the MF and ZF, prior channel equalization is necessary, while MMSE performs the channel equalization and the subcarriers and subsymbols decoupling simultaneously. In this paper, only ZF and MMSE detectors will be considered as linear receivers.

SCMA can be seamlessly integrated with GFDM using linear receivers, since the ZF and MMSE can successfully decouple the transmitted data on the receiver side. The SCMA-GFDM can use the same structure depicted in Fig. 3, where N/U SCMA encoders are employed and the buffer concatenates the \mathbf{s}_i vectors to generate \mathbf{d} . Clearly, instead of the Fourier matrix, the modulation matrix \mathbf{A} must be used in this case. Figure 6 shows the basic block diagram of the SCMA-GFDM receiver.

The interference introduced by the GFDM can be exploited as diversity, since the information of one data symbol will be spread among the adjacent subcarriers and subsymbols [42]. Although linear equalizers can harvest just some of this diversity gain, non-linear algorithms, such as expectation propagation, can be used to improve the GFDM performance under doubly dispersive channels. The factor graph for GFDM combined with SCMA is presented in Fig. 5 for $U = 4$, $Q = 2$, and $J = 4$. The SPA is able to resolve the interference introduced by GFDM and also perform the MUD. Since the non-linear GFDM detector is able to decouple the subcarrier and subsymbols, the SPA can be applied to each cluster individually to receive the SCMA codewords.

Because the subcarriers and subsymbols of the GFDM signal interact with each other, the number of links connecting the y_n nodes to the d_n node, which the weights are defined by $\tilde{\mathbf{H}}$, is very large. Therefore, the complexity of the SCMA-GFDM receiver is high. The size of the alphabet \mathcal{S}_u , corresponding to all possible values for s_u , is D . For the SCMA encoder presented in Fig. 5, $D = J^g = 64$. The number of links connecting a resource node is $\theta = N = 256$. The complexity for each resource node is $\mathcal{O}(D^\theta) = \mathcal{O}(64^{256})$, meaning that a SPA-based receiver for SCMA-GFDM is not practical.

The complexity of the SCMA-GFDM receiver can be substantially reduced if the GFDM symbol is equalized by a detector presented in (16) and parallel SCMA decoders are used for each cluster afterwards. The GFDM detector will decouple the subcarriers while the SCMA decoder will deal with the multi-user interference, as depicted in Fig. 6. This is a suboptimal approach, because the linear GFDM detector is unable to harvest the diversity introduced by the interaction among the subcarriers and subsymbols, but the significant complexity reduction achieved with this approach makes the SCMA-GFDM scheme feasible. This procedure is equivalent to the one described in Section 2, but due to the GFDM equalizer, the effective channel gain in (11) is $\mathcal{H}_{u,u} = 1$. The complexity of the GFDM equalizer is $\mathcal{O}(N^3)$, while the complexity of the SCMA decoders for N/U clusters is $\mathcal{O}(\tau N J^g)$. The total complexity of the SCMA-GFDM is $\mathcal{O}(N^3 + \tau N J^g)$.

In [34], the authors propose a new equalizer called BEP. This equalizer is based on the iterative expectation propagation (EP) algorithm, which approximates the true posterior distribution into exponential family distributions using the minimization of the Kullback-Leibler divergence [43]. The equalizer proposed in [34] achieves a trade-off between complexity and BER performance, and it can be employed in future wireless communication systems [34–36]. Next, this paper introduces two new approaches that improve the SCMA-GFDM SER performance, employing the BEP equalizer. The first one uses the BEP equalizer the GFDM self-interference, while SPA decoders are

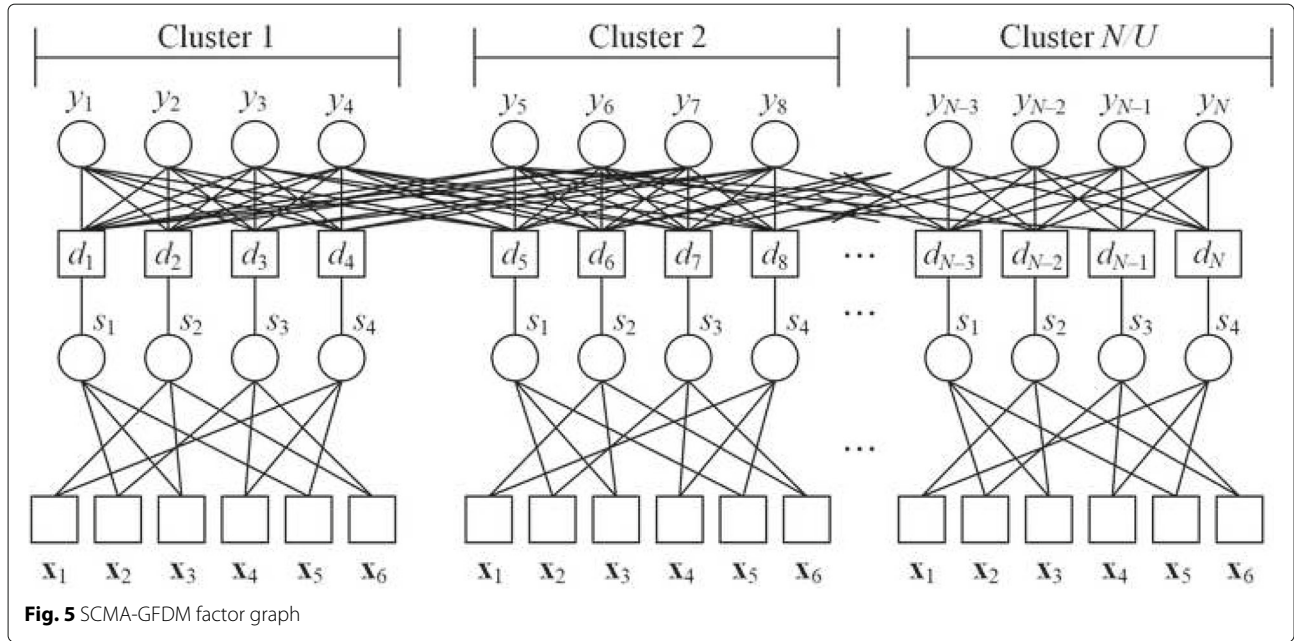


Fig. 5 SCMA-GFDM factor graph

employed to resolve the codeword collisions. This scheme improves the SCMA-GFDM SER performance at the cost of a high complexity. The second approach proposed in this paper uses the BEP equalizer with a SMUD that reduces the complexity of the SCMA decoder, at the cost of an affordable SER performance degradation. These new schemes are presented in the following sections.

3.1 SCMA-GFDM with BEP equalizer and SPA-based MUD

The scheme proposed for this approach is presented by Fig. 6. The BEP equalizer is used to detect the payload of the GFDM symbol while the SPA acts as the SCMA decoder. In this paper, the BEP equalizer proposed in [36] has been modified to recover the data transmitted by the SCMA-GFDM system. In [36], the posteriori probability of a given transmitted symbol vector, $p(\mathbf{d}|\mathbf{y}, \tilde{\mathbf{H}})$, was approximated by a non-normalized Gaussian $q(\mathbf{d})$ as:

$$\begin{aligned}
 p(\mathbf{d}|\mathbf{y}, \tilde{\mathbf{H}}) &= \frac{p(\mathbf{y}, \tilde{\mathbf{H}}|\mathbf{d})p(\mathbf{d})}{p(\mathbf{y}, \tilde{\mathbf{H}})} \approx q(\mathbf{d}) \\
 &= \mathcal{CN}(\mathbf{y}; \tilde{\mathbf{H}}\mathbf{d}, \sigma^2\mathbf{I}) \\
 &\prod_{n=1}^N \exp(d_n^H \gamma_n + \gamma_n^H d_n - \Gamma_n |d_n|^2),
 \end{aligned} \tag{18}$$

where $\gamma_n \in \mathbb{C}$ and $\Gamma_n \in \mathbb{R}_+$ are scalar values that are iteratively updated to approximate the non-normalized Gaussian $q(\mathbf{d})$ to the posteriori $p(\mathbf{d}|\mathbf{y}, \tilde{\mathbf{H}})$. The whole procedure of the iterative BEP equalizer is described in Algorithm 1. The iterative equalization process is repeated

ϕ times. According to [34], when ϕ increases, the BEP equalizer performance approaches the optimum performance of the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [44], but with a tractable complexity.

The whole BEP procedure can be divided in three stages. The first stage, indicated by line 3 in Algorithm 1, estimates the linear minimum mean square error (LMMSE) parameters, the mean vector $\boldsymbol{\mu}$, and the covariance matrix $\boldsymbol{\Sigma}$. The second stage, described by lines 4 to 14 in Algorithm 1, refines the estimates obtained in the first stage. The final stage, represented by lines 15 and 16, uses the outcome from the second stage to generate the complex Gaussian distributions that correspond to $q(\mathbf{d}) \propto \mathcal{CN}(\mathbf{d}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$. The mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$, obtained in line 3, are used to compute the n th marginal of distribution $q(\mathbf{d})$ given by $\mathcal{CN}(d_n; \mu_n^{(t)}, \Sigma_{n,n}^{(t)})$. In the next stage, the marginal cavity distribution $q_n^{(t)}(d_n)$ is found. This distribution is the baseline to obtain the approximations for the

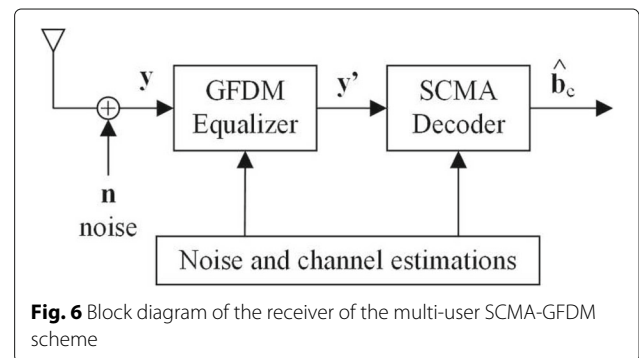


Fig. 6 Block diagram of the receiver of the multi-user SCMA-GFDM scheme

probability density function (PDF) of the received samples. The marginal cavity distribution is obtained by dividing $q(\mathbf{d})$ by the n th local likelihood approximation, given by $\exp\left(d_n^H \gamma_n^{(t)} + \gamma_n^{(t)} d_n - \Gamma_n^{(t)} |d_n|^2\right)$. The marginal cavity is a complex Gaussian distribution $\mathcal{CN}\left(d_n : b_n^{(t)}, c_n^{(t)}\right)$ with the mean $b_n^{(t)}$ and variance $c_n^{(t)}$ evaluated in line 5. This means that the marginal cavity distribution can be obtained by evaluating line 5 and plugging the mean values and covariances in the complex Gaussian distribution.

In the next step, the distribution $p^{(t)}(d_n)$ is evaluated in line 6. This distribution will be used to perform the moment matching described in the EP technique [43]. Line 6 brings an indicator function $\mathbb{I}_{d_n \in \mathcal{S}_n}$, which returns the value one if $d_n \in \mathcal{S}_u$ and zero otherwise. As already mentioned, \mathcal{S}_u is the alphabet of all possible values for s_u , defined by (4). Assuming $U = 4$, there are four different alphabets, namely $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$, and \mathcal{S}_4 . Figure 7 depicts these alphabets, where one can see that the alphabets \mathcal{S}_1 and \mathcal{S}_4 are identical, while the alphabets \mathcal{S}_2 and \mathcal{S}_3 are the same, as well. After computing $p^{(t)}(d_n)$, line 10 performs the moment matching.

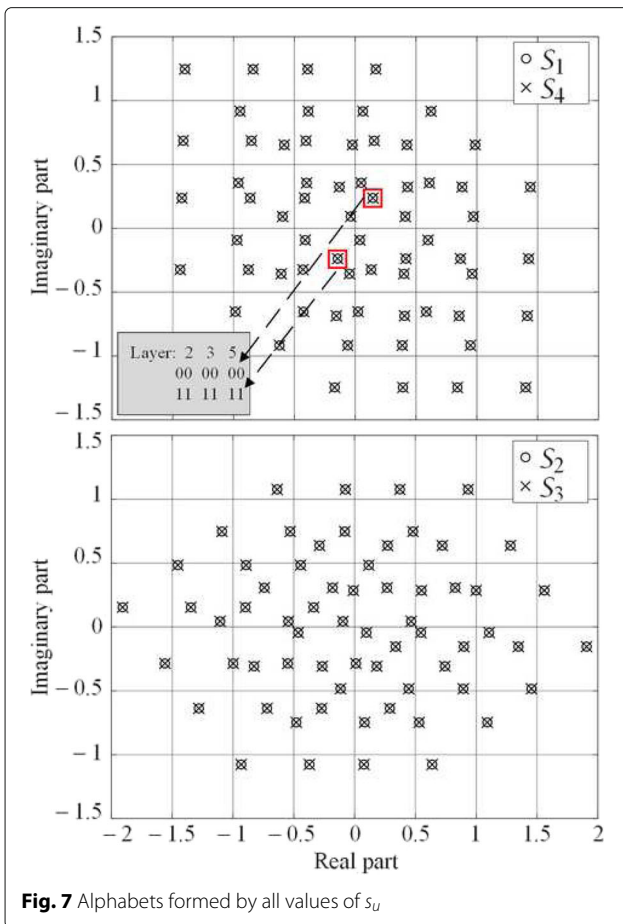


Fig. 7 Alphabets formed by all values of s_u

It is important to notice that there are two parameters in Algorithm 1 that control the numerical instabilities and the convergence rate. The parameters $\beta \in [0, 1]$ and $\epsilon \rightarrow 0$ are real-valued numbers that are empirically chosen, as described in [36]. The parameter ϵ is used in line 8 of Algorithm 1, where it can be seen as a minimum allowed variance for f_n . The parameter β is used in line 10, and it can be seen as the updating smoothing factor. If $\beta = 0$, there is no change in the parameters γ_n and Γ_n from one iteration to the next. If $\beta = 1$, the maximum updating rate is achieved for the parameters γ_n and Γ_n . If the iterative updating process leads to a negative value of Γ_n , line 11 keeps the previous value for the next iteration, i.e., the parameter Γ_n is not updated. Although these steps do not guarantee the convergence of the algorithm, they increase the probability of convergence to a level that allows it to be used in practical systems [36].

Regarding the complexity, for each iteration of Algorithm 1, there is a $N \times N$ matrix inversion in line 3, which dominates the stage 1 complexity. In the second stage, all N values of γ_n and Γ_n are simultaneously updated, leading to a complexity linearly proportional to NJ [34]. Hence, the overall complexity of the BEP equalizer is $\mathcal{O}(\phi(N^3 + NJ))$.

The vector delivered by the BEP equalizer, \mathbf{y}' , contains the estimates for $\hat{\mathbf{d}}$. This vector is used by the SPA algorithm to recover the information from the layers of each cluster. All SPA used as MUD can run in parallel. The complexity of the SCMA-GFDM with SPA-based MUD is $\mathcal{O}(\phi(N^3 + NJ) + \tau KJ^g)$.

The two iterative algorithms used to recover the users' information from the SCMA-GFDM signal result in a high complexity and high latency, but achieving good SER performance. Reducing the overall SCMA-GFDM complexity with manageable SER performance loss is the main goal of the approach presented in the next section.

3.2 SCMA-GFDM with BEP and SMUD

BEP is a soft or probabilistic equalizer that provides the posteriori probabilities, $q(\mathbf{d})$, of the received symbols. This additional information allows the implementation of the SMUD scheme proposed in this paper. Figure 6 shows the block diagram of the proposed SCMA-GFDM using SMUD. BEP is still used as the GFDM equalizer, providing the received samples of each cluster without the self-interference. These estimates and the posteriori distribution are used to recover the data sent at each SCMA layer.

SMUD relies on a look-up table (LUT) to recover the users' data, which is built according to the SCMA codebooks. Each value transmitted in a given time-frequency resource is a combination of non-zeroed values from the different codebooks, as defined in (4). One example is shown at the top of Fig. 8, where s_1 is a combination of

Algorithm 1 BEP Equalizer. The notation $(\cdot)^{(t)}$ indicates the value of a given variable (\cdot) at the t th iteration.

Require: Define the number of iterations ϕ

Require: Set $\beta \in [0, 1]$

Require: Set $\epsilon \leq \beta$

Require: Initialize $\gamma_n^{(0)} = 0$ and $\Gamma_n^{(0)} = 1, \forall n$

- 1: **for** $t = 0, \dots, \phi - 1$ **do**
- 2: **for** $n = 1, \dots, N$ **do**
- 3: Compute the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$ as follow:

$$\boldsymbol{\Sigma}^{(t)} = \left(\sigma^{-2} \tilde{\mathbf{H}}^H \tilde{\mathbf{H}} + \text{diag} \left(\Gamma^{(t)} \right) \right)^{-1}$$

$$\boldsymbol{\mu}^{(t)} = \boldsymbol{\Sigma} \left(\sigma^{-2} \tilde{\mathbf{H}}^H \mathbf{y} + \boldsymbol{\gamma}^{(t)} \right)$$

- 4: Compute the cavity marginal $q_{\setminus n}^{(t)}(d_n)$, defined as:

$$q_{\setminus n}^{(t)}(d_n) = \frac{\mathcal{CN}(d_n : \boldsymbol{\mu}_n^{(t)}, \boldsymbol{\Sigma}_{n,n}^{(t)})}{\exp \left(d_n^H \boldsymbol{\gamma}_n^{(t)} + \boldsymbol{\gamma}_n^H d_n - \Gamma_n^{(t)} |d_n|^2 \right)}$$

$$\approx \mathcal{CN} \left(d_n : b_n^{(t)}, c_n^{(t)} \right)$$

- 5: where $c_n^{(t)}$ and $b_n^{(t)}$ are, respectively, the variance and the mean of the cavity marginal $q_{\setminus n}^{(t)}(d_n)$ and are given by

$$c_n^{(t)} = \frac{\Sigma_{n,n}^{(t)}}{1 - \Sigma_{n,n}^{(t)} \Gamma_n^{(t)}} ; b_n^{(t)} = c_n^{(t)} \left(\frac{\mu_n^{(t)}}{\Sigma_{n,n}^{(t)}} - \gamma_n^{(t)} \right)$$

- 6: Obtain the distribution defined as

$$p^{(t)}(d_n) \propto q_{\setminus n}^{(t)}(d_n) \mathbb{I}_{d_n \in \mathcal{S}_u} \quad \forall u.$$

- 7: Estimate its mean $e_n^{(t)}$ and variance $f_n^{(t)}$.

- 8: **if** $f_n^{(t)} < \epsilon$ **then** $f_n^{(t)} = \epsilon$

- 9: **end if**

- 10: Update the new values of $\Gamma_n^{(t+1)}$ and $\gamma_n^{(t+1)}$ as follow

$$\Gamma_n^{(t+1)} = \beta \left(\frac{1}{f_n^{(t)}} - \frac{1}{c_n^{(t)}} \right) + (1 - \beta) \Gamma_n^{(t)}$$

$$\gamma_n^{(t+1)} = \beta \left(\frac{e_n^{(t)}}{f_n^{(t)}} - \frac{b_n^{(t)}}{c_n^{(t)}} \right) + (1 - \beta) \gamma_n^{(t)}$$

- 11: **if** $\Gamma_n^{(t+1)} < 0$ **then** $\Gamma_n^{(t+1)} = \Gamma_n^{(t)}$

- 12: **end if**

- 13: **end for**

- 14: **end for**

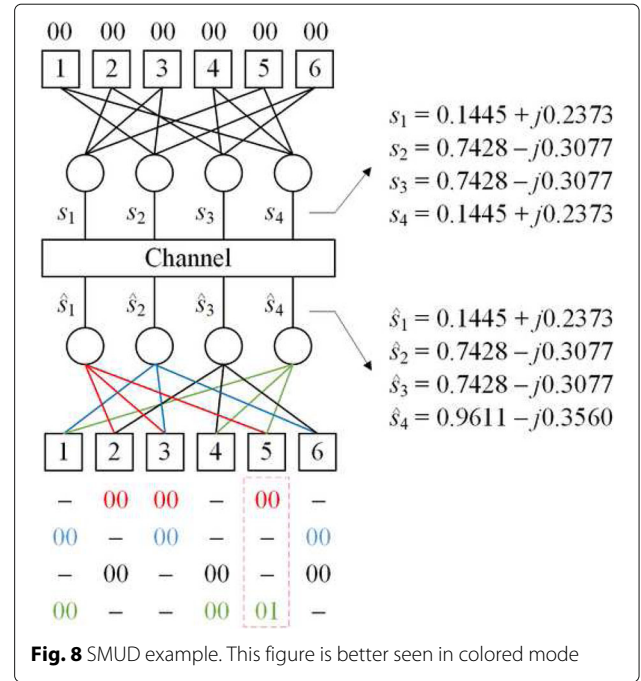
- 15: Compute the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$ of the Gaussian approximation $q(\mathbf{d}) \propto \mathcal{CN}(\mathbf{d} : \boldsymbol{\mu}, \boldsymbol{\Sigma})$ as follow

$$\boldsymbol{\Sigma} = \left(\sigma^{-2} \tilde{\mathbf{H}}^H \tilde{\mathbf{H}} + \text{diag} \left(\Gamma^{(\phi)} \right) \right)^{-1}$$

$$\boldsymbol{\mu} = \boldsymbol{\Sigma} \left(\sigma^{-2} \tilde{\mathbf{H}}^H \mathbf{y} + \boldsymbol{\gamma}^{(\phi)} \right)$$

- 16: Finally, the posteriori symbol probability are given by:

$$q(d_n = \mathcal{S}_u) \propto \mathcal{CN}(\mathcal{S}_u : \boldsymbol{\mu}_n, \boldsymbol{\Sigma}_{n,n}) \quad \forall u.$$



the values coming from layers 2, 3, and 5. These values are defined by the 4-QAM symbols in accordance with the codebook for each layer. The Appendix brings the set of codebooks used in this paper. For each combination of 4-QAM symbols, there will be a unique output in s_1 , and therefore, it is possible to build a LUT according to the alphabet \mathcal{S}_u shown in Fig. 7. As example, assuming that the sequence “00” is transmitted in layers 2, 3, and 5 in a given moment. Therefore, the correspondent outputs of each codebook will be 0.7851, $-0.6351 + j0.4615$, and $-0.0055 - j0.2242$ as can be seen in the Appendix. Hence, in this case, $s_1 = 0.1445 + j0.2373$. If the sequence “11” is transmitted in these layers, the correspondent outputs of each codebook will be $-0.7851, 0.6351 - j0.4615$, and $0.0055 + j0.2242$, leading to the output $s_1 = -0.1445 - j0.2373$. These two examples of s_1 are highlighted in Fig. 7. This procedure can be used to map all possible sequences transmitted in layers 2, 3, and 5 to all possible values that s_1 can assume. Table 3 in the Appendix shows the initial entries for the LUT developed for s_1 . Each time-frequency resource of a given cluster, s_u , will have a specific LUT.

Figure 8 shows an example where it is possible to visualize how the LUT can be used in this context. Consider that all users of a given cluster transmit the sequence “00”. The resulting transmitted values in the U time-frequency resources are as follows: $s_1 = s_4 = 0.1445 + j0.2373$ and $s_2 = s_3 = 0.7428 - j0.3077$, as highlighted in Fig. 8. Assuming, for now, a noiseless channel and a perfect GFDM equalization, then $\hat{s}_u = s_u$. In this case, according to the LUT, the \hat{s}_1 value will indicate that the sequence “00” was transmitted in layers 2, 3, and 5, and the \hat{s}_2

value will indicate that the sequence “00” was transmitted in layers 1, 3, and 6. The same is true for \hat{s}_3 and \hat{s}_4 , for the corresponding layers. In this example, the content received in layer 3 can be estimated from \hat{s}_1 or \hat{s}_3 . However, since both lead to the same sequence “00”, the information can be retrieved unequivocally. When a perfect equalization is performed, all duplicated sequences in each layer converge to the same value, leading to the correct decision.

Assume now that the distortions introduced by noise and channel equalization lead to an error in $\hat{s}_4 \neq s_4$. For example, consider that $\hat{s}_4 = 0.9611 - j0.3560$. Based on the \hat{s}_1 value, the LUT will indicate that the received sequence in layers 2, 3, and 5 is “00.” The same is true for the \hat{s}_2 and \hat{s}_3 values, for the corresponding layers. However, base on \hat{s}_4 , the LUT will indicate that the sequence “00” is received in layers 1 and 4, while the sequence “01” is received in layer 5, once this is the combination that results in $\hat{s}_4 = 0.9611 - j0.3560$ at the output of the SCMA encoder. There are two different sequences recovered for layer 5, which are “00” obtained from \hat{s}_1 and “01” obtained from \hat{s}_4 . The posteriori distribution $q(\mathbf{d})$ provided by the BEP soft equalizer can be used to define which sequence was transmitted with higher probability. In other words, the SMUD will use the sequence provided by the most reliable signal \hat{s}_u . In the example presented above, if $\max[q(d_1)] > \max[q(d_4)]$, the SMUD will decide for the sequence obtained from \hat{s}_1 . Otherwise, the SMUD will favor the sequence obtained from \hat{s}_4 . It is interesting to observe that the SMUD is able to correct the error introduced by the communication channel, meaning that it can behave as a forward error control (FEC).

Regarding the complexity evaluation, the LUT can be applied in parallel for each cluster and its complexity is linear in UJ . The total complexity of the SCMA-GFDM with SMUD is $\mathcal{O}(\phi(N^3 + NJ) + UJ)$, which is much smaller than the SCMA-GFDM with SPA-based MUD. Table 1 compares the complexity of each scheme discussed in this paper.

4 SER performance analysis

The SER performance evaluation has been realized using Monte Carlo simulations. The SNR is defined as E_s/N_0 ,

Table 1 Complexity of the schemes presented in this paper

Waveform	Equalizer	MUD	Complexity
OFDM	SPA		$\mathcal{O}(\tau KJ^g)$
	ZF	SPA	$\mathcal{O}(N^3 + \tau NJ^g)$
GFDM	MMSE	SPA	$\mathcal{O}(N^3 + \tau NJ^g)$
	BEP	SPA	$\mathcal{O}(\phi(N^3 + NJ) + \tau NJ^g)$
	BEP	SMUD	$\mathcal{O}(\phi(N^3 + NJ) + UJ)$

where E_s is the average data symbol energy and N_0 is the noise power spectral density.

A slow time-variant and frequency-selective channel with Rayleigh distribution is assumed, as described in (8). The SCMA-OFDM will be used as a benchmark for the proposed SCMA-GFDM schemes. The first step for the SER performance evaluation of the SCMA-GFDM with BEP equalizer and SPA-based MUD is to define the values for ϕ and τ . These parameters define the trade-off between complexity and SER performance. If they are too low, the SER performance losses will be significant. If they are too large, the complexity will increase without bringing any benefits in terms of SER performance. Figure 9 compares the SER performance of the SCMA-GFDM with BEP equalizer and SPA-based MUD, assuming the parameter presented in Table 2, for different values of ϕ and τ . From Fig. 9, it is possible to conclude that the worst performance is achieved when no iteration is employed, i.e., $\phi = \tau = 1$. When iteration is used only for the BEP, i.e., $\tau = 1$ and $\phi = 10$, it is possible to conclude that the performance improves mainly for high SNR. On the other hand, when iteration is used only for the SCMA MUD, i.e., $\tau = 10$ and $\phi = 1$, the performance for low SNR improves. Iterations are needed in both algorithms for good SER performance over the entire SNR range. The best performance among the simulated curves is achieved when $\tau = \phi = 10$, but the gain is marginal compared with the curve for $\tau = \phi = 5$. This allows one to conclude that values of τ and ϕ larger than 10 will increase the complexity without substantial increment in the SER performance.

Figure 10 shows the SER performance of the schemes proposed in this paper and compares them with the well-known SCMA-OFDM, in which the parameters have

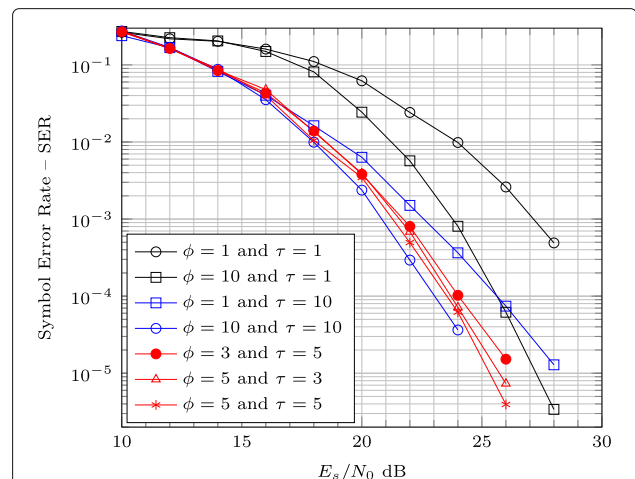


Fig. 9 SER performance for SCMA-GFDM with BEP equalizer and SPA multi-user detection for different values of ϕ and τ . This figure is better seen in colored mode

Table 2 Simulation parameters

Parameters	SCMA-OFDM	SCMA-GFDM
K	256	32
M	1	7
$N = MK$	256	224
Roll-off factor	0	0.5
C	6	6
Modulation (L)	QPSK (4)	QPSK (4)
U	4	4
β	0.5	0.5
ϵ	$1e^{-4}$	$1e^{-4}$

been defined in Table 2. From Fig. 10, it is possible to observe that SCMA-OFDM presents the best performance for low SNR, once this system does not suffer from self-interference and noise enhancement as the SCMA-GFDM. However, SCMA-OFDM also cannot benefit from the rich channel environment, being unable to exploit the diversity introduced by the interactions among the subsymbols and subcarriers. The ZF equalizer used for GFDM reception can harvest some diversity from the received signal, but this equalizer also introduces noise enhancement [45], resulting in a poor SER performance when compared with SCMA-OFDM. MMSE equalizer can deal with the GFDM subcarrier and subsymbol decoupling without enhancing the noise, leading to an acceptable performance for low SNR and outperforming SCMA-OFDM for high SNR.

The two schemes proposed in this paper can outperform all other analyzed schemes for high SNR because non-linear equalizers can better exploit the diversity introduced by GFDM. The SCMA-GFDM with BEP and

SPA-based MUD outperforms the SCMA-GFDM with MMSE equalizer for $\text{SNR} > 16$ dB for τ and ϕ larger than 5. It is also possible to observe that the performance is slightly improved when $\tau = \phi = 10$, when compared with the curve for $\tau = \phi = 5$. The price paid, in this case, is the high complexity of the receiver. The performance of the SCMA-GFDM with BEP and SMUD outperforms the SCMA-GFDM with MMSE for $\text{SNR} > 24$ dB, assuming $\phi = 10$, or $\text{SNR} > 25.5$ dB, assuming $\phi = 5$. The price for the considerable complexity reduction of this scheme is the SER poor performance at low SNR. If good performance at low SNR is mandatory and the complexity of the BEP equalizer with SPA-based MUD cannot be afforded, a hybrid solution using the MMSE equalizer combined with SPA MUD can be used for low SNR, while the BEP equalizer with SMUD can be employed for high SNR. Step 3 of the BEP Algorithm 1 can be used to evaluate the MMSE matrix at low SNR, simplifying the implementation of the hybrid scheme. In this case, the complexity between the two blocks of the receiver, equalizer and MUD, is balanced according to the SNR. For low SNR, the complexity of the equalizer is $\mathcal{O}(N^3 + NJ)$ and the complexity of the MUD is given by the SPA complexity, $\mathcal{O}(\tau NJ^g)$. As the SNR increases, the complexity of the equalizer also increases by a multiplicative factor ϕ , while the complexity of the MUD is drastically reduced to $\mathcal{O}(LJ)$. The complexity of the combined system is still lower than the SCMA-GFDM with BEP and SPA-based MUD, and its SER performance would be acceptable for the entire SNR range.

5 Conclusions

The demand for higher spectrum efficiency in future mobile communication systems is pushing the development of NOMA techniques. Among the several techniques proposed in the literature, SCMA is attracting the attention because of its SER performance over mobile communication channels. Also, GFDM is an interesting candidate for the next wireless communication standards because of its flexibility to address conflicting requirements in a scenario with multiple applications. This paper has proposed two SCMA-GFDM schemes with non-linear equalizers that are able to exploit the SCMA spectrum efficiency and the GFDM flexibility and good performance over rich multipath and time-variant channels. The first approach consists on using the BEP equalizer to recover the information from the GFDM signal, while a SPA-based algorithm is used as MUD to recover the user data at each SCMA layer. The SCMA-GFDM with BEP and SPA-based MUD presents acceptable SER performance for low SNR and outperforms all other analyzed schemes for high SNR. However, this scheme is highly complex. In order to overcome this problem, the paper also presents the SCMA-GFDM with BEP and SMUD, where instead of using the SPA as MUD, a LUT based

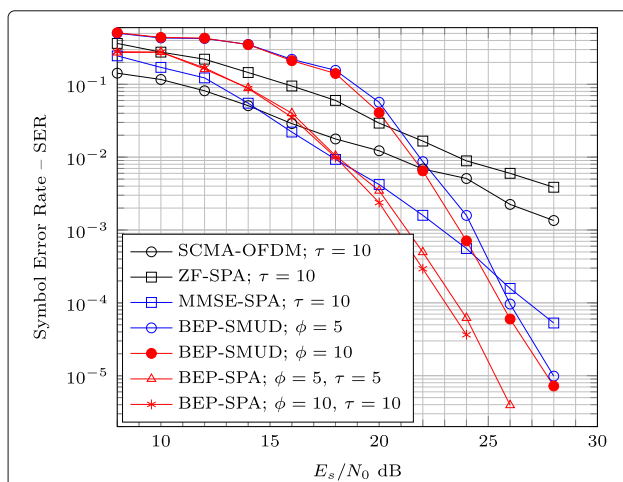


Fig. 10 SER performance for SCMA-OFDM and SCMA-GFDM. This figure is better seen in colored mode

Table 3 Examples of the LUT for alphabet \mathcal{S}_1

Layer 2	Layer 3	Layer 5	C_2	C_3	C_5	s_1
00	00	00	0.7851	$-0.6351 + j0.4615$	$-0.0055 - j0.2242$	$0.1445 + j0.2373$
00	00	01	0.7851	$-0.6351 + j0.4615$	$-0.0193 - j0.7848$	$0.1307 - j0.3233$
00	00	10	0.7851	$-0.6351 + j0.4615$	$0.0193 + j0.7848$	$0.1693 + j1.2463$
00	00	11	0.7851	$-0.6351 + j0.4615$	$0.0055 + j0.2242$	$0.1555 + j0.6857$
00	01	00	0.7851	$0.1815 - j0.1318$	$-0.0055 - j0.2242$	$0.9611 - 0.3560$
00	01	01	0.7851	$0.1815 - j0.1318$	$-0.0193 - j0.7848$	$0.9473 - j0.9166$
00	01	10	0.7851	$0.1815 - j0.1318$	$0.0193 + j0.7848$	$0.9859 + j0.6530$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
11	11	10	-0.7851	$0.6351 - j0.4615$	$0.0193 + j0.7848$	$-0.1307 + j0.3233$
11	11	11	-0.7851	$0.6351 - j0.4615$	$0.0055 + j0.2242$	$-0.1445 - 0.2373$

on the all possible results from the SCMA encoder is employed to recover the user data at each SCMA layer. The information received in each GFDM time-frequency resource will define the bit sequence for several layers, according to the SCMA encoder. Conflicting information about the received sequence in a given layer will be resolved using the BEP posteriori probability for each received sample of the GFDM signal. The sequence corresponding to the most reliable time-frequency resource will be considered at the SMUD output. This approach significantly reduces the overall system complexity at the cost of SER performance at low SNR, when compared with the SCMA-GFDM with BEP and SPA-based MUD. The proposed schemes show that it is possible to have the benefits of NOMA with the flexibility of a modern non-orthogonal waveform, opening new possibilities for the future wireless communication systems.

Appendix. SCMA codebooks and table of examples of the SMUD LUT

The following codebooks are used in this paper:

$$C_1 = \begin{bmatrix} 0 & -0.1815 - j0.1318 & 0 & 0.7851 \\ 0 & -0.6351 - j0.4615 & 0 & -0.2243 \\ 0 & 0.6351 + j0.4615 & 0 & 0.2243 \\ 0 & 0.1815 + j0.1318 & 0 & -0.7851 \end{bmatrix}$$

$$C_2 = \begin{bmatrix} 0.7851 & 0 & -0.1815 - j0.1318 & 0 \\ -0.2243 & 0 & -0.6351 - j0.4615 & 0 \\ 0.2243 & 0 & 0.6351 + j0.4615 & 0 \\ -0.7851 & 0 & 0.1815 + j0.1318 & 0 \end{bmatrix}$$

$$C_3 = \begin{bmatrix} -0.6351 + j0.4615 & 0.1392 - j0.1759 & 0 & 0 \\ 0.1815 - j0.1318 & 0.4873 - j0.6156 & 0 & 0 \\ -0.1815 + j0.1318 & -0.4873 + j0.6156 & 0 & 0 \\ 0.6351 - j0.4615 & -0.1392 + j0.1759 & 0 & 0 \end{bmatrix}$$

$$C_4 = \begin{bmatrix} 0 & 0 & 0.7851 & -0.0055 - j0.2242 \\ 0 & 0 & -0.2243 & -0.0193 - j0.7848 \\ 0 & 0 & 0.2243 & 0.0193 + j0.7848 \\ 0 & 0 & -0.7851 & 0.0055 + j0.2242 \end{bmatrix}$$

$$C_5 = \begin{bmatrix} -0.0055 - j0.2242 & 0 & 0 & -0.6351 + j0.4615 \\ -0.0193 - j0.7848 & 0 & 0 & 0.1815 - j0.1318 \\ 0.0193 + j0.7848 & 0 & 0 & -0.1815 + j0.1318 \\ 0.0055 + j0.2242 & 0 & 0 & 0.6351 - j0.4615 \end{bmatrix}$$

$$C_6 = \begin{bmatrix} 0 & 0.7851 & 0.1392 - j0.1759 & 0 \\ 0 & -0.2243 & 0.4873 - j0.6156 & 0 \\ 0 & 0.2243 & -0.4873 + j0.6156 & 0 \\ 0 & -0.7851 & -0.1392 + j0.1759 & 0 \end{bmatrix}$$

Table 3 brings some numerical examples of the LUT of the alphabet \mathcal{S}_1 . The most part of the LUT is omitted for the sake of brevity.

Abbreviations

3GPP: 3rd generation partnership project; 5G: fifth generation of mobile network; 6G: sixth generation of mobile networks; AWGN: additive white Gaussian noise; B5G: beyond 5G; BCJR: Bahl-Cocke-Jelinek-Raviv; BER: bit error rate; BEP: block expectation propagation; CDMA: code division multiple access; CP: cyclic prefix; CSI: channel state information; EP: expectation propagation; FBMC: filterbank multicarrier; FDMA: frequency division multiple access; FEC: forward error control; FFT: fast Fourier transform; GFDM: generalized frequency division multiplexing; HARQ: hybrid automatic repeat request; ICI: intercarrier interference; IDMA: interleaved division multiple access; iFFT: inverse fast Fourier transform; IoT: Internet of things; ISI: intersymbol interference; LMMSE: linear minimum mean square error; LPMA: lattice partition multiple access; LUT: look-up table; MF: matched filter; MMSE: minimum mean square error; MPA: message passing algorithm; MUD: multi user detection; MUSA: multi-user shared access; NR: new radio; NOMA: non-orthogonal multiple access; OFDM: orthogonal frequency division multiplexing; OMA: orthogonal multiple access; OOB: out-of-band emission; PDF: probability density function; PDMA: pattern division multiple access; QAM: quadrature amplitude modulation; RAN: radio access network; SC-FDE: single carrier frequency domain equalization; SCMA: sparse code multiple access; SER: symbol error rate; SIC: successive interference cancellation; SMUD: simplified multi user detection; SNR: signal-to-noise ratio; SPA: sum product algorithm; ZF: zero-forcing

Acknowledgements

Not applicable.

Authors' contributions

Both authors have contributed equally to the development of manuscript and of the experiments. The authors read and approved the final manuscript.

Authors' information

Guilherme Pedro Aquino received the M.Sc. degree in telecommunications engineering from Instituto Nacional de Telecomunicações (Inatel), Brazil, and

received the Doctor in Electrical Engineer degree from Universidade Federal de Itajubá, UNIFEI, Brazil. Since 2010, he is a professor and researcher at Inatel. His main research interests lie in the study of the non-orthogonal multiple access schemes and efficient cooperative spectrum sensing techniques for cognitive radio applications.

Luciano Leonel Mendes received the B.Sc. and M.Sc. in Electrical Engineer from Inatel, Brazil in 2001 and 2003, respectively. In 2007, he received the Doctor in Electrical Engineer degree from Unicamp, Brazil. Since 2001, he is a professor at Inatel, Brazil. He has coordinated the Master Program at Inatel and several research projects funded by FAPEMIG, FINEP, and BNDES. He has been a visiting researcher at Vodafone Chair Mobile Communications Systems at the TU Dresden between 2013 and 2015. His main area of research is waveforms for 5G networks and future mobile communication systems.

Funding

This work was partially supported by RNP, with resources from MCTIC, grant no. 01250.075413/2018-04, under the Radiocommunication Reference Center (Centro de Referência em Radiocomunicações - CRR) project of the National Institute of Telecommunications (Instituto Nacional de Telecomunicações - Inatel), Brazil, by the 5G-RANGE project and CNPq-Brasil under grant no. 305085/2018-2.

Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Competing interests

The authors declare that they have no competing interests.

Received: 19 February 2020 Accepted: 2 October 2020

Published online: 23 October 2020

References

1. G. Wunder, P. Jung, M. Kasparick, T. Wild, F. Schaich, Y. Chen, S. T. Brink, I. Gaspar, N. Michailow, A. Festag, L. Mendes, N. Cassiau, D. Ktenas, M. Dryjanski, S. Pietrzyk, B. Eged, P. Vago, F. Wiedmann, 5GNOW: non-orthogonal, asynchronous waveforms for future mobile applications. *IEEE Commun. Mag.* **52**(2), 97–105 (2014). <https://doi.org/10.1109/MCOM.2014.6736749>
2. F. Javed, M. K. Afzal, M. Sharif, B. Kim, Internet of Things (IoT) operating systems support, networking technologies, applications, and challenges: a comparative review. *IEEE Commun. Surv. Tutor.* **20**(3), 2062–2100 (2018). <https://doi.org/10.1109/COMST.2018.2817685>
3. F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, P. Popovski, Five disruptive technology directions for 5G. *IEEE Commun. Mag.* **52**(2), 74–80 (2014). <https://doi.org/10.1109/MCOM.2014.6736746>
4. A. Gupta, R. K. Jha, A survey of 5G network: architecture and emerging technologies. *IEEE Access.* **3**, 1206–1232 (2015). <https://doi.org/10.1109/ACCESS.2015.2461602>
5. A. Ghosh, A. Maeder, M. Baker, D. Chandramouli, 5G evolution: a view on 5G cellular technology beyond 3GPP release 15. *IEEE Access.* **7**, 127639–127651 (2019). <https://doi.org/10.1109/ACCESS.2019.2939938>
6. ETSI TS 123 501, 5G: system architecture for the 5G system. 3GPP TS 23.501 version 15.2.0 Release 15 (2018). https://www.etsi.org/deliver/etsi_ts/123500_123599/123501/15.02.00_60/ts_123501v150200p.pdf
7. D. Zhang, M. Matthé, L. L. Mendes, G. Fettweis, A study on the link level performance of advanced multicarrier waveforms under MIMO wireless communication channels. *IEEE Trans. Wirel. Commun.* **16**(4), 2350–2365 (2017). <https://doi.org/10.1109/TWC.2017.2664820>
8. H. Viswanathan, P. E. Mogensen, Communications in the 6G era. *IEEE Access.* **8**, 57063–57074 (2020). <https://doi.org/10.1109/ACCESS.2020.2981745>
9. P. Cheng, M. Tao, Y. Xiao, W. Zhang, V-OFDM: on performance limits over multi-path Rayleigh fading channels. *IEEE Trans. Commun.* **59**(7), 1878–1892 (2011). <https://doi.org/10.1109/TCOMM.2011.051711.100223>
10. P. Weikemper, J. Bazzi, K. Kusume, A. Benjebbour, Y. Kishiyama, On regular resource grid for filtered OFDM. *IEEE Commun. Lett.* **20**(12), 2486–2489 (2016). <https://doi.org/10.1109/LCOMM.2016.2572183>
11. L. Zhang, P. Xiao, A. Zafar, A. ul Qudus, R. Tafazolli, FBMC system: an insight into doubly dispersive channel impact. *IEEE Trans. Veh. Technol.* **66**(5), 3942–3956 (2017). <https://doi.org/10.1109/TVT.2016.2602096>
12. N. Michailow, G. Fettweis, in *Proceedings of the International Symposium on Intelligent Signal Processing and Communication Systems*. Low peak-to-average power ratio for next generation cellular systems with generalized frequency division multiplexing (IEEE, Naha, 2013), pp. 651–655. <https://doi.org/10.1109/ISAPCS.2013.6704629>
13. I. B. Franco de Almeida, L. Leonel Mendes, in *2018 IEEE 5G World Forum (5GWF)*. Linear GFDM: a low out-of-band emission configuration for 5G air interface, (2018), pp. 311–316. <https://doi.org/10.1109/5GWF.2018.8516993>
14. M. Matthé, D. Zhang, G. Fettweis, Low-complexity iterative MMSE-PIC detection for MIMO-GFDM. *IEEE Trans. Commun.* **66**(4), 1467–1480 (2018). <https://doi.org/10.1109/TCOMM.2017.2782339>
15. L. Zhu, Z. Xiao, X. Xia, D. Oliver Wu, Millimeter-wave communications with non-orthogonal multiple access for B5G/6G. *IEEE Access.* **7**, 116123–116132 (2019). <https://doi.org/10.1109/ACCESS.2019.2935169>
16. L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, Z. Wang, Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends. *IEEE Commun. Mag.* **53**(9), 74–81 (2015). <https://doi.org/10.1109/MCOM.2015.7263349>
17. M. Cheng, Y. Wu, Y. Chen, in *Proceedings of the International Conference on Wireless Communications Signal Processing*. Capacity analysis for non-orthogonal overloading transmissions under constellation constraints (IEEE, Nanjing, 2015), pp. 1–5. <https://doi.org/10.1109/WCSP.2015.7341294>
18. B. Wang, K. Wang, Z. Lu, T. Xie, J. Quan, in *Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*. Comparison study of non-orthogonal multiple access schemes for 5G (IEEE, Ghent, 2015), pp. 1–5. <https://doi.org/10.1109/BMSB.2015.7177186>
19. S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, K. Niu, Pattern division multiple access (PDMA) - a novel non-orthogonal multiple access for 5G radio networks. *IEEE Trans. Veh. Technol.* **66**(4), 3185–3196 (2016). <https://doi.org/10.1109/TVT.2016.2596438>
20. Y. Chen, F. Schaich, T. Wild, in *Proceedings of the IEEE 79th Vehicular Technology Conference*. Multiple access and waveforms for 5G: IDMA and universal filtered multi-carrier (IEEE, Seoul, 2014), pp. 1–5. <https://doi.org/10.1109/VTCSpring.2014.7022995>
21. Z. Yuan, G. Yu, W. Li, Y. Yuan, X. Wang, J. Xu, in *Proceedings of the IEEE Vehicular Technology Conference*. Multi-user shared access for Internet of things (IEEE, Nanjing, 2016), pp. 1–5. <https://doi.org/10.1109/VTCSpring.2016.7504361>
22. D. Fang, Y. Huang, Z. Ding, G. Geraci, S. Shieh, H. Claussen, in *2016 IEEE Global Communications Conference (GLOBECOM)*. Lattice partition multiple access: a new method of downlink non-orthogonal multiuser transmissions, (2016), pp. 1–6. <https://doi.org/10.1109/GLOCOM.2016.7841947>
23. H. Nikopour, H. Baligh, in *Proceedings of the IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications*. Sparse code multiple access (IEEE, London, 2013), pp. 332–336. <https://doi.org/10.1109/PIMRC.2013.6666156>
24. F. R. Kschischang, B. J. Frey, H. A. Loeliger, Factor graphs and the sum-product algorithm. *IEEE Trans. Inf. Theory.* **47**(2), 498–519 (2001). <https://doi.org/10.1109/18.910572>
25. Y. Chunlin, Y. Zhifeng, L. Weimin, Y. Yifei, Non-orthogonal multiple access schemes for 5G. *ZTE Communications.* **14**(4), 11–16 (2016)
26. Y. Han, S. Zhang, W. Zhou, Q. Ling, in *Proceedings of the International Conference on Wireless Communications Signal Processing*. Enabling SCMA long codewords with a parallel SCMA coding scheme (IEEE, Nanjing, 2015), pp. 1–6. <https://doi.org/10.1109/WCSP.2015.7341085>
27. M. Taherzadeh, H. Nikopour, A. Bayesteh, H. Baligh, in *Proceedings of the IEEE 80th Vehicular Technology Conference*. SCMA codebook design (IEEE, Vancouver, 2014), pp. 1–5. <https://doi.org/10.1109/VTCFall.2014.6966170>
28. J. Bao, Z. Ma, Z. Ding, G. K. Karagiannidis, Z. Zhu, On the design of multiuser codebooks for uplink SCMA systems. *IEEE Commun. Lett.* **20**(10), 1920–1923 (2016). <https://doi.org/10.1109/LCOMM.2016.2596759>
29. H. Nikopour, E. Yi, A. Bayesteh, K. Au, M. Hawryluck, H. Baligh, J. Ma, in *Proceedings of the IEEE Global Communications Conference*. SCMA for downlink multiple access of 5G wireless networks (IEEE, Austin, 2014), pp. 3940–3945. <https://doi.org/10.1109/GLOCOM.2014.7037423>
30. L. Yang, Y. Liu, Y. Siu, Low complexity message passing algorithm for SCMA system. *IEEE Commun. Lett.* **20**(12), 2466–2469 (2016). <https://doi.org/10.1109/LCOMM.2016.2609382>

31. Y. Long, Z. Chen, Z. Guo, J. Fang, A novel HARQ scheme for SCMA systems. *IEEE Wirel. Commun. Lett.* **5**(5), 452–455 (2016). <https://doi.org/10.1109/LWC.2016.2585110>
32. L. Lu, Y. Chen, W. Guo, H. Yang, Y. Wu, S. Xing, Prototype for 5G new air interface technology SCMA and performance evaluation. *China Commun.* **12**, 38–48 (2015). <https://doi.org/10.1109/CC.2015.7386169>
33. G. P. Aquino, L. Leonel Mendes, in *Proceedings of the IEEE 5G World Forum*. Sparse code multiple access applied in the generalized frequency division multiplexing (IEEE, Silicon Valley, 2018), pp. 49–54. <https://doi.org/10.1109/5GWF.2018.8517096>
34. I. Santos, J. J. Murillo-Fuentes, P. M. Olmos, in *Proceedings of the European Signal Processing Conference*. Block expectation propagation equalization for ISI channels (IEEE, Nice, 2015), pp. 379–383. <https://doi.org/10.1109/EUSIPCO.2015.7362409>
35. S. Wu, Z. Ni, X. Meng, L. Kuang, Block expectation propagation for downlink channel estimation in massive MIMO systems. *IEEE Commun. Lett.* **20**(11), 2225–2228 (2016). <https://doi.org/10.1109/LCOMM.2016.2598810>
36. I. Santos, J. J. Murillo-Fuentes, R. Boloix-Tortosa, E. Arias-de-Reyna, P. M. Olmos, Expectation propagation as turbo equalizer in ISI channels. *IEEE Trans. Commun.* **65**(1), 360–370 (2017). <https://doi.org/10.1109/TCOMM.2016.2616141>
37. A. P. Kannu, P. Schniter, Design and analysis of MMSE pilot-aided cyclic-prefixed block transmissions for doubly selective channels. *IEEE Trans. Signal Process.* **56**(3), 1148–1160 (2008). <https://doi.org/10.1109/TSP.2007.908969>
38. J. van de Beek, B. M. Popovic, in *Proceedings of the IEEE Global Telecommunications Conference*. Multiple access with low-density signatures (IEEE, Honolulu, 2009), pp. 1–6. <https://doi.org/10.1109/GLOCOM.2009.5425243>
39. Y. Wu, S. Zhang, Y. Chen, in *Proceedings of the IEEE International Conference on Communications*. Iterative multiuser receiver in sparse code multiple access systems (IEEE, London, 2015), pp. 2918–2923. <https://doi.org/10.1109/ICC.2015.7248770>
40. Z. Na, Z. Pan, M. Xiong, X. Liu, W. Lu, Y. Wang, L. Fan, Turbo receiver channel estimation for GFDM-based cognitive radio networks. *IEEE Access.* **6**, 9926–9935 (2018). <https://doi.org/10.1109/ACCESS.2018.2803742>
41. I. S. Gaspar, L. L. Mendes, N. Michailow, G. Fettweis, A synchronization technique for generalized frequency division multiplexing. *EURASIP J. Adv. Signal Process.* **2014**(1), 67 (2014). <https://doi.org/10.1186/1687-6180-2014-67>
42. D. Zhang, L. L. Mendes, M. Matthé, I. S. Gaspar, N. Michailow, G. P. Fettweis, Expectation propagation for near-optimum detection of MIMO-GFDM signals. *IEEE Trans. Wirel. Commun.* **15**(2), 1045–1062 (2016). <https://doi.org/10.1109/TWC.2015.2482479>
43. T. P. Minka, *A family of algorithms for approximate Bayesian inference*. PhD thesis. (Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, Cambridge, 2001). AAI0803033
44. L. Bahl, J. Cocke, F. Jelinek, J. Raviv, Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Trans. Inf. Theory.* **20**(2), 284–287 (1974). <https://doi.org/10.1109/TIT.1974.1055186>
45. N. Michailow, M. Matthé, I. S. Gaspar, A. N. Caldevilla, L. L. Mendes, A. Festag, G. Fettweis, Generalized frequency division multiplexing for 5th generation cellular networks. *IEEE Trans. Commun.* **62**(9), 3045–3061 (2014). <https://doi.org/10.1109/TCOMM.2014.2345566>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
