Fairness and Sum-Rate Maximization via Joint Channel and Power Allocation in Uplink SCMA Networks

Joao V.C. Evangelista, *IEEE Student Member,* Zeeshan Sattar, *IEEE Student Member,* Georges Kaddoum, *IEEE Member,* and Anas Chaaban, *IEEE Member*

Abstract—In this work, we consider a sparse code multiple access uplink network, where J users simultaneously transmit data over K subcarriers, such that $J > K$, with a constraint on the power transmitted by each user. To jointly optimize the codebook assignment and the transmitted power per subcarrier, two new iterative algorithms are proposed, the first one aims to maximize the sum-rate (Max-SR) of the network, while the second aims to maximize the minimum rate (Max-Min). In both cases, the optimization problem is of the mixed-integer nonlinear programming (MINLP) type, with non-convex objective functions. To address this issues, we employ an alternating optimization framework, where the codebook allocation problem is decoupled from the power allocation into two subproblems, and they are solved in an iterative alternating fashion. These subproblems are non-convex, therefore, we employ a sequential convex programming (SCP) method to solve them. An analysis of the convergence of the algorithms is provided. A comparison of the sum-rate and Jain's fairness index of the novel algorithms with three other algorithms proposed in the literature is presented for varying power budgets. The Max-SR algorithm outperforms the others in the sum-rate sense, while the Max-Min outperforms them in the fairness sense, achieving near perfect fairness.

Index Terms—SCMA, 5G, Power Allocation, Multiple Access.

I. INTRODUCTION

THE fifth generation (5G) of wireless networks is expected to deliver a better coverage and a higher capacity to massively connected users. One of the fundamental aspects to achieve this goal is the design of multiple access techniques. Orthogonal multiple access (OMA) techniques, allocate different users into orthogonal network resources, in order to minimize the interference between users. In time division multiple access (TDMA), users are scheduled into different time slots, in code division multiple access (CDMA), they are assigned orthogonal codes, and, in orthogonal frequency division multiple access (OFDMA) they are allocated orthogonal subcarriers. However, considering the data consumption trend and the introduction of new data-hungry technologies, such as virtual and augmented reality (VR/AR) and massively

J.V.C. Evangelista, G. Kaddoum and Z. Sattar were with the Department of Electrical Engineering, cole de Technologie Suprieure, Montreal, QC, H3C 1K3 CA, e-mail: {joao-victor.de-carvalho-evangelista.1, zeeshan.sattar.1}@ens.etsmtl.ca and georges.kaddoum@etsmtl.ca

Anas Chaaban is with the School of Engineering, University of British Columbia, Kelowna, BC, V1V 1V7 CA, email: anas.chaaban@ubc.ca

deployed internet of things (IoT) devices, a tenfold increase in traffic is expected by 2020 [1]. As the number of orthogonal network resources available is finite, this design paradigm is incompatible with the massive traffic and connectivity requirements of 5G networks. Recently, early information-theoretic works on multi-user communications [2], [3] have reemerged under the name non-orthogonal multiple access (NOMA) as a potential solution to deal with this requirement. Although NOMA methods are rooted in the information-theoretic literature, the recent interest has been focused on communication-theoretic aspects such as developing practical NOMA coding and modulation schemes, with desired error-rate performance and multi-user communication capabilities. Differently from OMA, in NOMA techniques, multiple users are allocated on the same network resources, permitting the allocation of more users and more efficient use of the available resources. In NOMA, each receiver must perform multiuser detection (MUD) in order to recover the intended transmitted signal. NOMA techniques can be classified into two different groups, power division NOMA (PD-NOMA) and code division NOMA (CD-NOMA).

In PD-NOMA, multiple users, with distinct channel conditions, are served under the same resource. In the downlink, users are sorted according to their channel quality, and more power is allocated to users with worst channel and less power is allocated to users with better channels. On the other hand, in the uplink users with different channel quality are scheduled on the same resource (e.g. a user close to the base station is paired with a user at the edge of the cell). In order to decode messages, successive interference cancellation (SIC) is performed on the decoding end. In the downlink, the users decode messages destined to users with worse channel conditions, and subtract them from the received signal. Messages from the users with the better channel conditions are treated as interference [4], [5, Ch. 6]. Similarly, in the uplink, the base station (BS), decodes the received messages successively, starting from the user with the best channel to the worse, and removing the already decoded messages at each step. A large number of contributions have been made in the development of PD-NOMA technology; the reader may refer to the survey in [6] for a comprehensive summary of the state-of-the-art.

In CD-NOMA, the transmitter introduces redundancy to the transmitted symbol, via code and/or spreading, in order to enable receivers to perform MUD and separate signals from different users. Furthermore, CD-NOMA has additional ad-

vantages in comparison to PD-NOMA [7], such as the coding gain and the shaping gain (i.e. methods using multidimensional constellations) [8]. Motivated by these advantages, this paper is focused on one of the promising CD-NOMA techniques, named sparse code multiple access (SCMA) [8]. In SCMA, sparse multidimensional codebooks are assigned to each user and each user's data layer is sparsely spread throughout the network resources. Detection techniques based on message passing algorithms (MPA) are applied on the receiver side to decode its information. Also, in comparison to OFDMA, SCMA allows for more users than subcarriers available to be served simultaneously, while reducing the peak average power ratio (PAPR) due to the sparsity of the subcarrier allocation. SCMA was first proposed in [8], as a multidimensional generalization of the low density spreading code division multiple access (LDS-CDMA) that yielded better results regarding detection error. In [9], a method to design SCMA codebooks based on lattice coding was proposed. In [10], a downlink SCMA system is considered, and an algorithm for user pairing along with rate adjustment and a detection strategy is proposed for a multiuser SCMA scheme. It is shown that this scheme is able to achieve robustness to mobility and high data rates. One of the main drawbacks of SCMA is its high complexity of decoding the received SCMA symbol with MPA, especially for devices with limited battery life or computing capabilities. However, this approach is suitable for the uplink transmission, as the decoding burden would mainly lie with the base station, which possesses significantly more computing capabilities than user equipment.

In this work, our main goal is to investigate the fairness in resource allocation in uplink SCMA networks. To tackle this challenge, we propose two algorithms, from now on referred as the Max-SR and the Max-Min, for joint channel and power allocation in SCMA networks. The first algorithm's goal is to maximize the sum-rate of the network. While this is an important criterion in cellular networks, fairness between users is equally important. Thus, to include fairness in the optimization, we propose the Max-Min algorithm aimed to maximize the minimum rate among the users. The obtained results demonstrate better fairness than the former, but, this fairness comes at the price of reducing the overall sum-rate performance.

A. Related Work

Regarding resource management and allocation in SCMA networks, an algorithm to maximize the rate of successful accesses on a random access massive machine communications network, is suggested in [11]. In [12], a resource allocation and subcarrier pairing scheme combining OFDMA and SCMA, for a dual-hop multiuser relay network, is proposed. The problem of assigning SCMA codebooks to maximize sum rate in uplink transmission is formulated as matching game in [13]. A grant-free contention based uplink SCMA scheme was proposed in [14]. In [15], the capacity of an SCMA cell with a Gaussian input is derived, and a joint codebook and power allocation algorithm is proposed. In [16], three algorithms for dynamic codebook allocation are presented, and their

link-level performance is evaluated, one of which takes user fairness in consideration. However, the system-level capacity of these algorithms is not investigated and their fairness is compared in terms of the BER difference of the best and the worst user. In [17], a low complexity bisection-based power allocation algorithm, aiming to maximize the capacity of the SCMA system with a finite alphabet is proposed. A stochastic geometry framework to obtain the system wide area spectral efficiency of underlaid and overlaid device-to-device (D2D) SCMA networks is developed in [18], and a power allocation strategy to minimize cross-tier interference in underlaid mode and an optimal codebook allocation for the overlaid mode are presented. To the best of our knowledge, no previous works have investigated the fairness in joint codebook and power allocation in SCMA networks, and, no algorithms for joint codebook and power allocation aiming to maximize the fairness have been proposed so far. Furthermore, in the Max-SR algorithm, we take a different approach from previously proposed solutions, by solving a relaxed approximation of the codebook allocation problem.

B. Contributions

In this paper we formulate two optimization problems for joint codebook and power allocation in SCMA networks, one aiming to maximize the sum-rate and another one for maximizing the fairness. Both problems are of the non-convex mixed integer nonlinear programming (MINLP) type. Currently, there are no known polynomial time algorithms to solve such problems, therefore, we propose some simplifications and heuristics, based on alternating optimization, and, sequential convex programming (SCP), that are able to find high quality solutions. The proposed algorithms split the codebook and power allocation subproblems and solve them in an alternating fashion, where the output of one is the input of the other, until convergence. However, these subproblems are non-convex, so, we employ SCP to obtain a locally optimal solution to the subproblems, by solving a sequence of convexified sub problems. Additionally, we compare both algorithms and the ones proposed in [16] in the sum-rate and in the Jain's fairness index sense. Results show that the Max-SR algorithm outperforms all other algorithms in terms of sum-rate, while the Max-Min algorithm outperforms all others in terms of fairness. To summarize, the list below presents the main accomplishments and insights developed in this work:

- The proposal of the Max-SR algorithm, which employs alternating optimization and SCP, to obtain a convex continuous approximation of the original MINLP. The algorithm achieves a better sum-rate in comparison to the ones proposed in [16].
- The proposal of the Max-Min algorithm, which employs alternating optimization and SCP, to obtain a convex continuous approximation of the original MINLP. The algorithm achieves a better fairness, using Jain's fairness index as the fairness metric, in comparison to the ones proposed in [16].

This paper is organized as follows: Section II contains a brief overview on the SCMA encoder and decoder structure

and a description of SCMA signals and the derivation of its sum-rate are presented. In Section III, the optimization problems are formulated and an algorithm for sum-rate maximization, and, another for fairness maximization are proposed. Furthermore, in Section IV numerical results are shown and the performance of the algorithm is evaluated. Also, a numerical analysis of the convergence is presented. Finally, in Section V the conclusions are presented.

C. Notation

Throughout this paper, italic lowercase letters denote real and complex scalar values, and x^* denotes the complex conjugate of x. Lower case boldface letters denote vectors, while upper case boldface denote matrices. A lowercase letter with one subscript, x_i, represents the i-th element of the vector $\mathbf{x}$, while both $x_{i,j}$ and $[\mathbf{X}]_{i,j}$ are used to denote the element on the i-th row and j-th column of matrix $\mathbf{X}$. The operators $\mathbf{x}^H$ and $\mathbf{X}^H$ denote the hermitian conjugate of a vector and of a matrix, respectively. The operator $\det(\mathbf{X})$ is the determinant of the square matrix $\mathbf{X}$ and $\text{tr}(\mathbf{X})$ is its trace. The operator $\text{diag}(\mathbf{x})$ denotes a square matrix with its diagonal components given by $\mathbf{x}$. The operator $E(\cdot)$ denotes the expected value of a random variable. The function $p(\cdot)$ represents the probability density function (PDF) of a random variable and $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{K})$, where $\mathbf{K} \in \mathbb{R}^n$, denotes that $\mathbf{x}$ is a complex Gaussian random vector, with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{K}$. The sets $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{B}$ are the sets of the reals, complex and binary integers, respectively. A calligraphic uppercase letter, such as $\mathcal{X}$, denotes a set and $|\mathcal{X}|$ is its cardinality. The function $\ln(\cdot)$ denotes the natural logarithm of its argument, while the function $I(\cdot; \cdot)$ is the mutual information between two random variables.

II. System Model

Consider a system consisting of one BS, and let $\mathcal{K}$ be the set of available resources (OFDMA subcarriers, MIMO spatial layers and so on), with $|\mathcal{K}| = K$, and $\mathcal{J}$ be the set of users served by the BS, with $|\mathcal{J}| = J$. Each user transmits a symbol from a multidimensional constellation with, order M. The SCMA encoder is a mapping $f : \mathbb{B}^{\log_2(M)} \rightarrow \mathcal{S}_j$, with $\mathbf{s}_j = f(\mathbf{b}_j)$, where $\mathbf{b}_j \in \mathbb{B}^{\log_2(M)}$ is a vector of bits taken at the output of a channel encoder, $\mathcal{S}_j \subset \mathbb{C}^K$, $|\mathcal{S}_j| = M$ and $\mathbf{s}_j$ is a sparse vector with $N < K$ nonzero elements for all $j \in \mathcal{J}$. Each user encodes its transmitted signal from a different multidimensional constellation [19] $\mathcal{S}_j$. In order to guarantee that the SCMA signals from different users are separable at the receiver each users' constellation has an unique sparsity pattern (i.e. no two users transmit on the same N subcarriers). Therefore, the BS serves up to $J = \binom{K}{N}$ users simultaneously and up to $d_f = \binom{K-1}{N-1}$ users are allocated on the same resource. The overloading factor of the cell is given by $\lambda = J/K$. Figure 1 shows an example of codebooks and a multiplexed codeword for an SCMA system with $K = 4$, $J = 6$, $N = 2$ and $d_f = 3$.

Optimal SCMA decoding is achieved by maximum a priori (MAP) decoding. However, due to the complexity of MAP decoding, message passing algorithms (MPA) that achieve near-optimal decoding, such as belief propagation (BP) [20] are employed, resulting in a complexity of $O(M^{d_f})$. In order

to reduce the decoding complexity of SCMA, alternative receiver architectures have been proposed, such as the SIC-MPA decoder [21], [22], which is an hybrid of the SIC and MPA procedure and the list spherical decoding (LSD) algorithm [23].

The structure of the SCMA code can be neatly conveyed through a factor graph representation. Let $\mathbf{F} \in \mathbb{B}^{K \times J}$ be the factor graph matrix, each element $f_{k,j}$ indicates if any information from the user j is transmitted on resource k. Figure 3 illustrates a factor graph with $J = 6$, $K = 4$, $N = 2$ and $d_f = 3$. Figure 2 shows the structure of a SCMA transceiver. The reader may refer to [8] for more details on the encoder/decoder structure of SCMA.

In a SCMA system, the signal received by the BS at the resource k can be written as

$$y_k = \sum_{j \in \mathcal{J}} f_{k,j} h_{k,j} s_{k,j} + n_k, \tag{1}$$

where $h_{k,j}$ is the channel coefficient, $s_{k,j}$ is the symbol transmitted from user j on the k-th resource, with average power $p_{k,j} = E(|s_{k,j}|^2)$, and n_k is the k-th component of $\mathbf{n} \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I})$. Here, we assume that $h_{k,j} = \frac{g_{k,j}}{\sqrt{1 + r_j^\alpha}}$. Without loss of generality, we assume $g_{k,j}$ is a Rayleigh distributed random variable representing the small scale fading of the channel of user j on subcarrier k, r_j is the distance of user j from the BS and α is the path loss exponent. Throughout this work, it is assumed that the users send a pilot sequence periodically, and the BS is able to perfectly estimate the channel state information (CSI). From (1), the received signal vector at the BS is written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{2}$$

where $\mathbf{y} \in \mathbb{C}^K$ is a complex vector, $\mathbf{H} \in \mathbb{C}^{K \times KJ}$ is a matrix composed of submatrices, such that, $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \cdots, \mathbf{H}_J]$, where, $\mathbf{H}_j = \text{diag}([h_{1,j}, h_{2,j}, \cdots h_{K,j}]^T) \; \forall \; j \in \mathcal{J}$. The vector $\mathbf{x} \in \mathbb{C}^{KJ}$ is given by $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1^T & \mathbf{x}_2^T & \cdots & \mathbf{x}_J^T \end{bmatrix}^T$, where $\mathbf{x}_j = \begin{bmatrix} f_{1,j} s_{1,j} & f_{2,j} s_{2,j} & \cdots & f_{K,j} s_{K,j} \end{bmatrix}^T \; \forall \; j \in \mathcal{J}$.

The sum-rate of a SCMA system is defined as the maximum mutual information between the received and transmitted signals. Therefore, assuming channel knowledge at the receiver we have that

$$
\begin{aligned}
R_{\text{SCMA}}^{\text{sum}} &= \max_{p(\mathbf{x})} I(\mathbf{x}; \mathbf{y}|\mathbf{H}) \\
&= \max_{p(\mathbf{x})} h(\mathbf{y}|\mathbf{H}) - h(\mathbf{n}) \\
&\overset{(a)}{\leq} \ln[(\pi e)^K \det(\sigma_n^2 \mathbf{I}_K + \mathbf{H}\mathbf{K}_x\mathbf{H}^H)] - K\ln[\pi e \sigma_n^2] \\
&= \ln\left[\det\left(\mathbf{I}_K + \frac{1}{\sigma_n^2}\mathbf{H}\mathbf{K}_x\mathbf{H}^H\right)\right]. \tag{3}
\end{aligned}
$$

In (3), the inequality in (a) comes from the fact that for a given variance a Gaussian input maximizes the entropy of a random vector [24]. In this paper we are concerned on maximizing this upper bound in the Max-SR algorithm, and, from now the upper bound is referred by C_{SCMA}. It is worth noting that for an increase on d_f the distribution of $\mathbf{y}$ approaches a multivariate Gaussian, due to the central limit theorem. Figure
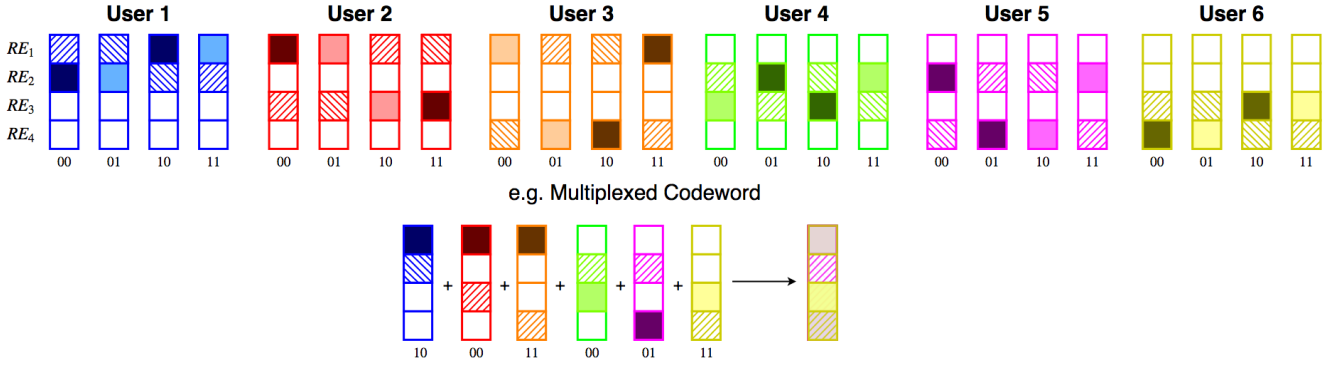
Fig. 1. Example of an SCMA uplink system with $J = 6$, $K = 4$, $N = 2$ and $d_f = 3$. The square arrays demonstrate the codebook of each user and each square represent the available resource elements (RE). An empty square indicates that no signal is transmitted in the RE and different filling patterns indicate a different complex value.
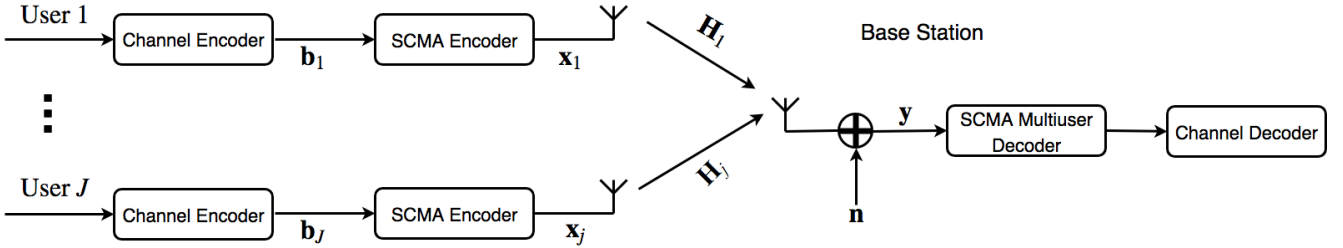

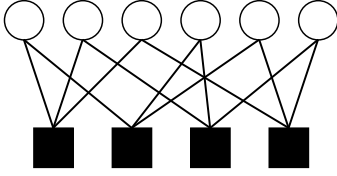Fig. 2. Illustration of an SCMA transceiver.

Fig. 3. Example factor graph with $J = 6$, $K = 4$, $N = 2$ and $d_f = 3$. The circles denote user nodes and the squares denote resource nodes.

Fig. 4. Histogram of $\mathbf{y}$, with $d_f = 3$, $K = 4$ and $J = 6$.

4 shows the histogram and a corresponding Gaussian fit for 10000 realizations of $\mathbf{y}$, with $d_f = 3$, $K = 4$ and $J = 6$, where the non zero coordinates of each $\mathbf{x}_j$ is drawn from a finite constellation of order 4. Furthermore, $\mathbf{K}_x$ is the covariance matrix of $\mathbf{x}$, and is given by

$$
\mathbf{K}_x = \begin{bmatrix} E(\mathbf{x}_1\mathbf{x}_1^H) & E(\mathbf{x}_1\mathbf{x}_2^H) & \cdots & E(\mathbf{x}_1\mathbf{x}_J^H) \\ \vdots & \vdots & \ddots & \vdots \\ E(\mathbf{x}_J\mathbf{x}_1^H) & E(\mathbf{x}_J\mathbf{x}_2^H) & \cdots & E(\mathbf{x}_J\mathbf{x}_J^H) \end{bmatrix}, \tag{4}
$$

where, each $E(\mathbf{x}_i\mathbf{x}_j^H)$ is given by

$$
E(\mathbf{x}_i\mathbf{x}_j^H) = \begin{bmatrix} E\left(x_{1,i}x_{1,j}^*\right) & E\left(x_{1,i}x_{2,j}^*\right) & \cdots & E\left(x_{1,i}x_{K,j}^*\right) \\ \vdots & \vdots & \ddots & \vdots \\ E\left(x_{K,i}x_{1,j}^*\right) & E\left(x_{K,i}x_{2,j}^*\right) & \cdots & E\left(x_{K,i}x_{K,j}^*\right) \end{bmatrix}, \tag{5}
$$

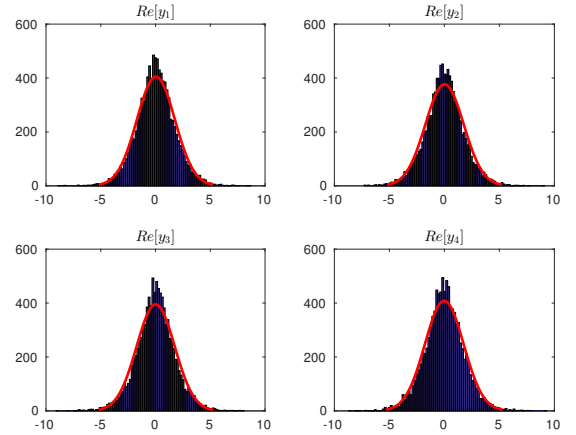therefore, $\mathbf{K}_x \in \mathbb{C}^{KJ \times KJ}$. The transmitted symbols from different users are uncorrelated, thus, all the non-diagonal entries of (4) are equal to zero. Furthermore, we have that

$$
\mathbf{H}\mathbf{K}_x\mathbf{H}^H = \sum_{j=1}^{J} \mathbf{H_j} E\left(\mathbf{x}_j\mathbf{x}_j^H\right) H_j^H =
$$
$$
\begin{bmatrix} \sum_{j=1}^{J} h_{1,j} E\left(x_{1,j}x_{1,j}^*\right) h_{1,j}^* & \cdots & \sum_{j=1}^{J} h_{1,j} E\left(x_{1,j}x_{k,j}^*\right) h_{k,j}^* \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{J} h_{k,j} E\left(x_{k,j}x_{1,j}^*\right) h_{1,j}^* & \cdots & \sum_{j=1}^{J} h_{k,j} E\left(x_{k,j}x_{k,j}^*\right) h_{k,j}^* \end{bmatrix} \tag{6}
$$
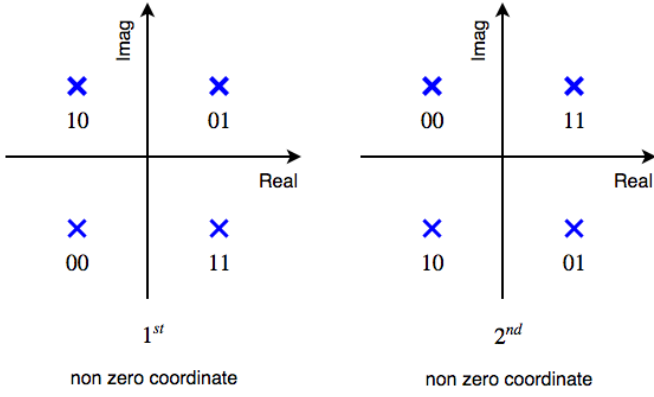
Fig. 5. Example of diagonally symmetric constellation for $N = 2$.

achievable rate of user j on resource k, $C_{k,j}$, as

$$
\ln\left(1 + \frac{\sum\limits_{j \in \mathcal{J}} |h_{k,j}|^2 f_{k,j} p_{k,j}}{\sigma_n^2}\right)
$$

$$
= \ln\left(\frac{\sigma_n^2 + \sum\limits_{j \in \mathcal{J}} |h_{k,j}|^2 f_{k,j} p_{k,j}}{\sigma_n^2}\right)
$$

$$
= \ln\left(\frac{\sigma_n^2 + |h_{k,1}|^2 f_{k,1} p_{k,1}}{\sigma_n^2}\right)
$$

$$
+ \ln\left(\frac{\sigma_n^2 + |h_{k,1}|^2 f_{k,1} p_{k,1} + |h_{k,2}|^2 f_{k,2} p_{k,2}}{\sigma_n^2 + |h_{k,1}|^2 f_{k,1} p_{k,1}}\right)
$$

$$
+ \cdots + \ln\left(\frac{\sigma_n^2 + \sum\limits_{i=1}^{J-1} |h_{k,i}|^2 f_{k,i} p_{k,i} + |h_{k,J}|^2 f_{k,J} p_{k,J}}{\sigma_n^2 + \sum\limits_{i=1}^{J-1} |h_{k,i}|^2 f_{k,i} p_{k,i}}\right)
$$

$$
= \sum_{j \in \mathcal{J}} \ln\left(1 + \frac{|h_{k,j}|^2 f_{k,j} p_{k,j}}{\sigma_n^2 + \sum\limits_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i} p_{k,i}}\right)
$$

$$
= \sum_{j \in \mathcal{J}} C_{k,j}. \tag{8}
$$

Therefore, the achievable rate of each user, C_j, is given by

$$
C_j = \sum_{k \in \mathcal{K}} C_{k,j} = \sum_{k \in \mathcal{K}} \ln\left(1 + \frac{|h_{k,j}|^2 f_{k,j} p_{k,j}}{\sigma_n^2 + \sum\limits_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i} p_{k,i}}\right). \tag{9}
$$

In the next section, we formulate the joint codebook and power allocation problems and present our proposed method to solve them.

As concluded in [15], if each nonzero coordinate of **s** is drawn from centrally symmetric constellations, the cross correlation between the signals transmitted in different subcarriers $E(x_{a,j} x_{b,j}*) = 0$, thus, (6) is diagonal. Figure 5 shows an example of centrally symmetric constellations for $N = 2$. It is worth noting that such codebook satisfies the design principles established in [9]. With that being said the SCMA sum-rate in (3) can be simplified as

III. Joint Codebook and Power Allocation

In this section, we formulate and propose two joint codebook and power allocation algorithms to solve two optimization problems: maximizing the sum-rate ($\mathbf{P_{Max\text{-}SR}}$) and maximizing the minimum rate of users ($\mathbf{P_{Max\text{-}Min}}$). The former can be formulated as

$$
\begin{aligned}
& C_{\text{SCMA}} \\
&= \ln\det\left(\mathbf{I}_K + \frac{1}{\sigma_n^2}\mathbf{H}\mathbf{K}_x\mathbf{H}^H\right) \\
&= \ln\det\begin{bmatrix} 1 + \frac{\sum\limits_{j \in \mathcal{J}} |h_{1,j}|^2 E(x_{1,j}^2)}{\sigma_n^2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 + \frac{\sum\limits_{j \in \mathcal{J}} |h_{K,j}|^2 E(x_{K,j}^2)}{\sigma_n^2} \end{bmatrix} \\
&= \sum_{k \in \mathcal{K}} \ln\left(1 + \frac{\sum\limits_{j \in \mathcal{J}} |h_{k,j}|^2 E(x_{k,j}^2)}{\sigma_n^2}\right) \\
&= \sum_{k \in \mathcal{K}} \ln\left(1 + \frac{\sum\limits_{j \in \mathcal{J}} |h_{k,j}|^2 f_{k,j} p_{k,j}}{\sigma_n^2}\right). \tag{7}
\end{aligned}
$$

Furthermore, by assuming a decoding order starting from user J to user 1, and, using the logarithm identity $\log_b(a + c) = \log_b(a) + \log_b(1 + \frac{c}{a})$, it is possible to obtain the

$\mathbf{P_{Max\text{-}SR}}$:

$$
\max_{\mathbf{P},\mathbf{F}} \quad C_{\text{SCMA}} = \sum_{k \in \mathcal{K}} \ln\left(1 + \frac{\sum\limits_{j \in \mathcal{J}} |h_{k,j}|^2 f_{k,j} p_{k,j})}{\sigma_n^2}\right) \tag{10}
$$

$$
\textbf{s.t.} \quad \sum_{k \in \mathcal{K}} f_{k,j} \leq N \ \forall \ j \in \mathcal{J} \tag{11}
$$

$$
\sum_{j \in \mathcal{J}} f_{k,j} \leq d_f \ \forall \ k \in \mathcal{K} \tag{12}
$$

$$
\sum_{k \in \mathcal{K}} f_{k,j} p_{k,j} \leq P_{\max}^{(j)} \ \forall \ j \in \mathcal{J} \tag{13}
$$

$$
f_{k,j} \in \{0, 1\} \ \forall \ k \in \mathcal{K} \text{ and } \forall \ j \in \mathcal{J}, \tag{14}
$$

where $\mathbf{P} \in \mathbb{R}^{K \times J}$ is the matrix of allocated power, (10) is the sum-rate, (11) is the constraint on the number of subcarriers

allocated per user. The constraint on the number of users per subcarrier is given by (12), while (13) is the constraint on the maximum transmitting power available per user. Finally, (14) is a binary constraint on the values of $f_{k,j}$. The objective function is concave in $\mathbf{P}$ and all the constraints are linear, however, it is non-convex due to the product between the entries of $\mathbf{F}$ and $\mathbf{P}$. Moreover, due to the presence of the integer constraint (14), this problem fall in the category of non-convex MINLP problems, which are in general NP-hard [25]. Furthermore, the problem $\mathbf{P_{Max\text{-}Min}}$ is formulated as

$$\mathbf{P_{Max\text{-}Min}}:$$

$$\max_{\mathbf{P},\mathbf{F}} \quad \min_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} \ln\left(1 + \frac{|h_{k,j}|^2 f_{k,j} p_{k,j}}{\sigma_n^2 + \sum\limits_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i} p_{k,i}}\right) \quad (15)$$

$$\mathbf{s.t.} \quad (11), (12), (13), (14),$$

where (15) is the max-min utility function of the rate per user. The objective function of this problem is non-concave and, similarly to $\mathbf{P_{Max\text{-}SR}}$, also has integer constraints on $\mathbf{F}$.

In order to provide good solutions to both problems in polynomial time, we propose two algorithms, employing heuristics, to approximately solve this problems. The first approach, common to both problems, is to divide the main problem into two subproblems, one optimizing the transmitted power while maintaining the codebook allocation fixed, and the other one optimizing the codebook allocation while keeping the transmitted power fixed. However, the codebook allocation problem remains an integer problem, thus, we solve a continuous relaxed version of the problem and add a penalty function to bias the solution towards an integer one. Then, the subproblems are solved in alternating fashion until convergence. This alternating optimization framework has been widely used to solve joint allocation problems in wireless networks [26]–[31]. The reformulated problems within the alternating framework are both non-convex, thus, in addition to the alternating framework, we propose a sequential convex programming (SCP) approach, to locally solve a sequence of concave approximations of the original problem until convergence. In the rest of this section, the reformulation of the problems, the details of the employed heuristics, and, the analysis of the convergence of the proposed algorithms are thoroughly discussed.

A. Alternating Optimization

Consider a generic optimization problem of the form

$$\begin{aligned} \max_{\mathbf{x},\mathbf{y}} \quad & f(\mathbf{x},\mathbf{y}) \\ \mathbf{s.t} \quad & g_i(\mathbf{x},\mathbf{y}) \le 0 \ \forall \ i = 1, \cdots, N \\ & \mathbf{x} \in \mathbb{R}^{D_1}, \ \mathbf{y} \in \mathbb{R}^{D_2} \end{aligned}$$

such that f and g_i's are convex on $\mathbf{x}$ for a constant $\mathbf{y}$, and on $\mathbf{y}$ with a constant $\mathbf{x}$, but are not convex on $\mathbf{x}$ and $\mathbf{y}$ together.

The alternating procedure consists on solving separate convex problems iteratively. At the n-th alternating step we obtain

$$\begin{aligned} \mathbf{x}_n &= \arg\max_{\mathbf{x}} f(\mathbf{x}, \mathbf{y}_{n-1}), \ \{\mathbf{x} \in \mathbb{R}^{D_1} \mid g_i(\mathbf{x}, \mathbf{y}_{n-1}) \le 0\} \\ \mathbf{y}_n &= \arg\max_{\mathbf{y}} f(\mathbf{x}_n, \mathbf{y}), \ \{\mathbf{y} \in \mathbb{R}^{D_2} \mid g_i(\mathbf{x}_n, \mathbf{y}) \le 0\}, \end{aligned}$$

where $\mathbf{x}_n$ and $\mathbf{y}_n$ are the values of the variables at the n-th alternating step. The algorithm stops when $\|\mathbf{x}_n - \mathbf{x}_{n-1}\| \le \epsilon_x$ and $\|\mathbf{y}_n - \mathbf{y}_{n-1}\| \le \epsilon_y$.

1) Max-SR Problem: Starting by $\mathbf{P_{Max\text{-}SR}}$, the original problem is split into two subproblems: $\mathbf{P^{(P)}_{Max\text{-}SR}}$ and $\mathbf{P^{(F)}_{Max\text{-}SR}}$. The former is given by

$$\mathbf{P^{(P)}_{Max\text{-}SR}}: \quad \max_{\mathbf{P}} \quad (10) \quad \mathbf{s.t.} \quad (13).$$

This new problem is convex and can be readily solved with an optimization package, such as CVX [32]. As the codebook allocation $\mathbf{F}$ is integer a few modifications are made to turn $\mathbf{P^{(F)}_{Max\text{-}SR}}$ into a continuous problem. Firstly, the integer variables $f_{k,j}$ are relaxed to continuous ones inside the interval $[0,1]$. Secondly, a penalty term of the form $\gamma(\mathbf{F}) = \lambda \sum\limits_{k \in \mathcal{K}} \sum\limits_{j \in \mathcal{J}} f_{k,j}^2 - f_{k,j}$, where λ is chosen to be a large number, is added to the objective function. As $f_{k,j} \in [0,1]$, we have that $\gamma(\mathbf{F}) \le 0$, with equality if $f_{k,j} \in \{0,1\}$. This penalty term drives the solution towards an integer value of $f_{k,j}$. Finally, the second problem, $\mathbf{P^{(F)}_{Max\text{-}SR}}$, is given by

$$\mathbf{P^{(F)}_{Max\text{-}SR}}:$$

$$\max_{\mathbf{F}} \quad \sum_{k \in \mathcal{K}} \ln\left(1 + \frac{\sum\limits_{j \in \mathcal{J}} |h_{k,j}|^2 f_{k,j} p_{k,j}}{\sigma_n^2}\right) + \gamma(\mathbf{F}) \quad (16)$$

$$\mathbf{s.t.} \quad 0 \le f_{k,j} \le 1 \ \forall \ k \in \mathcal{K} \text{ and } \forall \ j \in \mathcal{J} \quad (17)$$
$$(11), (12), (13).$$

Due to the fact that $\gamma(\mathbf{F})$ is a convex function, (16) is non-concave. The procedure to turn the objective function into a concave is discussed in the next subsection.

2) Max-Min Problem: A similar procedure is applied to $\mathbf{P_{Max\text{-}Min}}$, by splitting it into $\mathbf{P^{(P)}_{Max\text{-}Min}}$ and $\mathbf{P^{(F)}_{Max\text{-}Min}}$. Furthermore, an auxiliary variable $z \in \mathbb{R}$ is added to rewrite the original formulation of the objective function and a new set of constraints is added to guarantee the equivalence between the problems. Then, $\mathbf{P^{(P)}_{Max\text{-}Min}}$ is given by

$$\mathbf{P^{(P)}_{Max\text{-}Min}}:$$

$$\max_{\mathbf{P},z} \quad z \quad (18)$$

$$\mathbf{s.t} \quad z \le \sum_{k \in \mathcal{K}} \ln\left(1 + \frac{|h_{k,j}|^2 f_{k,j} p_{k,j}}{\sigma_n^2 + \sum\limits_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i} p_{k,i}}\right) \ \forall \ j \in \mathcal{J}$$

$$(19)$$

$$(13).$$

The set of constraints given by (19) are non-convex. Similarly to $\mathbf{P^{(F)}_{Max\text{-}SR}}$, the integer variables $f_{k,j}$ are relaxed into contin-

uous ones and a penalty factor $\gamma(\mathbf{F})$ is added to the objective, resulting in

$$
\mathbf{P}_{\mathbf{Max\text{-}Min}}^{(\mathbf{F})}:
$$
$$
\max_{\mathbf{F}} \quad z + \gamma(\mathbf{F}) \tag{20}
$$
$$
\mathbf{s.t} \quad (19), (11), (12), (13), (17)
$$

Both (16) and (20) are non-concave objective functions, and the constraints defined by (19) are non-convex, which makes the subproblems hard to solve. In the next section we present a sequential convex programming (SCP) approach to iteratively solve convex approximate versions of these problems, that can be solved efficiently.

B. Sequential Convex Programming

In order to solve the non-convex optimization problems presented in the last section, we employ the SCP framework [33]. In the SCP formulation, a sequence of subproblems, consisting of convexifications of the original problem, are solved. The solution of each subproblem is used as the input to the next one. The sequence of solutions converges to a local optimum of the original problem.

For a general constrained optimization problem consider

$$
\max_{x} \ f(x) \quad \mathbf{s.t.} \quad g_i(x) \le 0 \ , \ \forall i \in \{1, 2, \dots, N\},
$$

where $g_i(x) \ \forall \ i = 1 \cdots M$, with $M < N$, are convex functions, and $g_i(x) \ \forall \ i = M + 1 \cdots N$ are non-convex functions and the objective function $f(x)$ is non-concave. Let $\tilde{f}(x, y)$ be a lower bound approximation of $f(x)$ in the neighborhood of y, and, $\tilde{g}_i(x, y) \ \forall \ i = M + 1 \cdots N$ be upper bound approximations of the non-convex constraints in the neighborhood of y, satisfying the following properties:

(a) Both $\tilde{f}(x, y)$ and $\tilde{g}_i(x, y)$ are continuous and differentiable in x.
(b) The function $\tilde{f}(x, y)$ is concave in x.
(c) The functions $\tilde{g}_i(x, y)$ are convex in x
(d) $\tilde{f}(x, x) = f(x)$ and $\tilde{g}_i(x, x) = g(x)$, $\forall x$.
(e) $\nabla \tilde{f}(x, x) = \nabla f(x)$ and $\nabla \tilde{g}_i(x, x) = \nabla g_i(x)$, $\forall x$.
(f) $\tilde{f}(x, y) \le f(x)$ and $\tilde{g}_i(x, y) \ge g_i(x)$, $\forall x, y$.

Summarizing, $\tilde{f}(x, y)$ is a first order, concave, lower bound approximation of $f(x)$ and $\tilde{g}_i(x, y)$ are first order, convex, upper bound approximations of $g_i(x)$.

Theorem 1. *Let $\hat{\mathbf{x}}_n = \{x^0, x^1, \dots, x^{n-1}\}$ be a sequence such that $x^n = x_*^{n-1}$, where x_*^{n-1} is a solution to the problem*

$$
\mathbf{P}(n): \quad \max_{x} \quad \tilde{f}(x, x^{n-1}) \ , \ x^{n-1} \in \hat{\mathbf{x}}_n \tag{21}
$$
$$
\mathbf{s.t.} \quad g_i(x) \le 0 \ , \ \forall i \in \{1, \dots, M\}.
$$
$$
\tilde{g}_i(x, x^{n-1}) \le 0 \ , \ \forall i \in \{M + 1, \dots, N\}. \tag{22}
$$

Assuming the sequence $\hat{\mathbf{x}}_n$ converges to $\bar{x}$ and Slater's condition hold on $\bar{x}$, then $\bar{x}$ is a KKT point of the original problem, thus, a local maximum.

Proof. The proof is relegated to Appendix A. $\qquad\square$

1) Max-SR Problem: In the sum-rate maximization problem, the power optimization step, $\mathbf{P}_{\mathbf{Max\text{-}SR}}^{(\mathbf{P})}$, is already convex and does not need any modifications. In order to solve $\mathbf{P}_{\mathbf{Max\text{-}SR}}^{(\mathbf{F})}$, a sequence of convex sub problems denoted by $\mathbf{P}_{\mathbf{Max\text{-}SR}}^{(\mathbf{F})}(t)$ are solved to obtain a locally optimal solution to the non-convex problem $\mathbf{P}_{\mathbf{Max\text{-}SR}}^{(\mathbf{F})}$, where the index t represents the t-th SCP subproblem. Both (16) and (20), the objective functions of $\mathbf{P}_{\mathbf{Max\text{-}SR}}^{(\mathbf{F})}$ and $\mathbf{P}_{\mathbf{Max\text{-}Min}}^{(\mathbf{F})}$ respectively, consist of the sum of a concave function with a convex one so they follow the same convexification procedure. The objective function of the subproblem $\mathbf{P}_{\mathbf{Max\text{-}SR}}^{(\mathbf{F})}$, given in (16), consist of the sum of a concave function with a convex one.

Lemma 1. *Let $\mathbf{F}^{(t)}$ be the codebook allocation matrix at the t-th SCP iteration, $\Omega(\mathbf{F})$ be a concave function, and, $f(\mathbf{F}) = \Omega(\mathbf{F}) + \gamma(\mathbf{F})$. Then, the function*

$$
\tilde{f}\left(\mathbf{F}, \mathbf{F}^{(t)}\right) = \Omega(\mathbf{F}) + \gamma\left(\mathbf{F}^{(t)}\right) + tr\left[\nabla\gamma\left(\mathbf{F}^{(t)}\right)^T \left(\mathbf{F} - \mathbf{F}^{(t)}\right)\right],
$$

where $\nabla\gamma\left(\mathbf{F}^{(t)}\right) \in \mathbb{R}^{K \times J}$ is a matrix such that

$$
\left[\nabla\gamma\left(\mathbf{F}^{(\mathbf{t})}\right)\right]_{k,j} = \frac{\partial\gamma\left(\mathbf{F}^{(\mathbf{t})}\right)}{\partial f_{k,j}^{(t)}} \tag{23}
$$

satisfies properties (a), (b), (d), (e) and (f).

Proof. Properties (a), (d), (e) are trivial to prove. The function $\tilde{f}\left(\mathbf{F}, \mathbf{F}^{(t)}\right)$ is the sum of a concave and an affine function, therefore, it is concave and property (b) is satisfied. Moreover, as $\gamma(\mathbf{F})$ is a convex function, we have that

$$
\gamma(\mathbf{F}) \ge \gamma\left(\mathbf{F}^{(t)}\right) + tr\left[\nabla\gamma\left(\mathbf{F}^{(t)}\right)^T \left(\mathbf{F} - \mathbf{F}^{(t)}\right)\right],
$$

thus, property (f) is also satisfied, completing the proof. $\quad\square$

From Lemma 1 the subproblems to solve $\mathbf{P}_{\mathbf{Max\text{-}SR}}^{(\mathbf{F})}$ in the SCP framework are given by

$$
\mathbf{P}_{\mathbf{Max\text{-}SR}}^{(\mathbf{F})}(t):
$$
$$
\max_{\mathbf{F}} \quad \sum_{k \in \mathcal{K}} \ln\left(1 + \frac{\sum_{j \in \mathcal{J}} |h_{k,j}|^2 f_{k,j} p_{k,j}}{\sigma_n^2}\right) +
$$
$$
\gamma\left(\mathbf{F}^{(t-1)}\right) + tr\left[\nabla\gamma\left(\mathbf{F}^{(t)}\right)^T \left(\mathbf{F} - \mathbf{F}^{(t)}\right)\right] \tag{24}
$$
$$
\mathbf{s.t.} \quad (11), (12), (13), (17).
$$

The complete algorithm to solve the sum-rate maximization problem, $\mathbf{P}_{\mathbf{Max\text{-}SR}}$, is summarized in Algorithm 1.

2) Max-Min Problem: For the max-min problem, both steps $\mathbf{P}_{\mathbf{Max\text{-}Min}}^{(\mathbf{P})}$ and $\mathbf{P}_{\mathbf{Max\text{-}Min}}^{(\mathbf{F})}$ are non-convex. In both steps, a locally optimal solution for each of the problems is obtained by solving a sequence of convex sub problems $\mathbf{P}_{\mathbf{Max\text{-}Min}}^{(\mathbf{P})}(t)$ and $\mathbf{P}_{\mathbf{Max\text{-}Min}}^{(\mathbf{F})}(t)$. Both (16) and (20), the objective functions of $\mathbf{P}_{\mathbf{Max\text{-}SR}}^{(\mathbf{F})}$ and $\mathbf{P}_{\mathbf{Max\text{-}Min}}^{(\mathbf{F})}$, respectively, consist of the sum of a concave function with a convex one so they follow the same convexification procedure based on Lemma 2. A

Algorithm 1: Maximization of sum-rate

Variable Definition

1. $\mathbf{P}_n$ is the power allocation matrix at the n-th alternating step.
2. $\mathbf{F}_n$ is the codebook allocation matrix at the n-th alternating step.
3. $\mathbf{F}^{(t)}$ is the codebook allocation matrix at the t-th SCP step.
4. $C^{(t)}$ is the sum-rate at the t-th SCP step.

Initialization

1. Set the initial values for the power allocation matrix $\mathbf{P}_0$ randomly, within the set defined by constraint (13), and, the codebook allocation matrix $\mathbf{F}_0$ within the set defined by constraints (11), (12) and (17).
2. Set the convergence tolerance for the codebook allocation ϵ_F and for the power allocation ϵ_P.
3. Set the tolerance ϵ_1 for the convergence of the SCP sub problems of $P_{Max\text{-}SR}^{(F)}$
4. $n \leftarrow 0$

Output

1. Optimized power allocation $\mathbf{P}^*$.
2. Optimized codebook allocation $\mathbf{F}^*$.

while $\|\mathbf{F}_n - \mathbf{F}_{n-1}\| > \epsilon_F$ and $\|\mathbf{P}_n - \mathbf{P}_{n-1}\| > \epsilon_P$ **do**

> $n \leftarrow n + 1$;
> $\mathbf{F}^{(0)} \leftarrow \mathbf{F}_{n-1}$;
> $C^{(0)} \leftarrow C_{\text{SCMA}}\left(\mathbf{F}^{(0)}, \mathbf{P}_{n-1}\right)$; (see (7))
> $t \leftarrow 0$;
> **while** $\|C^{(t)} - C^{(t-1)}\| > \epsilon_1$ **do**
>> $t \leftarrow t + 1$;
>> $\mathbf{F}^{(t)} \leftarrow \mathbf{P}_{Max\text{-}SR}^{(F)}(t-1)$; (see (24))
>> $C^{(t)} \leftarrow C_{\text{SCMA}}\left(\mathbf{F}^{(t)}, \mathbf{P}_\mathbf{n}\right)$; (see (7))
> **end**
> $\mathbf{F}_n \leftarrow \mathbf{F}^{(t)}$;
> $\mathbf{P}_n \leftarrow \mathbf{P}_{Max\text{-}SR}^{(P)}$; (see (10))

end
$\mathbf{P}^* \leftarrow \mathbf{P}_n$;
$\mathbf{F}^* \leftarrow \mathbf{F}_n$;

similar procedure is employed to convexify the constraints of $\mathbf{P}_{Max\text{-}Min}^{(F)}$. Firstly, notice that

$$\sum_{k \in \mathcal{K}} \ln\left(1 + \frac{|h_{k,j}|^2 f_{k,j} p_{k,j}}{\sigma_n^2 + \sum_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i} p_{k,i}}\right)$$

$$= \sum_{k \in \mathcal{K}} \ln\left(|h_{k,j}|^2 f_{k,j} p_{k,j} + \sigma_n^2 + \sum_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i} p_{k,i}\right) -$$
$$\sum_{k \in \mathcal{K}} \ln\left(\sigma_n^2 + \sum_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i} p_{k,i}\right), \qquad (25)$$

which is the difference of concave functions. The constraints given by (19) in $\mathbf{P}_{Max\text{-}Min}^{(F)}$ can be approximated as follows

Lemma 2. *Let* $\theta_j\left(\mathbf{F}^{(t)}\right) = \sum_{k \in \mathcal{K}} \ln\left(\sigma_n^2 + \sum_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i}^{(t)} p_{k,i}\right)$, *then the function*

$$z - \left[\sum_{k \in \mathcal{K}} \ln\left(|h_{k,j}|^2 f_{k,j} p_{k,j} + \sigma_n^2 + \sum_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i} p_{k,i}\right) - \theta_j\left(\mathbf{F}^{(t)}\right) - tr\left(\nabla\theta_j\left(\mathbf{F}^{(t)}\right)^T \left(\mathbf{F} - \mathbf{F}^{(t)}\right)\right)\right], \qquad (26)$$

where $\nabla\theta_j\left(\mathbf{F}^{(t)}\right) \in \mathbb{R}^{K \times J}$ *is a matrix such that*

$$\left[\nabla\theta_j\left(\mathbf{F}^{(t)}\right)\right]_{k,n} = \frac{\partial\theta_j\left(\mathbf{F}^{(t)}\right)}{\partial f_{k,n}^{(t)}}$$
$$= \begin{cases} \frac{|h_{k,n}|^2 p_{k,n}}{\sigma_n^2 + \sum_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i}^{(t)} p_{k,i}} & , \forall\, n < j \\ 0 & , otherwise \end{cases}, \qquad (27)$$

satisfies the properties (a), (c), (d), (e) and (f) for the constraints of $\mathbf{P}_{Max\text{-}Min}^{(F)}$.

Proof. The proof of properties (a), (d) and (e) is straightforward. The term inside the outermost bracket in (26) consists of the subtraction of an affine function from a concave function in $\mathbf{F}$, thus, the result is concave. Furthermore, the whole expression (26) consists of an affine function minus a concave one, therefore, it is a convex function of $\mathbf{F}$ and (c) is satisfied. Also, as $\theta_j(\mathbf{F}^{(t)})$ is concave we have that

$$\theta_j(\mathbf{F}) \leq \theta_j\left(\mathbf{F}^{(t)}\right) - tr\left[\nabla\theta_j\left(\mathbf{F}^{(t)}\right)^T \left(\mathbf{F} - \mathbf{F}^{(t)}\right)\right],$$

so,

$$(26) \geq$$
$$z - \left[\ln\left(|h_{k,j}|^2 f_{k,j} p_{k,j} + \sigma_n^2 + \sum_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i} p_{k,i}\right) - \theta_j(\mathbf{F})\right],$$

and property (f) is satisfied, completing the proof. $\square$

Similarly, the constraints given by (19) in $\mathbf{P}_{Max\text{-}Min}^{(P)}$ can be approximated as follows

Lemma 3. *Let* $\theta_j(\mathbf{P}^{(t)}) = \sum_{k \in \mathcal{K}} \ln\left(\sigma_n^2 + \sum_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i} p_{k,i}^{(t)}\right)$, *then the function*

$$z - \left[\sum_{k \in \mathcal{K}} \ln\left(|h_{k,j}|^2 f_{k,j} p_{k,j} + \sigma_n^2 + \sum_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i} p_{k,i}\right) - \theta_j\left(\mathbf{P}^{(t)}\right) - tr\left(\nabla\theta_j\left(\mathbf{P}^{(t)}\right)^T \left(\mathbf{P} - \mathbf{P}^{(t)}\right)\right)\right], \qquad (28)$$

where $\nabla\theta_j\left(\mathbf{P}^{(t)}\right) \in \mathbb{R}^{K \times J}$ *is a matrix such that*

$$\left[\nabla\theta_j\left(\mathbf{P}^{(t)}\right)\right]_{k,n} = \frac{\partial\theta_j\left(\mathbf{P}^{(t)}\right)}{\partial p_{k,n}^{(t)}}$$
$$= \begin{cases} \frac{|h_{k,n}|^2 f_{k,n}}{\sigma_n^2 + \sum_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i} p_{k,i}^{(t)}} & , \forall\, n < j \\ 0 & , otherwise \end{cases}, \qquad (29)$$

satisfies the properties (a), (c), (d), (e) and (f) for the constraints of $P_{Max\text{-}Min}^{(P)}$.

Proof. The proof follow the same procedure as Lemma 2. □

The power allocation SCP subproblems of $\mathbf{P_{Max\text{-}Min}}$ are given by

$$\mathbf{P}_{\mathbf{Max\text{-}Min}}^{(\mathbf{P})}(t) :$$

$$\max_{\mathbf{P}} \quad z$$

$$\text{s.t.} \quad z - \sum_{k \in \mathcal{K}} \ln \left(h_{k,j} f_{k,j} p_{k,j} + \sigma_n^2 + \sum_{i=1}^{j-1} h_{k,i} f_{k,i} p_{k,i} \right) +$$

$$\theta_j \left(\mathbf{P}^{(t-1)} \right) + \text{tr} \left[\nabla \theta_j \left(\mathbf{P}^{(t-1)} \right)^T \left(\mathbf{P} - \mathbf{P}^{(t-1)} \right) \right] \leq 0$$

$$\forall \; j \in \mathcal{J} \quad\quad\quad (30)$$

$$(13).$$

The codebook allocation SCP subproblems of $\mathbf{P_{Max\text{-}Min}}$ are given by

$$\mathbf{P}_{\mathbf{Max\text{-}Min}}^{(\mathbf{F})}(t) :$$

$$\max_{\mathbf{F}} \quad z + \gamma \left(\mathbf{F}^{(t-1)} \right) + \text{tr} \left[\nabla \gamma \left(\mathbf{F}^{(t-1)} \right)^T \left(\mathbf{F} - \mathbf{F}^{(t-1)} \right) \right]$$

$$\text{s.t.} \quad z - \sum_{k \in \mathcal{K}} \ln \left(|h_{k,j}|^2 f_{k,j} p_{k,j} + \sigma_n^2 + \sum_{i=1}^{j-1} |h_{k,i}|^2 f_{k,i} p_{k,i} \right) +$$

$$\theta_j \left(\mathbf{F}^{(t-1)} \right) + \text{tr} \left[\nabla \theta_j \left(\mathbf{F}^{(t-1)} \right)^T \left(\mathbf{F} - \mathbf{F}^{(t-1)} \right) \right] \leq 0$$

$$\forall \; j \in \mathcal{J} \quad\quad\quad (31)$$

$$(13), (11), (12), (17).$$

The complete alternating SCP algorithm to solve the Min-Max problem, $\mathbf{P_{Max\text{-}Min}}$, is described in Algorithm 2.

IV. NUMERICAL RESULTS

In this section the performance of the algorithms proposed in Section III is presented. Additionally, we compare our results with the three algorithms proposed in [16]: the fixed user order (FUO), opportunistic allocation (OA) and proportional fair (PF) algorithms. In these three algorithms, the users are sorted according to a different criteria, and the resources are allocated sequentially by picking the best available resource for the user in the sorted order. The FUO algorithm performs the allocation in a random order, while the OA algorithm sort the users according to their overall channel qualities, prior to the channel allocation. The PF algorithm takes into account the L past channel qualities when sorting the allocation order in order to improve fairness. In our evaluation we consider $L = 30$.

We consider a scenario where one BS is serving 6 users over 4 subcarriers in a circular cell of radius $R = 500$ m, and the users are uniformly distributed inside the cell. We consider a path loss exponent $\alpha = 4$. We consider a noise power density of -140 dBm/Hz and a bandwidth of 15 kHz. Also, we consider a normalized slow fading Rayleigh channel, such that the channel remains constant for the duration of each transmitted symbol. Furthermore, we simulate the algorithms

---

Algorithm 2: Maximization of the minimum rate

Variable Definition

1. $\mathbf{P}_n$ *is the power allocation matrix at the n-th alternating step.*
2. $\mathbf{F}_n$ *is the codebook allocation matrix at the n-th alternating step.*
3. $\mathbf{P}^{(t)}$ *is the power allocation matrix at the t-th SCP step.*
4. $\mathbf{F}^{(t)}$ *is the codebook allocation matrix at the t-th SCP step.*
5. $C^{(t)}$ *is the sum-rate at the t-th SCP step.*

Initialization

1. *Set the initial values for the power allocation matrix $\mathbf{P}_0$ randomly, and respecting the constraint (13), z_0 respecting the constraints (19), and, the codebook allocation matrix $\mathbf{F}_0$ randomly, and respecting the constraints (11), (12) and (17).*
2. *Set the convergence tolerance for the codebook allocation ϵ_F and for the power allocation ϵ_P in the alternating step.*
3. *Set the tolerance ϵ_1 for the convergence of the SCP sub problems of $P_{Max\text{-}Min}^{(F)}$, and ϵ_2 for $P_{Max\text{-}Min}^{(P)}$.*
4. $n \leftarrow 0$

Output

1. *Optimized power allocation $\mathbf{P}^*$.*
2. *Optimized codebook allocation $\mathbf{F}^*$.*

while $\|\mathbf{F}_n - \mathbf{F}_{n-1}\| > \epsilon_F$ *and* $\|\mathbf{P}_n - \mathbf{P}_{n-1}\| > \epsilon_P$ **do**

   $n \leftarrow n + 1;$

   $\mathbf{F}^{(0)} \leftarrow \mathbf{F}_{n-1};$

   $C^{(0)} \leftarrow \min \left(C_{\mathbf{SCMA}} \left(\mathbf{F}^{(0)}, \mathbf{P}_{n-1} \right) \right);$ (see (7))

   $t \leftarrow 0;$

   while $\left\| C^{(t)} - C^{(t-1)} \right\| > \epsilon_1$ **do**

     $t \leftarrow t + 1;$

     $\mathbf{F}^{(t)} \leftarrow \mathbf{P}_{\mathbf{Max\text{-}Min}}^{(\mathbf{F})}(t - 1)$ (see (20));

     $C^{(t)} \leftarrow \min \left(C_{\mathbf{SCMA}} \left(\mathbf{F}^{(t)}, \mathbf{P}_{n-1} \right) \right)$ (see (7));

   end

   $\mathbf{F}_n \leftarrow \mathbf{F}^{(t)};$

   $\mathbf{P}^{(0)} \leftarrow \mathbf{P}_{n-1};$

   $C^{(0)} \leftarrow \min \left(C_{\mathbf{SCMA}} \left(\mathbf{F}_n, \mathbf{P}^{(0)} \right) \right)$

   $t \leftarrow 0;$

   while $\left\| C^{(t)} - C^{(t-1)} \right\| > \epsilon_2$ **do**

     $t \leftarrow t + 1;$

     $\mathbf{P}^{(t)}, z^{(t)} \leftarrow \mathbf{P}_{\mathbf{Max\text{-}Min}}^{(\mathbf{P})}(t - 1);$ (see (18))

     $C^{(t)} \leftarrow \min \left(C_{\mathbf{SCMA}} \left(\mathbf{F}_n, \mathbf{P}^{(t)} \right) \right)$ (see (7));

   end

   $\mathbf{P}_n \leftarrow \mathbf{P}^{(t)};$

end

$\mathbf{P}^* \leftarrow \mathbf{P}_n;$

$\mathbf{F}^* \leftarrow \mathbf{F}_n;$

---

performance for a maximum transmit power per user varying between -4 dBm and 10 dBm. We evaluate the performances according to two metrics: the sum-rate and the Jain's fairness index. The former, is a measure of the overall achievable throughput of the network and the latter is a measure of the fairness of the resource allocation between the users based
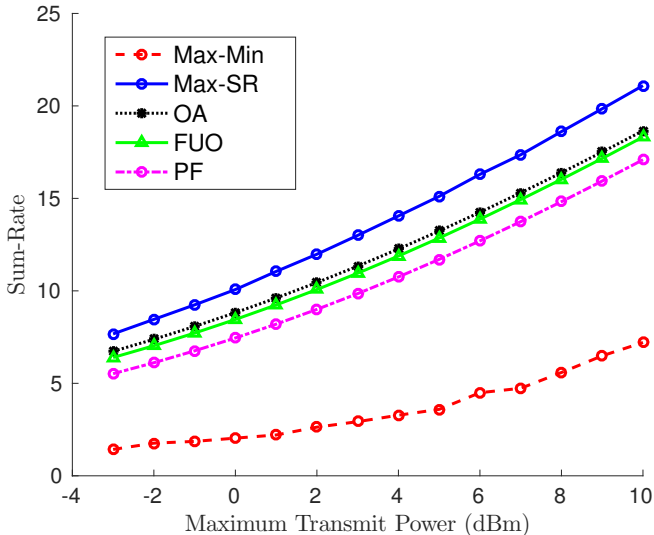
Fig. 6. Sum-rate comparison.

Fig. 7. Jain's fairness index comparison.

on their individual achievable throughputs. Let $\mathbf{c} \in \mathbb{R}^J$ be a vector, such that, its i-th coordinate, c_i, corresponds to the throughput of the i-th user. The Jain's fairness index for a given rate vector, $\mathbf{c} = [c_1, \cdots, c_J]^T$, is

$$J(\mathbf{c}) = \frac{\left(\sum\limits_{j=1}^{J} c_j\right)^2}{J \sum\limits_{j=1}^{J} c_j^2}. \tag{32}$$

This index varies from $\frac{1}{J}$, meaning no fairness, to 1, meaning perfect fairness. Furthermore, we consider a normalized Rayleigh fading channel, and the performances are averaged over several channel realizations.

The comparison of the sum-rate of the five algorithms is shown in Figure 6. The sum-rate of the Max-SR algorithm outperforms the other ones for the whole ranges of transmitted power evaluated. The three algorithms proposed in [16] present similar performance with the proportional fairness being slightly worse than the other two. Finally, the Max-Min algorithm proposed in this is greatly outperformed by the other ones. This result is expected since the Max-Min gives up on maximizing the sum-rate in favor of improving the fairness.

Figure 7 shows the Jain's fairness index achieved by the algorithms being considered. The Max-Min algorithm greatly outperforms the alternatives for the whole range of transmitted powers. Theoretically, the global solution to the original Max-Min problem should lead to perfect fairness, however, the algorithm proposed in this paper employs approximations to convexify the original problem and probably converges to local solutions of the original problem. Furthermore, with increasing maximum transmit power, the fairness of the Max-Min algorithm approaches one. The other algorithms achieve similar fairness performance, with the PF algorithm achieving slightly better fairness than the others.
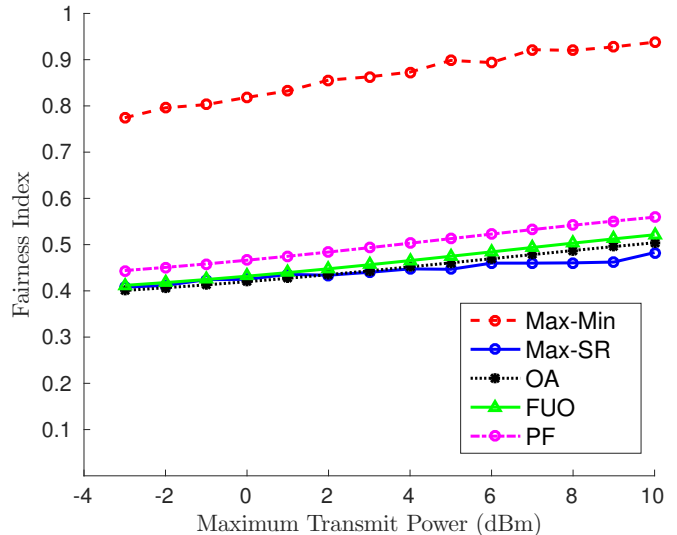
A. Convergence Analysis

In this section, the convergence of the alternating optimization iterations is investigated, based on simulations of the system. The stopping criteria for the procedure is based on the difference between successive iterations of the alternating optimization steps $\mathbf{F}_n$ and $\mathbf{P}_n$. When this difference falls below a threshold, ϵ_F and ϵ_P, respectively, the algorithm has converged. In order to illustrate the convergence, Algorithms 1 and 2 are simulated with $P_{\max}^{(j)} = 10$ dBm $\forall \; j \in \mathcal{J}$, with the same channel gains and users location, but with three randomly chosen initial conditions ($\mathbf{F}_0, \mathbf{P}_0$). As both algorithms are based on successively approximating the original problem, and each alternating step converges to a local optimum, the sum rate and fairness index value will be different for distinct initial conditions. Figure 8 shows the convergence of Algorithm 1. For all the initial conditions the algorithm converges after three alternating steps to similar values of sum-rate, However, there are significant variations on the fairness index depending on the initialization of the algorithm. In Figure 9, the convergence of Algorithm 2 is shown. The algorithm converges in three steps for all three initial conditions, but, it converges to different sum-rate and fairness values. Furthermore, despite the dependence on the initial conditions, Algorithm 1 consistently converges to a higher sum-rate than Algorithm 2, while Algorithm 2 converges to a higher fairness index, as expected.

V. CONCLUSIONS

In this paper, two joint channel and power allocation algorithms are proposed, the Max-SR and the Max-Min algorithms. The former aims for sum-rate maximization, while the latter aims for maximizing fairness. Alternating optimization and sequential convex programming are employed to obtain algorithms converging to locally optimal points of the original problems. The performance of the proposed algorithms is compared with other ones recently proposed in the literature.
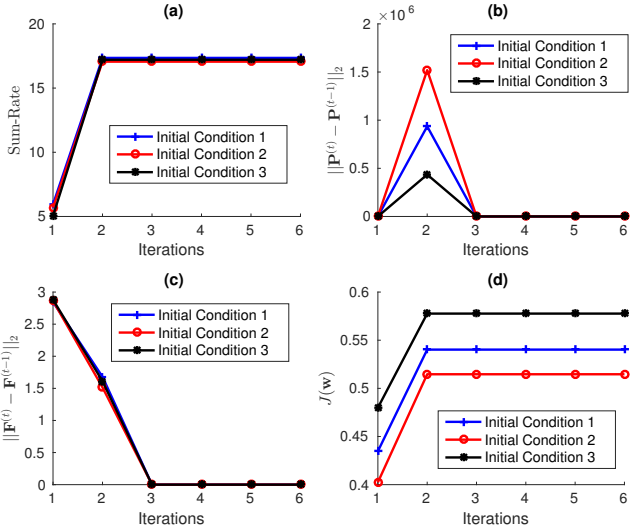
Fig. 8. Convergence of the $\mathbf{P_{Max\text{-}SR}}$, with maximum transmit power of 10 dbm, algorithm for three different initial conditions. In **(a)** and **(d)**, the values of the sum-rate and the Jain's fairness index, respectively, at each alternating iteration are shown. In **(b)** and **(c)**, the values of the stopping criteria of the alternating optimization algorithm at each alternating iteration is shown.
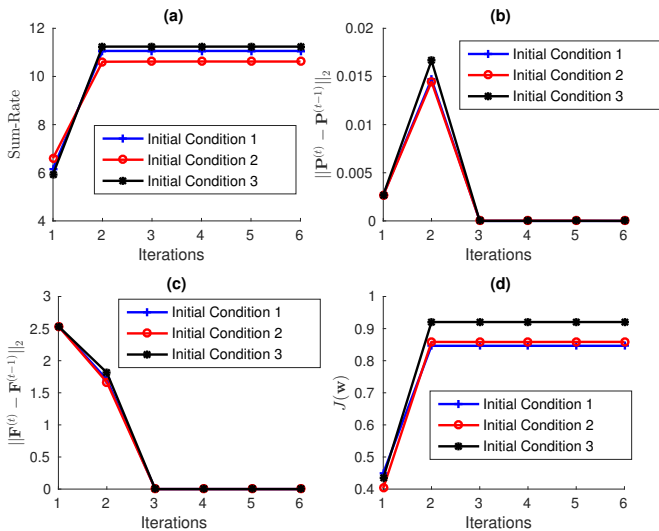

Fig. 9. Convergence of the $\mathbf{P_{Max\text{-}Min}}$ algorithm, with maximum transmit power of 10 dBm, for three different initial conditions. In **(a)** and **(d)**, the values of the sum-rate and the Jain's fairness index, respectively, at each alternating iteration are shown. In **(b)** and **(c)**, the values of the stopping criteria of the alternating optimization algorithm at each alternating iteration is shown.

The results show that the Max-SR algorithms has better performance on the sum-rate sense, while the Max-Min has better performance on the fairness sense. Finally, a numerical analysis of the convergence of the algorithms is presented, and a sensitivity of the performance to the initial condition is identified. The results show that the Max-SR consistently achieve a better sum-rate, while the Max-Min achieves better fairness, despite their sensitivity to the initial conditions. Furthermore, there is a tradeoff between the fairness and the

sum-rate, therefore, for future works, we intend to investigate new algorithms that can improve the fairness of the network without losing much on the sum-rate.

APPENDIX A
PROOF OF SCP CONVERGENCE

As $g(x, y) \geq g(x)$ all the points in the sequence $\hat{\mathbf{x}}_n$ are feasible solutions to the original problem. As x^n is a global solution to $\mathbf{P}(n - 1)$, x^n is a feasible point of $\mathbf{P}(n)$, so if $x^{n+1} \neq x^n$ we have that $f(x^{n+1}) > f(x^n)$. On the other hand, if $x^{n+1} = x^n = \bar{x}$ (i.e. $\hat{\mathbf{x}}_n$ has converged to $\bar{x}$), we have that the Lagrangian of $\mathbf{P}(n)$

$$
\begin{aligned}
\nabla_x \mathcal{L}(\bar{x}, \lambda) &= \nabla \tilde{f}(\bar{x}, \bar{x}) + \sum_{i=1}^{M} \lambda_i \nabla g_i(\bar{x}) + \\
&\qquad \sum_{i=M+1}^{N} \lambda_i \nabla \tilde{g}_i(\bar{x}, \bar{x}) \\
&= \nabla f(\bar{x}) + \sum_{i=1}^{N} \lambda_i \nabla g_i(\bar{x}) = 0. \quad (33)
\end{aligned}
$$

This completes the proof.

REFERENCES

[1] "Ericsson Mobility Report," https://www.ericsson.com/assets/local/mobility-report/documents/2015/ericsson-mobility-report-june-2015.pdf, Tech. Rep., 2015.

[2] T. Cover, "Broadcast channels," *IEEE Trans. Info. Theory*, vol. 18, no. 1, pp. 2–14, Jan 1972.

[3] R. Ahlswede, *Multi-way communication channels*. Akadmiai Kiad, 1973, pp. 23–51.

[4] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec 2014.

[5] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. New York, NY, USA: Cambridge University Press, 2005.

[6] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181–2195, Oct 2017.

[7] M. Moltafet, N. M. Yamchi, M. R. Javan, and P. Azmi, "Comparison study between PD-NOMA and SCMA," *IEEE Trans. on Veh. Tech.*, vol. 67, no. 2, pp. 1830–1834, Feb 2018.

[8] H. Nikopour and H. Baligh, "Sparse code multiple access," in *IEEE 24th Annu. Int. Symp. on Personal, Indoor, and Mobile Radio Commun. (PIMRC)*, Sept 2013, pp. 332–336.

[9] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, "SCMA codebook design," in *IEEE 80th Veh. Technol. Conf.*, Sept 2014, pp. 1–5.

[10] H. Nikopour, E. Yi, A. Bayesteh, K. Au, M. Hawryluck, H. Baligh, and J. Ma, "SCMA for downlink multiple access of 5G wireless networks," in *IEEE Global Commun. Conf.*, Dec 2014, pp. 3940–3945.

[11] T. Xue, L. Qiu, and X. Li, "Resource allocation for massive M2M communications in SCMA network," in *IEEE 84th Veh. Technol. Conf.*, Sept 2016, pp. 1–5.

[12] L. Luo, L. Li, and X. Su, "Optimization of resource allocation in relay assisted multi-user SCMA uplink network," in *Int. Conf. on Computing, Networking and Commun.*, Jan 2017, pp. 282–286.

[13] B. Di, L. Song, and Y. Li, "Radio resource allocation for uplink sparse code multiple access (SCMA) networks using matching game," in *IEEE Int. Conf. on Commun.*, May 2016, pp. 1–6.

[14] K. Au, L. Zhang, H. Nikopour, E. Yi, A. Bayesteh, U. Vilaipornsawai, J. Ma, and P. Zhu, "Uplink contention based SCMA for 5G radio access," in *IEEE Globecom Workshops*, Dec 2014, pp. 900–905.

[15] Z. Li, W. Chen, F. Wei, F. Wang, X. Xu, and Y. Chen, "Joint code-book assignment and power allocation for SCMA based on capacity with gaussian input," in *2016 IEEE/CIC International Conference on Communications in China (ICCC)*, July 2016, pp. 1–6.

[16] M. Dabiri and H. Saeedi, "Dynamic scma codebook assignment methods: A comparative study," *IEEE Commun. Letters*, vol. 22, no. 2, pp. 364–367, Feb 2018.

[17] J. Cui, P. Fan, X. Lei, Z. Ma, and Z. Ding, "Downlink power allocation in scma with finite-alphabet constraints," in *IEEE 85th Veh. Technol. Conf.*, June 2017, pp. 1–5.

[18] J. Liu, M. Sheng, L. Liu, Y. Shi, and J. Li, "Modeling and analysis of SCMA enhanced D2D and cellular hybrid network," *IEEE Trans. on Commun.*, vol. 65, no. 1, pp. 173–185, Jan 2017.

[19] G. D. Forney Jr. and L.-F. Wei, "Multidimensional Constellations-Part I : Introduction , Figures of Merit, and Generalized Cross Constellations," *IEEE J. Sel. Areas Commun.*, vol. I, no. 6, 1989.

[20] R. J. McEliece, D. J. C. MacKay, and J.-F. Cheng, "Turbo decoding as an instance of pearl's 'belief propagation' algorithm," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 2, pp. 140–152, Feb 1998.

[21] "Discussion on the feasibility of advanced MU-detector," 3GPP, Tech. Rep. TSG RAN WG1 Meeting 86, Aug 2016.

[22] J. Zou, H. Zhao, and W. Zhao, "Low-complexity interference cancellation receiver for sparse code multiple access," in *IEEE 6th Int. Symp. on Microwave, Antenna, Propagation, and EMC Technologies*, Oct 2015, pp. 277–282.

[23] F. Wei and W. Chen, "Low complexity iterative receiver design for sparse code multiple access," *IEEE Trans. on Commun.*, vol. 65, no. 2, pp. 621–634, Feb 2017.

[24] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.

[25] C. A. Floudas, *Deterministic Global Optimization: Theory, Methods and Applications*, ser. Nonconvex Optimization and Its Applications. Berlin, Heidelberg: Springer-Verlag, 2005, vol. 37.

[26] K. Shen and W. Yu, "Distributed pricing-based user association for downlink heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1100–1113, June 2014.

[27] H. H. M. Tam, H. D. Tuan, D. T. Ngo, T. Q. Duong, and H. V. Poor, "Joint load balancing and interference management for small-cell heterogeneous networks with limited backhaul capacity," *IEEE Trans. on Wireless Commun.*, vol. 16, no. 2, pp. 872–884, Feb 2017.

[28] Q. Kuang, J. Speidel, and H. Droste, "Joint base-station association, channel assignment, beamforming and power control in heterogeneous networks," in *IEEE 75th Veh. Technol. Conf.*, May 2012, pp. 1–5.

[29] K. Chitti, Q. Kuang, and J. Speidel, "Joint base station association and power allocation for uplink sum-rate maximization," in *IEEE 14th Workshop on Signal Process. Advances in Wireless Commun.*, June 2013, pp. 6–10.

[30] R. Sun, M. Hong, and Z. Q. Luo, "Joint downlink base station association and power control for max-min fairness: Computation and complexity," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1040–1054, June 2015.

[31] V. N. Ha and L. B. Le, "Distributed base station association and power control for heterogeneous cellular networks," *IEEE Trans. on Veh. Technol.*, vol. 63, no. 1, pp. 282–296, Jan 2014.

[32] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[33] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Oper. Res.*, vol. 26, no. 4, pp. 681–683, 1978.