Joint User Grouping, Scheduling, and Precoding for Multicast Energy Efficiency in Multigroup Multicast Systems

Ashok Bandi, Bhavani Shankar Mysore R, Symeon Chatzinotas and Björn Ottersten
Interdisciplinary Centre for Security, Reliability and Trust (SnT), the University of Luxembourg, Luxembourg.
Email: {ashok.bandi, bhavani.shankar, symeon.chatzinotas, bjorn.ottersten}@uni.lu

Abstract—This paper studies the joint design of user grouping, scheduling (or admission control) and precoding to optimize energy efficiency (EE) for multigroup multicast scenarios in single-cell multiuser MISO downlink channels. Noticing that the existing definition of EE fails to account for group sizes, a new metric called multicast energy efficiency (MEE) is proposed. In this context, the joint design is considered for the maximization of MEE, EE, and scheduled users. Firstly, with the help of binary variables (associated with grouping and scheduling) the joint design problem is formulated as a mixed-Boolean fractional programming problem such that it facilitates the joint update of grouping, scheduling and precoding variables. Further, several novel optimization formulations are proposed to reveal the hidden difference of convex/ concave structure in the objective and associated constraints. Thereafter, we propose a convex-concave procedure framework based iterative algorithm for each optimization criteria where grouping, scheduling, and precoding variables are updated jointly in each iteration. Finally, we compare the performance of the three design criteria concerning three performance metrics namely MEE, EE, and scheduled users through Monte-Carlo simulations. These simulations establish the need for MEE and the improvement from the system optimization.

Index Terms—User grouping, scheduling, Precoding, Multicasting, Energy efficiency and difference-of-concave programming.

I. INTRODUCTION

The mobile data traffic is exploding unprecedentedly due to the exponential increase in mobile devices and their demand for throughput hungry service/ applications [3]. This has led to the adaptation of full-spectrum reuse and multi-antenna technologies, which result in significantly improved spectral efficiency (SE). On the other hand, the demand for green communications necessitates achieving these high SEs with limited energy [4]. In this regard, energy efficiency (EE), which measures the performance in throughput/Watts, becomes a key factor to be considered in the next-generation wireless networks [5], [6]. Notice that the power minimization or energy minimization is also referred to as energy efficiency maximization in the literature. The previously mentioned EE that measures the performance in throughput/Watts is the focus of this paper.

On the other hand, in some scenarios like live streaming of popular events, multiple users are interested in the same data. Realizing that multicasting such information to groups of users leads to better utilization of the resources, physical layer multigroup multicasting (MGMC) has been proposed in [7], [8]. Noticing the significant improvement in EE, multicasting has been adopted into 3GPP standards [9]. However, the following challenges need to be addressed to fully leverage the gains of MGMC:

- Inter-group interference: The co-channel users in different groups generate interference across the groups; this is referred to as inter-group interference (IGI). A study of IGI is essential as it fundamentally limits the minimum rate of the groups that can be achieved [10] and, hence, the total throughput of the network. In this context, *user grouping* is a pivotal factor to be considered since it dominantly influences IGI [11], [12].

- Infeasibility: In a real scenario, each user needs to be served with a certain quality-of-service (QoS); failing to meet the QoS leads to retransmissions that significantly decrease the EE of the network. The severe IGI and/ or poor channel gains may thwart some users from meeting their QoS [10]. On the contrary, even in the cases with lower IGI, limited power may restrain the users from meeting their QoS [7]. Due to a combination of these three factors, the system may fail to satisfy the QoS requirements of all the users in all groups. This scenario is referred to as the infeasibility of the MGMC in the literature [7], [10]. The infeasibility of MGMC is crucial to the design and is, therefore, typically addressed by *user scheduling* (also referred to as admission control in the literature) [12], [13].

A. Joint user grouping, group scheduling and user scheduling for message-based MGMC systems

In this work, similar to [12], [14], a message based user grouping and scheduling are considered. In the message based MGMC model, each group is associated with a different message, and a user may be interested in multiple messages. Despite the user's interest in multiple message the user can subscribe to utmost one group per transmit slot, and, thus requiring user grouping. Unlike [14] and similar to [12], a limited antenna BS system is considered with the number

of groups being larger than the number of antennas at BS. Therefore, in a given transmit slot, only a few groups (equal to the number of antennas of BS) are scheduled; this is referred to as group scheduling. Further, users that fail to satisfy the QoS requirements of all the interested groups are simply excluded from the grouping; this is referred to as user scheduling. Therefore, the considered model requires the design of user grouping, group scheduling, and user scheduling.

User grouping, group scheduling, and user scheduling are inter-related. To see this, user grouping decides the achievable minimum signal-to-interference and noise ratio (SINR) of the groups (or IGI) which influences the group scheduling and user scheduling. Further, omitting or adding a user (i.e., user scheduling) to a group changes the IGI, thereby influencing group scheduling. Similarly, user scheduling in a group might necessitate the re-grouping of users (i.e., user grouping); this affects IGI and group scheduling. Furthermore, IGI is a function of precoding [11]. Therefore, the optimal performance requires the *joint design of user grouping, group scheduling, user scheduling, and precoding; this is compactly referred to as joint design in this paper*.

B. EE in the context of the joint design of user grouping, scheduling, and precoding

In this work, for the reasons mentioned earlier, EE is considered as the measure of the system performance. All the existing works on EE maximization in MGMC systems [8], [15]–[18], presume a particular user grouping and scheduling, therefore, the EE for MGMC systems is defined as the ratio of the sum of minimum throughput within each group and the total consumed energy. Notice that the existing EE definition accounts for only the minimum rate of a group ignoring the number of users in the groups (group sizes). However, in the context of user grouping and scheduling, group sizes need to be accounted for, in addition to their minimum rates. To comprehend the necessity, consider two groups with an equal minimum rate and a large difference in group sizes. According to the existing EE definition, the group with few users could be scheduled as the EE maximization is not biased to schedule a group with the larger size. However, from the network operator perspective, scheduling the group with more users leads to efficient utilization of the resources. Moreover, the event of scheduling the group with few users is likely for EE maximization since it usually consumes less energy. However, if a large number of users can be served with a slight increase in energy, scheduling such a larger group improves the efficiency in the utilization of resources. The existing frameworks can not handle these scenarios as the number of users is not included in the EE definition. Noticing the drawbacks of existing EE definition, in this work, a new metric called multicast energy efficiency (MEE) is proposed to account for the group sizes along with the minimum rates. In contrast to EE, MEE considers, as numerator, the product of minimum throughput within a group times number of users in that group. Realizing the importance of MEE, in this work, we consider the joint design of user grouping, scheduling, and precoding for MGMC systems subject to grouping, scheduling,

quality-of-service (QoS) and total power constraints for the maximization of three design criteria: MEE, EE and scheduled users. In this context, related works in the literature and contributions/novelty of the paper are summarized in the sequel of this section.

C. Related works

Energy efficiency for MGMC systems: The EE maximization problem for MGMC systems was first addressed in the context of coordinated beamforming for multicell networks [15]. By definition, EE maximization belongs to a class of fractional programming. The authors in [15] used Dinkelbach's method to transform this fractional program to a subtractive non-linear form and further solved the problem with an iterative algorithm where precoding and power vectors are updated alternatively in each iteration. Later, this work is extended to the case of imperfect channel state information in [16]. Dinkelbach's method based transformation works efficiently if the denominator is a simple linear objective and numerator is convex; otherwise it leads in a multi-level parametric iterative algorithm [19]. In [19], the authors optimized the EE for MGMC in multicell networks considering the rate-dependent processing power, and using successive convex approximation (SCA) to transform fractional EE maximization problem as convex-concave programming and thereby solved iteratively wherein each iterate the problem is convexified using Charnes-Cooper transformation (CCT). In [17], EE maximization in a large antenna system with antenna selection is solved using SCA based CCT. Further, the authors in [17] addressed the Boolean nature stemming from antenna selection by continuous relaxation and followed by thresholding. In [18], EE maximization for non-orthogonal layered-division multiplexing based joint multicast and the unicast system is considered. A pseudoconvex approach based parallel solution is developed for EE maximization in MIMO interference channels in [20]. EE user scheduling and power control is considered for multi-cell OFDMA networks for a unicast scenario in [21]. Moreover, precoding is not considered in [21]. The methodologies used in all these works are either SCA based CCT [15], [17], [19] or SCA based Dinkelbach's method [18], [20], [22]. Unlike the majority of the aforementioned works that restrict the rate-dependent processing power to be a convex function of rate, a non-convex power consumption model is considered in this work. Therefore, unlike EE maximization considered in the literature, MEE maximization considered in this work belongs to mixed-Boolean fractional programming where the numerator is mixed-Boolean non-convex and denominator is also non-convex. Hence, the SCA based CCT can not be applied and Dinkelbach's method yields parametric multilevel iterative algorithms [20]. Moreover, the Boolean nature stemming from the user grouping and scheduling is different to the antenna selection problems (see [17] and references therein) and, hence, problem formulation and solution methodologies used in the antenna selection literature can not be employed.

User grouping, scheduling, and precoding for EE maximization in MGMC systems: In this work, we consider the joint design of user grouping, group scheduling, user scheduling, and

precoding for MEE maximization in message-based MGMC systems. Joint design of admission control and beamforming for MGMC systems was initially addressed in [13] for the power minimization problem. The authors in [13] addressed the admission control using binary variables and transformed the resulting mixed-Boolean non-linear problem (MINLP) into a convex problem using semidefinite programming (SDP) transformations; SDR based precoding is likely to include high-rank matrices for MGMC systems [23]; hence, the solutions may become infeasible to the original problems. Later, the design of user grouping and precoding without admission control is considered in [11] for satellite systems. However, the authors in [11] adopted the decoupled approach of heuristic user grouping followed by a semidefinite relaxation (SDR) based precoding. In [14], the authors considered the joint user grouping and beamforming without user scheduling for massive multiple-input multiple-output (MIMO) systems and proved that arbitrary user grouping is asymptotically optimal for max-min fairness criteria. However, the arbitrary grouping is not optimal for other design criteria and also not optimal for max-min criteria in the finite BS antenna system. The authors in [12] addressed the joint design for power minimization by decoupling the user grouping, scheduling and precoding. Moreover, the underlying precoding problem in [11]–[13] is solved by SDR transformation which is likely to include high-rank matrices for MGMC systems [23]; hence, the solutions may become infeasible to the original problems. In [24], joint adaptive user grouping and beamforming is considered for the MGMC scenario in a massive MIMO system. The authors in [24] adapted iterative approach wherein each iterate user grouping and beamforming are solved separately by decoupling the two problems. In [1], user scheduling and frame-based precoding is considered for MGMC scenario in satellite system for the maximization of sum-rate. In [2], user scheduling and precoding is considered for multicast spectral efficiency maximization. Notice that user grouping and group scheduling are not considered in [1], [2]. Moreover, the EE or MEE maximization problem in the context of user grouping, scheduling, and precoding is not considered in the literature.

In our previous work [25], we considered the joint design of scheduling and precoding for the unicast scenario to optimize sum rate, Max-min SINR, and network power. In [25], scheduling is addressed by bounding the power of the precoder with the help of a binary variable. However, in the MGMC system, each precoder is associated with a group of users, hence, the same method can not be employed. Moreover, the MEE or EE maximization in the context of user grouping, scheduling, and precoding belongs to mixed-Boolean fractional programming which is not dealt in [25], and its extension to proposed system model is not clear. Furthermore, the proposed MEE belongs to Mixed-Boolean fractional programming problems with a mixed-Boolean non-convex objective in the numerator and a non-convex objective in the denominator. Therefore, the MEE maximization problem considered in this work is significantly different from [25] in terms of the system model, performance metric of optimization, problem formulation, and the nature of the optimization problem. Hence, the solution in [25] can not be applied directly here. The MEE maximization problem considered for the joint design in this work is highly complex as it inherits the complications of user grouping, group scheduling, user scheduling, and EE problems, and poses additional challenges

D. Contributions

Below we summarize the contribution on the joint design of user grouping, scheduling, and precoding for the MGMC system to maximize the MEE and EE as follows:

- Noticing that the existing EE definition accounts only for the minimum rate of groups ignoring group size, a new metric called MEE is proposed to account for the group sizes along with the minimum rate of the groups in the message-based MGMC systems [12]. Unlike the existing works e.g., [12], [19], [25], this results in a new mixed Boolean-fractional objective function that is not considered in the literature. This fundamental departure precludes the direct application existing results in user grouping, scheduling, and EE designs.
- Further, unlike existing models which restrict rate-dependent processing power for energy calculations to be a convex function [15], [17]–[19], [21], our work is applicable to a general class of non-convex functions (in rate) which admit difference-of-convex decomposition.
- Inspired by the work in [25], user grouping, group scheduling, and user scheduling are addressed with the help of binary variables. Further, unlike [25], the joint design problem is formulated as mixed Boolean-fractional programming that facilitates the joint update of the variable unlike the decoupled approaches [11], [12], [21], [23]. Since the Boolean and continuous variables are coupled the existing SCA based CCT [19] can not be applied. On the contrary, Dinkelbach [15] method results in the parameter-dependent multilevel iterative algorithm.
- The resulting mixed-Boolean fractional formulations are non-convex and NP-hard. Noticing that exploiting the structure in non-convexity helps in obtaining high-quality solutions, the hidden difference-of-convex structure is revealed with the help of novel reformulations. Further, Boolean nature is handled with appropriate relaxation and penalization. These reformulations render the joint design as a DC problem, a fact hitherto not considered.
- Finally, towards obtaining a low-complexity stationary point, within the framework of the convex-concave procedure (CCP) [26] (which is a special case of SCA [27]), an iterative algorithm is proposed to solve the resulting DC problem wherein a convex problem is solved in each iterate. Thereafter, a simple low-complexity non-iterative procedure which solves a linear programming problem to obtain a feasible initial point, which inherently establishes convergence of the proposed algorithms to a stationary point [28], is proposed.
- The performance of the proposed algorithms affecting the three design aspects (i.e., MEE, EE, and number of scheduled users), and their typical quick convergence behavior(which confirms the low-complexity nature) are numerically evaluated through Monte-Carlo simulations.

The sequel is organized as follows. Section II presents MGMC system. Further, the joint design for MEE problem in Section III, EE problem in Section IV and SUM problem in Section IV-B. Section V presents simulations and Section VI concludes the work.

Notations: Lower or upper case letters represent scalars, lower case boldface letters represent vectors, and upper case boldface letters represent matrices. $\|\cdot\|$ represents the Euclidean norm, $|\cdot|$ represents the cardinality of a set or the magnitude of a scalar, $(\cdot)^H$ represents Hermitian transpose, $(\cdot)^T$ represents transpose, $\binom{a}{b}$ represents a choose b, and $\mathbb{R}\{\}$ represents real operation, $\mathbb{E}\{\}$ represents expectation operator and ∇ represents the gradient.

II. SYSTEM MODEL

A. Message based user grouping and scheduling

We consider the downlink scenario of a single cell multiuser MISO system with M transmit base station (BS) antennas and $N (\geq M)$ users each equipped with a single receive antenna. In this work, similar to [12], [14], message-based user grouping, and scheduling is considered. In this context, it is assumed that each user is interested in at least one message and a user may be interested in multiple messages. Despite the users interest in multiple messages, we assume that the user has capability to process one message in a slot (possibly causes include single antenna, simple decoding strategy). Hence, the user can subscribe to utmost one message in any given transmit slot. Depending on the network conditions, he could subscribe to other messages in subsequent time slots. This per time-slot constraint is simply referred to as user grouping constraint (UGC) where users subscribed to a message is referred to as a group. Naturally, the number of groups, say G, is equal to the number of messages. Letting $\mathcal{S}_i$ to be the set of users belonging to message (group) i and ϕ to be the empty set, the UGC is formulated as $\mathcal{S}_i \cap \mathcal{S}_j = \phi$, for $i \neq j$. Further, to establish the relevance of the design to the real scenarios, a certain QoS requirement (typically depending on the type of service/application) on the messages is assumed. UGC also captures the worst-case scenario of a user failing to meet any QoS requirement associated with any of the interesting messages: hence, the user is simply not scheduled in the current slot. Therefore, UGC naturally leads to $\sum_{i=1}^{G}|\mathcal{S}_i| \leq N$. Further, it is assumed that $G \geq M$, hence, scheduling of exactly M groups out of G is considered. This constraint is simply referred to as group scheduling constraint (GSC).

User channels are assumed to constant and perfectly known. The noise at all users is assumed to be independent and characterized as additive white complex Gaussian with zero mean and variance σ^2. Furthermore, total transmit power at the BS is limited to P_T for each transmission. Finally, the BS is assumed to transmit independent data to different groups with $\mathbb{E}\{|x_i|^2\} = 1, \forall i$, where x_i is the message associated with group i. The intra-group interference i.e., the interference among the users within a group does not arise in the considered system model. All the users in a group receive the same message, and, hence, the users within a group do not cause interference to each other [7]. Let $\mathbf{w}_i \in \mathbb{C}^{M \times 1}$ be the precoding vector with group i and $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_G]$, $\mathbf{h}_i \in \mathbb{C}^{M \times 1}$ be the downlink channel of user i, and $\gamma_{ij} = \dfrac{|\mathbf{h}_i^H \mathbf{w}_j|^2}{\sum_{l \neq j}|\mathbf{h}_i^H \mathbf{w}_l|^2 + \sigma^2}$ be the SINR of user i belonging to group j.

B. Power consumption model, Energy efficiency, and Multicast energy efficiency

Power consumption model: In this work, we adopt the power consumption model proposed in [19]. Let B be the bandwidth of the channel and $r_j = B \log_2 \left(1 + \min_{i \in \mathcal{S}_j} \gamma_{ij}\right)$ be the minimum rate of group j. Notice that all the users in a group receive exactly the same message associated with the group. Therefore, the transmission rate of the message j to group j at the BS is simply the minimum rate of the group i.e., r_j. With the notations defined, the power consumption at the BS is defined as

$$g(\mathbf{W}, \boldsymbol{r}) \triangleq P_0 + \sum_{j=1}^{G} \left(\frac{1}{\rho} \|\mathbf{w}_i\|^2 + \mu p\left(r_j\right)\right),$$

where $\boldsymbol{r} = [r_1, \ldots, r_G]$, P_0 is the static power spent by the cooling systems, power supplies etc., $\rho < 1$ is the power amplifier efficiency, and $\mu \geq 0$ is a constant accounting for coding and decoding power loss, and $p(r_j) = p_1\left(r_j\right) - p_2\left(r_j\right)$ is a differentiable non-negative difference-of-convex function of r_j reflecting the rate-dependent processing power of group j with $p(0) = 0$, and p_1 and p_2 being any convex functions. Notice that unlike previous works e.g. [17], [19], [21] where $p(r_j)$ is assumed to be a convex functions, the considered model $g(\mathbf{W}, \boldsymbol{r})$ represents relatively broader class of rate-dependent power consumption models.

Energy efficiency: EE for MGMC systems is typically defined as a ratio of the throughput of the network to the energy consumed at the BS in the literature. Letting $\mathcal{T}$ to be the set of scheduled groups and B be the bandwidth of the channel, the EE is defined as (1).

$$\text{EE} \triangleq \frac{\sum_{j \in \mathcal{T}} B \log \left(1 + \min_{i \in \mathcal{S}_j} \gamma_{ij}\right)}{g(\mathbf{W}, \boldsymbol{r})}. \tag{1}$$

The numerator of the EE in (1) models the network's multicast throughput as the sum of the minimum throughput of all groups. So, this definition only accounts for the minimum throughput of a group ignoring its size.

Multicast energy efficiency : In the context of user grouping and scheduling for MGMC systems, the standard EE metric needs to be redefined to account for the size of the group. To understand this, consider a scenario of scheduling a group between two groups having the same minimum throughput, consuming the same energy and large difference in group sizes. The EE criterion does not discriminate between two groups. However, scheduling a group with a large number of users leads to better utilization of resources. So, to account for the number of users being served in each group along with its minimum rate, we propose a new metric called MEE for the MGMC systems. With the help of defined notations, MEE is formally defined as,

$$\text{MEE} \triangleq \frac{\sum_{j \in \mathcal{T}} \left(\Psi_j |\mathcal{S}_j| B \log \left(1 + \min_{i \in \mathcal{S}_j} \gamma_{ij}\right)\right)}{g(\mathbf{W}, \boldsymbol{r})}. \tag{2}$$

where Ψ_j is the weight associated with group j. The weights i.e., Ψ_js are introduced in MEE to address the fairness among the groups. For example, by choosing Ψ_1 to be relatively much larger than $\{\Psi_i\}_{j=2}^{G}$ scheduling of group 1 can be prioritized.

Interpretation of MEE as total received bits/Joule: From the physical layer transmission perspective, the network throughput (number of transmitted bits per second) in MGMC systems is same as unicast systems. In unicast scenario, the transmitted information is received by only one user. However, in MGMC scenario, the information transmitted to group j is received by $|\mathcal{S}_j|$ users. Hence, from the network operator perspective, throughput of group j in this multicast scenario is $|\mathcal{S}_j|B \log \left(1 + \min_{i \in \mathcal{S}_j} \gamma_{ij}\right)$ received bits per second. Motivated by this, the numerator of equation (2) i.e., $\sum_{j \in \mathcal{T}} \left(|\mathcal{S}_j|B \log \left(1 + \min_{i \in \mathcal{S}_j} \gamma_{ij}\right)\right)$ reflects the combined multicast throughput of all the groups i.e., network throughput, henceforth, simply referred to in this work as multicast throughput. Similarly, MEE defined in (2) reflects MEE for MGMC systems. Thus, MEE can be seen as number of received bits for one joule of transmitted energy.

III. MULTICAST ENERGY EFFICIENCY

In this section, at first, the joint design of user grouping, scheduling, and precoding is mathematically formulated to maximize the MEE subject to appropriate constraints on the number of groups, users per group, number of scheduled groups, power, and QoS constraints. This problem is simply referred to as the MEE problem. Further, with the help of useful relaxations and reformulations, the MINLP NP-hard MEE problem is transformed as a DC programming problem. Finally, within the framework of CCP, an iterative algorithm is proposed which guarantees to attain a stationary point of the original problem.

A. Problem formulation: MEE

The EE maximization problem, with the notations defined, in the context of user grouping, scheduling and precoding for the MGMC scenario in Section II is formulated as,

$$\mathcal{P}_1^{\text{MEE}} : \max_{\{\mathbf{w}_j, \mathcal{S}_j\}_{j=1}^{G}} \frac{\sum_{j=1}^{G} \left(\Psi_j |\mathcal{S}_j| B \log \left(1 + \min_{i \in \mathcal{S}_j} \gamma_{ij}\right)\right)}{g(\mathbf{W}, \boldsymbol{r})}$$

(3)

s.t. $C_1 : \mathcal{S}_i \cap \mathcal{S}_j = \phi, i \neq j, \forall i, \forall j,$

$\quad C_2 : \left\| [|\mathcal{S}_1|, \ldots, |\mathcal{S}_G|] \right\|_0 = M,$

$\quad C_3 : \log \left(1 + \gamma_{ij}\right) \geq \epsilon_j, \ i \in \mathcal{S}_j, \ \forall j;$

$\quad C_4 : \sum_{j=1}^{G} \left\| \mathbf{w}_j \right\|^2 \leq P_T,$

$\quad C_5 : \mathcal{S}_i \subset \{1, \ldots, N\}, \forall i,$

where ϵ_j is the QoS requirement of group j, $\forall i$ refers to $i \in \{1, \ldots, N\}$ and $\forall j$ refers to $j \in \{1, \ldots, G\}$

Remarks:

- Constraint C_1 is the UGC; constrains a user to be a member of at most one group.

- Constraint C_2 is the GSC; it ensures the design to schedule exactly M groups.
- Constraint C_3 is the QoS constraint; it enforces the scheduled users in each group to satisfy the corresponding minimum rate requirement associated with the group. This enables the flexibility to support different rates on different groups. Hereafter, the constraint C_3 is simply referred to as QoS constraint. Moreover, the constraint C_3 together with C_1 implicitly ensure the USC. In fact, user who do not satisfy the QoS constraint of the group $\mathcal{S}_j$ are not scheduled (i.e., USC).
- Constraint C_4 is the total power constraint (TPC); precludes the design from consuming the power in excess of available power i.e., P_T.
- Constraint C_5 ensures that the group $\mathcal{S}_i$ is a subset of superset containing all the users i.e., $\{1, \ldots, N\}$.

Notice that the problem $\mathcal{P}_1^{\text{MEE}}$ is a joint optimization of precoders, user grouping, group scheduling, and user scheduling. While the user grouping and group scheduling are expressed explicitly, the user scheduling is ensured implicitly through the constraints C_1 and C_3.

Necessity of low-complexity algorithms for joint design: The problem $\mathcal{P}_1^{\text{MEE}}$ is combinatorial due to constraints C_1 and C_2. Hence, obtaining the optimal solution to $\mathcal{P}_1^{\text{MEE}}$ requires an exhaustive search-based user grouping and scheduling. To understand the complexity of the exhaustive search methods, assume that each user is interested in only one message. Further, let N_i be the number of users in the group i. Let $\mathcal{T}_i$ be the all possible scheduling subsets of $\mathcal{S}_i$, so the number of sets in $\mathcal{T}_i$ is $\sum_{j=0}^{N_i} \binom{N_i}{j}$ for $i \in \{1, \ldots, G\}$. So, the exhaustive search needs to be performed over the Cartesian product of sets $\mathcal{T}_i$s i.e., $\times_{i=1}^{G} (\mathcal{T}_i)$. It is easy to see that the exhaustive search algorithms quickly become impractical due to exponential complexity. This case merely a simple case of the problem considered in $\mathcal{P}_1^{\text{MEE}}$. Additionally, for each scheduled combination, the corresponding precoding problems in $\mathcal{P}_1^{\text{MEE}}$ need to be solved. Moreover, these precoding problems are generally not only NP-hard but also non-convex [7]. Thus, in the sequel, we focus on developing low-complexity algorithms that are guaranteed to obtain a stationary point of the NP-hard and non-convex problem $\mathcal{P}_1^{\text{MEE}}$.

B. A mixed Boolean difference of concave formulation: MEE

In this section, firstly, avoiding the set notation by using binary variables the problem $\mathcal{P}_1^{\text{MEE}}$ is equivalently reformulated as an MINLP problem without the set notations. Further, with the help of a minimal number of slack variables and novel reformulations, the resulting MINLP problem is transformed as a difference-of-concave (DC) problem subject to binary constraints.

Towards transforming the MEE problem in $\mathcal{P}_1^{\text{MEE}}$ as DC a problem, let η_{ij} be the binary variable indicating the membership of user $i \in \{1, \ldots, N\}$ in group $j \in \{1, \ldots, G\}$. In other words, $\eta_{ij} = 1$ indicates that user i is a member of the group j and not a member otherwise. Since a user may not be interested in some groups, the η_{ij}s corresponding to these groups is fixed beforehand to zero. Hence, only a subset of the

entries in $\boldsymbol{\eta}_i$ are the variables of the optimization. However, for the ease of notation, without the loss of generality, henceforth, we assume that each user is interested in all the groups. In other words, all the entries in $\boldsymbol{\eta}_i$ become variables of optimization. It is easy to see that this is only a generalization to the aforementioned case. Hence, a solution to this generalized problem is a solution to the aforementioned problem.

Letting Θ_j and ζ_j be the slack variables associated with minimum rate of group j respectively, and α_{ij} be the slack variable associated with SINR of user i of group j, the problem $\mathcal{P}_1^{\mathrm{MEE}}$ is equivalently reformulated as,

$$\mathcal{P}_2^{\mathrm{MEE}} : \max_{\mathbf{W},\boldsymbol{\Theta},\boldsymbol{\eta},\boldsymbol{\alpha},\boldsymbol{\zeta}} \frac{\sum_{j=1}^M \left(\sum_{i=1}^{N_i}\eta_{ij}\right) B\Psi_j\Theta_j}{g(\mathbf{W},\boldsymbol{r})} \quad (4)$$

$$\text{s.t. } C_1 : \eta_{ij} \in \{0,1\}, \ \forall i, \ \forall j,$$

$$C_2 : \sum_{j=1}^G \eta_{ij} \leq 1, \forall i,$$

$$C_3 : \left\| \left[\sum_{i=1}^N \eta_{i1}, \dots, \sum_{i=1}^N \eta_{iG}\right] \right\|_0 = M,$$

$$C_4 : 1 + \gamma_{ij} \geq \alpha_{ij}, \ \forall i, \ \forall j,$$

$$C_5 : \log \alpha_{ij} \geq \eta_{ij}\Theta_j, \ \forall i, \ \forall j,$$

$$C_6 : \Theta_j \geq \epsilon_j, \ \forall j,$$

$$C_7 : \sum_{i=1}^M \|\mathbf{w}_i\|_2^2 \leq P_T,$$

where $\boldsymbol{\Theta} = [\Theta_1,\dots,\Theta_M]$, $\boldsymbol{\eta} = [\boldsymbol{\eta}_1,\dots,\boldsymbol{\eta}_G]$, $\boldsymbol{\eta}_i = [\eta_{i1},\dots,\eta_{iG}]$, $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_G]$, $\boldsymbol{\alpha}_i = [\alpha_{i1},\dots,\alpha_{iG}]$, and $g(\mathbf{W},\boldsymbol{r}) \triangleq P_0 + \sum_{j=1}^G \left(\frac{1}{\rho}\|\mathbf{w}_i\|^2 + \mu \left\|\sum_{i=1}^N \eta_{ij}\right\|_0 p\left(\Theta_j\right)\right)$
Remarks:

- Constraints C_1 and C_2 in $\mathcal{P}_2^{\mathrm{MEE}}$ ensures the UGC. The constraint C_3 is the equivalent reformulation of GSC constraint C_2 in $\mathcal{P}_1^{\mathrm{MEE}}$.
- For all the users that are not subscribed to group j (i.e., users with $\eta_{ij} = 0$), the constraint C_5 implies $\log \alpha_{ij} \geq 0$ which is satisfied by the definition of rate. On the contrary, for all the users subscribed to group j (i.e., users with $\eta_{ij} = 1$) constraint C_4 implies $\log \alpha_{ij} \geq \Theta_j$. Hence, Θ_j provides the lower bound for the minimum rate of the group. Moreover, at the optimal solution of $\mathcal{P}_2^{\mathrm{MEE}}$, Θ_j is equal to the minimum rate of group j i.e., $\Theta_j = \min_{j\in\mathcal{S}_i}\log\alpha_{ij}$.
- In the objective of $\mathcal{P}_2^{\mathrm{MEE}}$, the term $\sum_{i=1}^{N_i}\eta_{ij}$ is equivalent to $|\mathcal{S}_i|$. Since at the optimal solution $\Theta_j = \min_{j\in\mathcal{S}_i}\log\alpha_{ij}$, the objective in $\mathcal{P}_2^{\mathrm{MEE}}$ is equivalent to the objective in $\mathcal{P}_1^{\mathrm{MEE}}$.

Transforming problem $\mathcal{P}_1^{\mathrm{MEE}}$ as an equivalent mixed-Boolean non-linear programming with the help of binary variables $\boldsymbol{\eta}$, $\boldsymbol{\delta}$s that facilitates the joint update of $\boldsymbol{\eta}$, $\boldsymbol{\delta}$ and $\mathbf{W}$. Particularly, formulation of the constraints C_1, C_2 and C_5 in problem $\mathcal{P}_1^{\mathrm{MEE}}$ are unlike any constraints in [25], [28]. Moreover, the equivalent reformulation of C_1, C_2 and C_5 using binary constraints and the associated QoS constraint C_3

as shown in problem $\mathcal{P}_2^{\mathrm{MEE}}$ are crucial to getting a joint update of all the variables.

Notice that the problem $\mathcal{P}_2^{\mathrm{MEE}}$ is significantly different and much more complex than problems dealt in references [12], [14], [15], [17], [19], [20], [25]. The MEE objective in $\mathcal{P}_2^{\mathrm{MEE}}$ is unlike any EE objective in the literature (see [15], [17], [19], [20] and references therein). The power consumption model $g(\mathbf{W},\boldsymbol{r})$ and multicast throughput i.e., $\sum_{j=1}^M \left(\sum_{i=1}^{N_i}\eta_{ij}\right)B\Theta_j$ considered in this work are non-convex and multicast throughput is a function of binary variables. Since the objective $\mathcal{P}_2^{\mathrm{MEE}}$ is a function of both coupled binary and continuous variables and that $g(\mathbf{W},\boldsymbol{r})$ is also non-convex the well-known SCA based CCT [17] can not applied $\mathcal{P}_2^{\mathrm{MEE}}$. Further, Dinkelbach methods [15] results in a parametric multi level iterative algorithm. Further, problem $\mathcal{P}_2^{\mathrm{MEE}}$ differs from [25]where the binary variables are only associated with precoding and SINR terms. The transformation to deal with the MEE objective, constraint C_3 and C_8 are not dealt in [25]. The problem $\mathcal{P}_2^{\mathrm{MEE}}$ inherit inherits the complexities associated with both the EE and user grouping/ scheduling problems, thereby rendering it significantly challenging than the individual problems (solved in literature).

The reformulation given in $\mathcal{P}_2^{\mathrm{MEE}}$ is equivalent to $\mathcal{P}_1^{\mathrm{MEE}}$ that the optimal solution of $\mathcal{P}_2^{\mathrm{MEE}}$ is also the optimal solution of $\mathcal{P}_1^{\mathrm{MEE}}$. Hence, the problem $\mathcal{P}_2^{\mathrm{MEE}}$ is an equivalent reformulation of $\mathcal{P}_1^{\mathrm{MEE}}$. The problem $\mathcal{P}_2^{\mathrm{MEE}}$ is combinatorial due to constraint C_1 and C_3, and non-convex due to constraint C_3, C_4 and the objective. Letting δ_i to be the slack variable associated with group i and t to be the slack variable associated with power consumption, $\mathcal{P}_2^{\mathrm{MEE}}$ is transformed into a DC problem subject to binary constraints as,

$$\mathcal{P}_3^{\mathrm{MEE}} : \max_{\mathbf{W},\boldsymbol{\Theta},\boldsymbol{\eta},\boldsymbol{\delta},\boldsymbol{\alpha},\boldsymbol{\zeta}} \sum_{i=1}^N \sum_{j=1}^G f\left(\eta_{ij},\Theta_j,t\right) \quad (5)$$

$$\text{s.t. } C_1 : \eta_{ij} \in \{0,1\}, \ \forall i, \ \forall j,$$

$$C_2 : \sum_{j=1}^G \eta_{ij} \leq 1, \forall i,$$

$$C_3 : \sum_{i=1}^N \eta_{ij} \leq \delta_j N, \forall j,$$

$$C_4 : \delta_j \in \{0,1\}, \forall j,$$

$$C_5 : \sum_{l\neq i}|\mathbf{h}_i^H\mathbf{w}_l|^2 + \sigma^2 \leq \mathcal{J}_{ij}(\mathbf{W},\alpha_{ij}), \ \forall i, \ \forall j,$$

$$C_6 : \sum_{j=1}^M \|\mathbf{w}_j\|_2^2 \leq P_T,$$

$$C_7 : \left(\eta_{ij} + \Theta_j\right)^2 - 2\log\alpha_{ij} \leq \eta_{ij}^2 + \Theta_j^2, \ \forall i, \ \forall j,$$

$$C_8 : \Theta_j \geq \epsilon_j, \ \forall j,$$

$$C_9 : \tilde{g}(\mathbf{W},\boldsymbol{r}) \leq t,$$

$$C_{10} : \sum_{j=1}^G \delta_j = M, \ \forall j,$$

where $\quad f\left(\eta_{ij},\Theta_j,t\right) \quad \triangleq \quad B\Psi_j \frac{\left(\eta_{ij}+\Theta_j\right)^2 - \eta_{ij}^2 - \Theta_j^2}{2t},$

$\mathcal{J}_{ij}(\mathbf{W}, \alpha_{ij}) \triangleq \dfrac{\sum_{l=1}^{G} |\mathbf{h}_i^H \mathbf{w}_l|^2 + \sigma^2}{\alpha_{ij}}, \quad \tilde{g}(\mathbf{W}, \mathbf{r}) =$

$P_0 + \sum_{j=1}^{G} \left(\dfrac{1}{\rho} \|\mathbf{w}_j\|^2 + \mu \left(p_1(\Theta_j) - p_2(\Theta_j) \right) \right)$, and

$\boldsymbol{\delta} = [\delta_1, \ldots, \delta_G]^T$. In constraint C_3 in $\mathcal{P}_3^{\mathrm{MEE}}$, the binary slack variable δ_j is used for controlling the scheduling of group j. In other words, $\delta_j = 0$ indicates that group j is not scheduled else scheduled. However, for scheduled group j (i.e., $\delta_j = 1$) constraint C_3 becomes superfluous as it is always satisfied. With the help C_3 and C_4, constraint C_{10} ensures that number of scheduled groups is exactly M. Further, it is easy to see that $\tilde{g}(\mathbf{W}, \mathbf{r}) \geq g(\mathbf{W}, \mathbf{r})$. In other words, the objective in $\mathcal{P}_3^{\mathrm{MEE}}$ is strictly less than the objective in $\mathcal{P}_2^{\mathrm{MEE}}$. Therefore, the problem $\mathcal{P}_3^{\mathrm{MEE}}$ provides a strict lower bound to the problem in $\mathcal{P}_2^{\mathrm{MEE}}$.

C. Continuous DC using relaxation and penalization: MEE

Ignoring the combinatorial constraints C_1 and C_4, the constraint set of $\mathcal{P}_3^{\mathrm{MEE}}$ can be seen as a DC problem. So, the stationary points of such DC problems can be efficiently obtained by convex-concave procedure (CCP). With the aim of adopting the CCP framework, the binary constraints C_1 and C_4 in $\mathcal{P}_4^{\mathrm{MEE}}$ are relaxed to box constraint between 0 and 1 i.e., $[0, 1]$. The CCP framework can be readily applied to this relaxed continuous problem; however, the obtained stationary points might yield non-binary δ_js and η_{ij}s. Although, a quantization procedure can be used to obtain binary δ_js and η_{ij}s, the resulting solutions may not be even feasible to $\mathcal{P}_1^{\mathrm{MEE}}$. Therefore, obtaining binary δ_js and η_{ij} in the relaxed problem is crucial to ensure that the obtained solution are feasible to the original problem $\mathcal{P}_2^{\mathrm{MEE}}$. Therefore, the relaxed variables δ_js and η_{ij}s are further penalized to encourage the relaxed problem to include binary δ_js and η_{ij}s in the final solutions. Letting $\lambda_1 > 0$ and $\lambda_2 > 0$ be the penalty parameters respectively and $\mathbb{P}(.)$ be the penalty function, the penalized continuous formulation of $\mathcal{P}_4^{\mathrm{MEE}}$ is,

$$\mathcal{P}_4^{\mathrm{MEE}}: \max_{\mathbf{W}, \boldsymbol{\Theta}, \boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\alpha}, \boldsymbol{\zeta}, t} \sum_{i=1}^{N} \sum_{j=1}^{G} f(\eta_{ij}, \Theta_j, t)$$
$$+ \sum_{j=1}^{G} \sum_{i=1}^{N} \lambda_1 \mathbb{P}(\eta_{ij}) + \lambda_2 \sum_{j=1}^{G} \mathbb{P}(\delta_j) \qquad (6)$$
$$\text{s.t. } C_1 : 0 \leq \eta_{ij} \leq 1, \ \forall i, \ \forall j,$$
$$C_4 : 0 \leq \delta_j \leq 1, \forall j,$$
$$C_2, C_3, C_5 \text{ to } C_{10} \text{ in } (5)$$

It is easy to see that any choice of convex function $\mathbb{P}(\eta_{ij})$ that promotes the binary solutions suffice to transform $\mathcal{P}_4^{\mathrm{MEE}}$ as a DC problem of our interest. The entropy based penalty function proposed in [25] i.e., $\mathbb{P}(\eta_{ij}) \triangleq \eta_i \log \eta_{ij} + (1 - \eta_{ij}) \log (1 - \eta_{ij})$ is considered for this work. With this choice of $\mathbb{P}(\eta_{ij})$, the problem $\mathcal{P}_4^{\mathrm{MEE}}$ becomes a DC problem. In order to apply the CCP framework to the problem $\mathcal{P}_4^{\mathrm{MEE}}$, a feasible initial point (FIP) needs to supplied. However, the constraint C_5 in $\mathcal{P}_4^{\mathrm{MEE}}$ limits the choices of FIPs. For ease of finding the FIPs, the constraint C_{10} is brought into the

objective with another penalty parameter $\Omega_1 > 0$ as,

$$\mathcal{P}_5^{\mathrm{MEE}}: \max_{\mathbf{W}, \boldsymbol{\Theta}, \boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\alpha}, \boldsymbol{\zeta}, t} \sum_{i=1}^{N} \sum_{j=1}^{G} \left(f(\eta_{ij}, \Theta_j, t) + \lambda_1 \mathbb{P}(\eta_{ij}) \right)$$
$$+ \lambda_2 \sum_{j=1}^{G} \mathbb{P}(\delta_j) - \Omega_1 \left\| \sum_{j=1}^{G} \delta_j - M \right\|^2 \qquad (7)$$
$$\text{s.t. } C_1 \text{ to } C_9 \text{ in } (6)$$

D. A CCP based Joint Design Algorithm: MEE

In this section, a CCP based algorithm is proposed for joint user grouping, scheduling and precoding for MEE (JGSP-MEE) problem given in problem (7). CCP proposed in [26] is a special case of successive convex approximation framework [27] designed for DC programming problem. So, CCP is an iterative framework where in each iteration convexification and optimization steps are applied to the DC problem until the convergence. The convexification and optimization steps of $\mathcal{P}_5^{\mathrm{MEE}}$ of JGSP-MEE at the iteration k is given as,

- Convexification: Let $(\mathbf{W}, \boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\Theta}, \boldsymbol{\alpha}, , t)^{k-1}$ be the estimates of $(\mathbf{W}, \boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\Theta}, \boldsymbol{\alpha}, t)$ in iteration $k-1$ respectively. In iteration k, the functions $\mathbb{P}(\delta_j), \mathbb{P}(\eta_{ij}), p_2(\zeta_j)$ and $f(\eta_{ij}, \Theta_j, t)$ are replaced their first Taylor approximations $\tilde{\mathbb{P}}^k(\eta_{ij}), \tilde{\mathbb{P}}^k(\delta_j), \tilde{p}_2(\zeta_j)$, and $f^k(\eta_{ij}, \Theta_j, t)$ respectively which are given in Appendix I. Similarly, the concave parts in of C_5, C_7 and C_9 in $\mathcal{P}_5^{\mathrm{MEE}}$ are replayed their first Taylor approximations $\tilde{\mathcal{G}}_{ij}^k(\eta_{ij}, \Theta_j), \tilde{\mathcal{K}}_{ij}^k(\delta_j, \Theta_j), \tilde{\mathcal{J}}_{ij}^k(\mathbf{W}, \alpha_{ij})$ respectively given in Appendix I.
- Optimization: Updated $(\mathbf{W}, \boldsymbol{\alpha}, \boldsymbol{\Theta}, \boldsymbol{\eta}, \boldsymbol{\delta}, t)^{k+1}$ is obtained by solving the following convex problem,

$$\mathcal{P}_6^{\mathrm{MEE}}: \max_{\mathbf{W}, \boldsymbol{\Theta}, \boldsymbol{\zeta}, \boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\alpha}, t} \sum_{j=1}^{G} \lambda_2 \tilde{\mathbb{P}}^k(\delta_j) - \Omega_1 \left\| \sum_{j=1}^{G} \delta_j - M \right\|^2$$
$$+ \sum_{i=1}^{N} \sum_{j=1}^{G} \left(f^k(\eta_{ij}, \Theta_j, t) + \lambda_1 \tilde{\mathbb{P}}^k(\eta_{ij}) \right) \qquad (8)$$
$$\text{s.t. } C_1 \text{ to } C_4 \text{ and } C_6, C_8 \text{ in } (7),$$
$$C_5 : \sum_{l \neq i} |\mathbf{h}_j^H \mathbf{w}_l|^2 + \sigma^2 \leq \tilde{\mathcal{J}}_{ij}^k(\mathbf{W}, \alpha_{ij}), \ \forall i, \ \forall j,$$
$$C_7 : (\eta_{ij} + \Theta_j)^2 \leq 2 \log \alpha_{ij} + \tilde{\mathcal{G}}_{ij}^k(\eta_{ij}, \Theta_j), \ \forall i, \ \forall j.$$
$$C_9 : P_0 + \sum_{j=1}^{G} \left(\dfrac{1}{\rho} \|\mathbf{w}_j\|^2 + \mu \left(p_1(\zeta_j) - \tilde{p}_2(\zeta_j) \right) \right) \leq t,$$

The proposed CCP based JGSP-MEE algorithm iteratively solves the problem in $\mathcal{P}_6^{\mathrm{MEE}}$. However, to guarantee its convergence to a stationary point JGSP-MEE needs to be initialized with a FIP (kindly refer [28]). In this case, $\boldsymbol{\delta} = \boldsymbol{\eta} = \mathbf{0}$ results a trivial FIP. Although the trivial solution is a valid FIP to the problem $\mathcal{P}_5^{\mathrm{MEE}}$, it is observed through simulations that it usually converges to a poorly performing stationary point with the poor objective function value. This behavior might be due to the fact the trivial FIP has the lowest objective (i.e., zero), therefore, the JGSP-MEE initialized with the trivial FIP may

converge to a stationary point around this lowest objective value. Since, FIP is crucial for JGSP-MEE's performance, in the sequel, a simple procedure is proposed to obtain a FIP that promises the convergence to stationary points which yield better performance.

E. Feasible Initial Point: MEE

Since, the quality of the solution depends on the FIP, the harder task of finding a better FIP is considered through the following procedure.

- Step 1: Initialize $\mathbf{W}^0$ with complex random values subject to $\left\| \mathbf{W}^0 \right\|_2^2 \leq P_T$ and calculate initial SINRs $\boldsymbol{\gamma}^0$.
- Step 2: Solve the following optimization:

$$\mathcal{P}_{\text{FES}} : \{\boldsymbol{\delta}_0, \boldsymbol{\eta}_0\} : \quad \max \sum_{j=1}^{G} \delta_j + \sum_{j=1}^{G} \sum_{i=1}^{N} \eta_{ij} \quad (9)$$

$$\text{s.t. } C_1 : 0 \leq \eta_{ij} \leq 1, \; \forall i, \; \forall j,$$

$$C_2 : \sum_{j=1}^{G} \eta_{ij} \leq 1, \forall i,$$

$$C_3 : \sum_{i=1}^{N} \eta_{ij} \leq \delta_j N, \forall j,$$

$$C_4 : 0 \leq \delta_j \leq 1, \forall j,$$

$$C_5 : \log \left(1 + \gamma_{ij}^0 \right) \geq \eta_{ij} \epsilon_j, \; \forall i, \; \forall j.$$

- Step 3: The parameters $\boldsymbol{\Theta}^0, \boldsymbol{\zeta}^0, \boldsymbol{\alpha}^0, t^0$ can easily be derived from $\mathbf{W}^0, \boldsymbol{\delta}^0$ and $\boldsymbol{\eta}^0$.

Remarks:

- The problem $\mathcal{P}_{\text{FES}}$ is a linear programming problem and always feasible since trivial solution $\boldsymbol{\delta}^0 = \boldsymbol{\eta}^0 = \mathbf{0}$ is also a feasible solution. However, the optimization problem $\mathcal{P}_{\text{FES}}$ usually results a better solution than trivial one. Therefore, initial parameters $\mathbf{W}^0, \boldsymbol{\delta}^0, \boldsymbol{\Theta}^0, \boldsymbol{\eta}^0, \boldsymbol{\alpha}^0, t^0$ are always feasible. Different $\mathbf{W}^0$ in step 1 may lead to different FIPs.
- The optimization problem in Step 2 is a linear programming problem which can be solved efficiently to large dimensions with many of the existing tools like CVX.
- The FIP obtained by this procedure may not be feasible for the original MEE problem $\mathcal{P}_1^{\text{MEE}}$ unless $\mathbf{W}^0, \boldsymbol{\eta}^0, \boldsymbol{\delta}^0$ becomes feasible to $\mathcal{P}_2^{\text{MEE}}$.
- Although the FIP obtained by this method is not feasible for $\mathcal{P}_1^{\text{MEE}}$, the final solution obtained by JGSP-MEE with this FIP becomes a feasible for $\mathcal{P}_1^{\text{MEE}}$ since the final solution satisfies the group scheduling constraint C_2 in $\mathcal{P}_1^{\text{MEE}}$.

Letting $\mathcal{P}_6^{\text{MEE}} (k)$ be the objective value of the problem $\mathcal{P}_6^{\text{MEE}}$ at iteration k, the pseudo code of JGSP-MEE for the joint design problem is given in algorithm 1.

F. Complexity of JGSP-MEE

Since JGSP-MEE is a CCP based iterative algorithm, its complexity depends on complexity of the convex sub-problem $\mathcal{P}_6^{\text{MEE}}$ and $\mathcal{P}_{\text{FES}}$. $\mathcal{P}_6^{\text{MEE}}$ is a convex problem

---

Algorithm 1 JGSP-MEE

> **Input**: $\mathbf{H}, \{\epsilon_i\}_{i=1}^N, P_T, \Delta, \mathbf{W}^0, \boldsymbol{\delta}^0, \boldsymbol{\Theta}^0, \boldsymbol{\zeta}^0, \boldsymbol{\eta}^0, \boldsymbol{\alpha}^0, t^0,$ $\lambda_1 = 0, k = 1;$
>
> **Output**: $\mathbf{W}, \boldsymbol{\eta}$
>
> **while** $|\mathcal{P}_6^{\text{MEE}} (k) - \mathcal{P}_6^{\text{MEE}} (k - 1)| \geq \Delta$ **do**
>
>   **Convexification:** Convexify the problem (7)
>
>   **Optimization**: Update $(\mathbf{W}, \boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\alpha}, \boldsymbol{\zeta}, \boldsymbol{\Theta}, t)^k$ by solving $\mathcal{P}_6^{\text{MEE}}$
>
>   **Update :** $\mathcal{P}_6^{\text{MEE}} (k), \lambda_1, \lambda_2, \Omega_1, k$
>
> **end while**

---

with $(MG + 2NG + 3G + 1)$ decision variables and $(2NG + 2 + G)$ convex constraints and $(2NG + 4G + N)$ linear constraints. Hence, the complexity of $\mathcal{P}_6^{\text{MEE}}$ is $\mathcal{O} \left((MG + 2NG + 2G + 1)^3 (4NG + 4G + N + 2) \right)$ [29]. Similarly, $\mathcal{P}_{\text{FES}}$ is a linear programming problem with $(NG + G)$ decision variables and $(3NG + N + 3G)$ linear constraints, and its complexity is $\mathcal{O} \left((NG + G)^3 (3NG + N + 3G) \right)$. Similarly, Commercial software such as CVX can solve the convex problem of type $\mathcal{P}_6^{\text{MEE}}$ efficiently to a large dimension. Besides the complexity per iteration, the overall complexity also depends on the convergence speed of the algorithm. Through simulations, we observe that the JGSP-MEE converges typically in 15-20 iterations.

IV. VARIANTS OF MULTICAST ENERGY EFFICIENCY

Multicast energy efficiency combines three different aspects: the number of scheduled users, minimum throughput of groups and total consumed power. To understand the interplay or trade-off achieved among these factors, JSP-MEE is compared with the cases that solely maximizes the number of scheduled users and energy efficiency. Therefore, in this section, the two special cases of the MEE problem namely the maximization of EE and the number of scheduled users are considered.

A. Energy efficiency

In this section, we focus on developing a CCP based low-complexity algorithm for the joint design of user grouping, scheduling, and precoding for maximization of weighted EE (defined in (1)) subject to grouping, scheduling, precoding, power, and QoS constraints. This problem is simply referred to as the EE problem.

1) Problem formulation: EE: With the defined slack variables in Section III, the EE problem is mathematically formulated as,

$$\mathcal{P}_1^{\text{EE}} : \max_{\mathbf{W}, \boldsymbol{\Theta}, \boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\alpha}} \frac{\sum_{j=1}^{M} B\Psi_j \Theta_j}{t} \quad (10)$$

$$\text{s.t. } C_1 \text{ to } C_8 \text{ and } C_{11} \text{ in (5),}$$

$$C_9 : P_0 + \sum_{j=1}^{G} \left(\frac{1}{\rho} \left\| \mathbf{w}_j \right\|^2 + \mu \, p \left(\Theta_j \right) \right) \leq t,$$

$$C_{10} : \Theta_j \leq \delta_j \Theta^*, \forall j,$$

where $\Theta^* \geq \max_{j=1}^{G} \Theta_j$ is a constant. The constant Θ^* in constraint C_{12} in $\mathcal{P}_1^{\text{EE}}$ is used for forcing Θ_j to zero when the group is not scheduled i.e., $\delta_j = 0$. For the scheduled group i.e., $\delta_j = 1$ constraint C_{12} becomes superficial as $\Theta_j \leq \Theta*$ is always true. Without the constraint C_{12} the problem $\mathcal{P}_1^{\text{EE}}$ becomes unbounded as the Θ_j can be infinity for the unscheduled group j thus yielding the highest EE which is infinity. The constraint C_{12} helps in containing Θ_j to zero for the unscheduled group j. Therefore the problem $\mathcal{P}_1^{\text{EE}}$ becomes bounded due to C_{12}. Notice the difference between the constraint C_9 in $\mathcal{P}_1^{\text{EE}}$ and C_{10} in $\mathcal{P}_3^{\text{MEE}}$. Due to C_{10} in $\mathcal{P}_1^{\text{EE}}$ for an unscheduled group j the minimum rate of the group i.e., Θ_j is zero. Therefore, the power consumption can be modelled simply using Θ_j unlike ζ_j in MEE case.

Nature of EE in the context of grouping and scheduling: EE problem is not biased to favor the solutions with more number of users since it only considers the minimum rate of the group ignoring its size. Typically, adding more users to groups either leads to increased inter-group interference and/or lower minimum rate of the group due to lower channel gains. Hence, to obtain the same rate as with few users extra power needs to be used. Since the linear increase in rate is achieved at the cost of exponential increase power, newly added users result in lower EE.

2) DC formulation and CCP based algorithm: EE: The problem $\mathcal{P}_1^{\text{EE}}$ is combinatorial and non-convex similar to the problem $\mathcal{P}_3^{\text{MEE}}$. With the help of a slack variable Γ, and applying reformulations and relaxations proposed in Section III, the problem $\mathcal{P}_1^{\text{EE}}$ is reformulated into a DC problem as,

$$\mathcal{P}_2^{\text{EE}} : \max_{\mathbf{W},\boldsymbol{\Theta},\boldsymbol{\eta},\boldsymbol{\delta},\boldsymbol{\alpha},\Gamma,t} \frac{\Gamma^2}{t} - \Omega_2 \left\| \sum_{j=1}^{G} \delta_j - M \right\|^2$$
$$+ \sum_{j=1}^{G}\sum_{i=1}^{N} \lambda_3 \mathbb{P}\left(\eta_{ij}\right) + \lambda_4 \sum_{j=1}^{G} \mathbb{P}\left(\delta_j\right) \quad (11)$$
$$\text{s.t. } C_1 \text{ to } C_{10} \text{ in (10)},$$
$$C_{11} : \sum_{j=1}^{G} B\Psi_j \Theta_j \geq \Gamma^2.$$

where $\lambda_3 > 0, \lambda_4 > 0$ and $\Omega_2 > 0$ are the penalty parameters.

Notice that the DC problem $\mathcal{P}_2^{\text{EE}}$ resembles the DC problem $\mathcal{P}_5^{\text{MEE}}$, hence, the CCP framework proposed in Section III-D can be simply be adapted. The proposed CCP framework based algorithm for the EE problem is simply referred to as JGSP-EE. Since JGSP-EE is a CCP based iterative algorithm at iteration k it executes the following convex problem:

$$\mathcal{P}_3^{\text{EE}} : \max_{\mathbf{W},\boldsymbol{\Theta},\boldsymbol{\eta},\boldsymbol{\delta},\boldsymbol{\alpha},\Gamma,t} \frac{2\Gamma^{k-1}\Gamma}{t} - \Omega_2 \left\| \sum_{j=1}^{G} \delta_j - M \right\|^2$$
$$+ \lambda_3 \sum_{j=1}^{G}\sum_{i=1}^{N} \tilde{\mathbb{P}}^k\left(\eta_{ij}\right) + \sum_{j=1}^{G} \lambda_4 \tilde{\mathbb{P}}^k\left(\delta_j\right) \quad (12)$$
$$\text{s.t. } C_1 \text{ to } C_8 \text{ in (7)},$$
$$C_9 : P_0 + \sum_{j=1}^{G} \left(\frac{1}{\rho} \left\| \mathbf{w}_j \right\|^2 + \mu \left(p_1\left(\Theta_j\right) - \tilde{p}_2\left(\Theta_j\right) \right) \right) \leq t,$$

$$C_{10} : \Theta_j \leq \delta_j \Theta^*, \forall j,$$
$$C_{11} : \sum_{j=1}^{G} B\Psi_j \Theta_j \geq \Gamma^2.$$

Letting $\mathcal{P}_2^{\text{EE}}(k)$ be the objective value of the problem $\mathcal{P}_2^{\text{EE}}$ at iteration k, the pseudo code of JGSP-EE for the joint design problem is given in algorithm 2.

Algorithm 2 JGSP-EE-SR

Input: $\mathbf{H}, [\epsilon_1, \ldots, \epsilon_N], P_T, \Delta, \mathbf{W}^0, \boldsymbol{\delta}^0, \boldsymbol{\Theta}^0, \boldsymbol{\eta}^0, \boldsymbol{\alpha}^0, t^0,$
$\lambda_3, \lambda_4, \Omega_2, k = 1;$
Output: $\mathbf{W}, \boldsymbol{\eta}$
while $|\mathcal{P}_2^{\text{EE}}(k) \mathcal{P}_2^{\text{EE}}(k-1)(k-1)| \geq \Delta$ **do**
   Convexification: Convexify the problem (7)
   Optimization: Update $(\mathbf{W},\boldsymbol{\eta},\boldsymbol{\delta},\boldsymbol{\alpha},\boldsymbol{\Theta},\Gamma,t)^k$ by solving $\mathcal{P}_2^{\text{EE}}(k)$
   Update : $\mathcal{P}_2^{\text{EE}}(k), \lambda_3, \lambda_4, \Omega_2, k$
end while

B. Maximization of scheduled users

In this section, the problem of maximizing the scheduled users (SUM) is considered subject to grouping, scheduling, precoding, total power, and QoS constraints. This problem is simply referred to as SUM problem in this paper and is formulated as,

$$\mathcal{P}_1^{\text{SUM}} : \max_{\mathbf{W},\boldsymbol{\Theta},\boldsymbol{\eta},\boldsymbol{\delta}} \sum_{i=1}^{N}\sum_{j=1}^{G} \eta_{ij} \quad (13)$$
$$\text{s.t. } C_1 : \eta_{ij} \in \{0,1\}, \ \forall i, \ \forall j,$$
$$C_2 : \sum_{j=1}^{G} \eta_{ij} \leq 1, \forall i,$$
$$C_3 : \sum_{i=1}^{N} \eta_{ij} \leq \delta_j N, \forall j,$$
$$C_4 : \delta_j \in \{0,1\}, \forall j,$$
$$C_5 : \sum_{j=1}^{G} \delta_j = M, \ \forall j,$$
$$C_6 : \sum_{i=1}^{M} \left\| \mathbf{w}_i \right\|_2^2 \leq P_T,$$
$$C_7 : 1 + \gamma_{ij} \geq 1 + \eta_{ij}\epsilon_j, \ \forall i, \ \forall j.$$

Notice that except for constraint C_6 all the constraints and the objective in $\mathcal{P}_1^{\text{SUM}}$ are linear and convex. Further, similar to the constraint C_4 in $\mathcal{P}_2^{\text{EE}}$, the constraint C_6 can be easily equivalently transformed as a DC. Therefore, with the help of the relaxations and penalization approach provided in Section III and IV, the problem can be transformed as a DC programming problem. Hence, the CCP framework can be adapted to solve the resulting DC problem. The transformed DC problem and the convexified problem to be solved in the CCP framework for the SUM problem are given appendix VI. The CCP framework based algorithm proposed for the SUM problem is simply referred to as JGSP-SUM.

V. SIMULATION RESULTS

A. Simulation setup and parameter initialization

Simulation setup: In this section, the performance of the proposed algorithms JGSP-MEE, JGSP-EE and JGSP-SUM is evaluated. The system parameters discussed in this paragraph are common for all the figures. Bandwidth for all the groups is assumed to be 1 Hz i.e., $B = 1$ Hz. The coefficients of the channel matrix, i.e., h_{ij} are drawn from the complex normal distribution with zero mean and unit variance and noise variances at the receivers are considered to be unity i.e., $\sigma^2 = 1$. All the simulation results are averaged over 100 different channel realizations (CRs). Weights are assumed to be unity i.e., $\{\Psi_j = 1\}_{j=1}^{G}$. Following are the acronyms/definitions commonly used for all simulation results: 1) *Number of scheduled users*: the sum of all the scheduled users in all scheduled groups. 2) *Consumed power* $\triangleq P_0 + \sum_{j=1}^{G}\left(\frac{1}{\rho}\|\mathbf{w}_i\|^2 + \mu p\left(r_j\right)\right)$. 3) *Throughput* $\triangleq \sum_{j=1}^{M} B\Theta_j$.

Parameter initialization: Power amplifier efficiency i.e., ρ is assumed to 0.2, and $B = 1$MHz. The penalty parameters responsible for binary nature of $\boldsymbol{\eta}, \boldsymbol{\delta}$ are initialized as $\lambda_1 = \lambda_2 = 0.01$, $\lambda_3 = \lambda_4 = 0.5$ and $\lambda_5 = \lambda_6 = 0.05$. Further $\{\lambda_i\}_{i=1}^{6}$ are incremented by factor 1.2 each iteration. Notice the small (large) increments of the penalty values result in relatively slower (faster) convergence but lead to high (low) quality solutions. Furthermore, penalty parameters corresponding to group scheduling constraint are initialized to relatively larger values such as $\Omega_1 = 2.5$, $\Omega_2 = 5$, and $\Omega_3 = 1$ and are incremented by 1.5 in each iteration. MEE and SUM maximization criteria naturally encourage the solutions towards to non-zero $\boldsymbol{\eta}$ and $\boldsymbol{\delta}$. Therefore, low initial values and slow update of penalty parameters corresponding to MEE and SUM problems eventually result in a binary solution of $\boldsymbol{\eta}, \boldsymbol{\delta}$. Further, relatively large initial value and larger increments for Ω_1 and Ω_2 in each iteration, along with binary nature of $\boldsymbol{\delta}$, eventually ensures the group scheduling constraint i.e., $\sum_{j=1}^{G} \delta_j = M$. On the contrary, as discussed in Section IV, the EE problem is not biased towards the solutions with a higher number of users since it only considers the minimum rate of the group ignoring its size. Moreover, the solutions with $\eta_{ij} < 1$ might be encouraged as it would facilitate larger Θ_j hence better objective in $\mathcal{P}_1^{\text{EE}}$ (from constraint C_5 in the problem $\mathcal{P}_1^{\text{EE}}$). Hence, to ensure the group scheduling constraint and the binary nature of $\boldsymbol{\eta}, \boldsymbol{\delta}$, the penalty parameters are initialized to relatively larger values in EE than in MEE and SUM problems and incremented in large steps.

B. Performance as a function of total users (N)

In figure 1, the performance of the proposed algorithm i.e., JGSP-MEE, JGSP-EE, and JGSP-SUM is illustrated as a function of N varying from 20 to 32 in steps of 4 for $M = 5$, $G = 8$, $P_T = 10$dBW and $\epsilon_j = 1$ bps/Hz, $\forall j$. Fixed static power i.e., P_0 is assumed to be 12 dBW, $\mu = 2.4$Watts/$\left(\text{bits/sec}\right)^2$ [17], and $p\left(x\right) = x^2$ [20].

1) Number of scheduled users versus N: In figure 1a, the number of scheduled users is illustrated as function of N. Since JGSP-SUM directly maximizes the number of scheduled
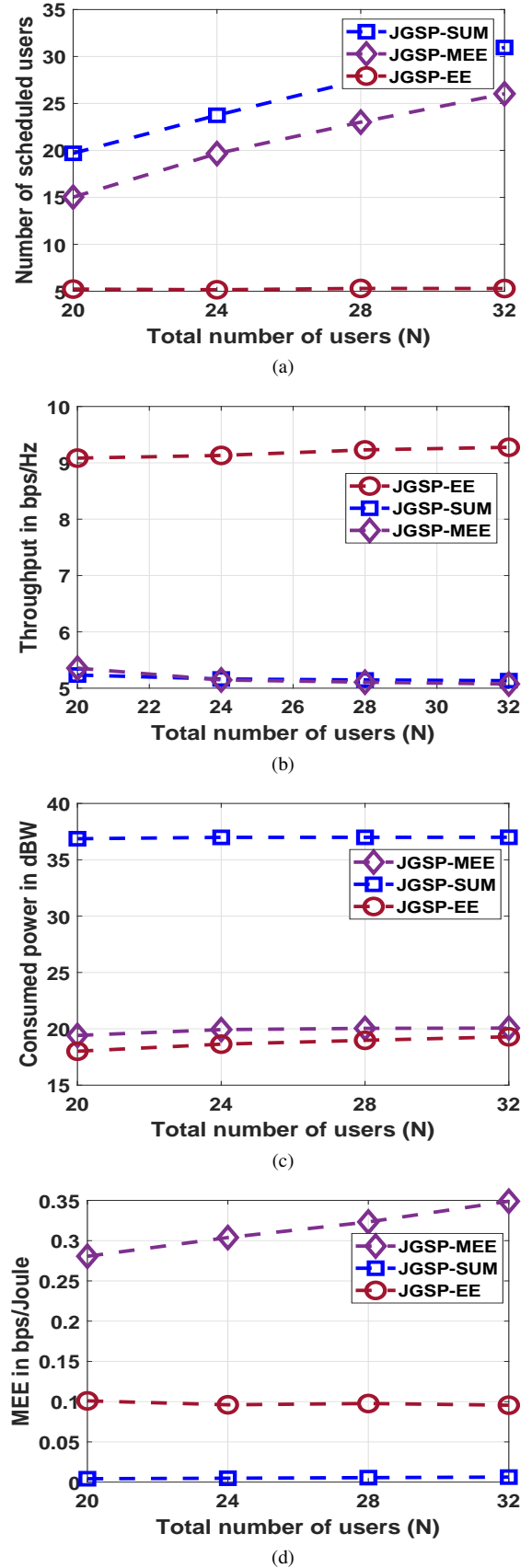

Fig. 1: Comparison of proposed algorithms as a function of N varying from 20 to 32 for $M = 5, G = 8, \{\epsilon_i = 1\text{bps/Hz}\}_{j=1}^{N}$, and $P_T = 20$dBW (a) number of scheduled users (b) throughput in bps (c) Consumed power in dBW (d) MEE in bits/Joule

users, it schedules the maximum number of users compared to JGSP-MEE and JGSP-EE. Moreover, due to low QoS requirement and availability of resources to satisfy the QoS requirement, JGSP-SUM schedules almost all the users despite the increase in N. Since the number of scheduled users contribute linearly to MEE objective a similar increase in the number of scheduled users versus N in JSP-MEE can be observed in figure 1a. However, JSP-MEE also considers the power consumed by the scheduled users, hence, JSP-MEE schedules fewer users than JSP-SUM as scheduling these excess users requires huge power which can be observed in figure 1c. On the contrary, the EE objective is not accounting for the number of scheduled users, hence, JGSP-EE schedules the lowest number of users i.e. $M = 5$. In other words, it is serving one user per group which is nothing but a unicast scenario. Furthermore, despite the increase in N, the number of users scheduled by JGSP-EE remains the same. This can be attributed to three reasons: 1) scheduling any user whose channel is highly correlated with users in other groups increases interference which decreases the minimum rate of the influenced groups hence decreases EE. 2) Correlated channels across the groups with with un-equal gains: EE swaps the existing user with the best available user in the pool as scheduling the second best user decreases the minimum rate of the group hence lower EE. 3) Correlated channels with equal gains: This is an unlikely event; even if such users exist, as mentioned earlier, their scheduling is not guaranteed as the EE objective is unaffected.

2) Throughput versus N: In figure 1b, the throughput in bps obtained by JGSP-MEE, JGSP-EE, and JGSP-SUM is illustrated as a function of N. The nature of JSP-MEE to schedule more users and consume fewer power results in lower throughput than JSP-SUM and JSP-EE. On the other hand, as the JSP-EE objective includes throughput in the objective, hence, it naturally achieves higher throughput than JSP-SUM. Moreover, as N increases the probability of finding M orthogonal users with good channels increases. This leads to a better throughput in JSP-EE with an increase of N. However, the gains in throughput for JGSP-EE diminishes as the gains in multiuser diversity diminin. On the contrary, an increase in multiuser diversity with N is utilized to schedule a higher number of users by JSP-MEE and JSP-SUM which can be observed in figure 1a. Moreover, the degradation in throughput in JSP-MEE and JSP-SUM is due to the combination of two factors: 1) for relatively lower N i.e., 20, after scheduling the maximum number of users, resources could be used to improve minimum throughput of the groups. 2) for relatively higher N, as scheduling higher users improve the objectives of JSP-SUM and JSP-EEE, the available power is used to schedule more users and this is also achieved by keeping their achieved minimum rate close to the required rates of the groups. Hence, the throughput by JSP-MEE and JSP-SUM decreases slightly with an increase of N.

3) Consumed power versus N: In figure 1b, the consumed power in Watts by JSP-SUM, JSP-EE and JSP-MEE is illustrated as a function of N. As the JSP-SUM does not optimize power, in the process of scheduling the maximum number of users (as shown in figure 1a) it inefficiently utilizes the

power by consuming all of the available power as depicted in figure 1c. On the contrary, as the EE and MEE objectives are penalized inversely for excess usage of power, both JSP-EE and JSP-MEE utilize power efficiently as shown in figure 1c. However, JSP-MEE slightly utilizes more power than JSP-EE as illustrated in figure 1c to schedule a higher number of users (as shown in figure 1a) as it improves over the MEE.

4) MEE versus N: Recall that MEE can be interpreted as the number of received bits for one joule of transmitted energy as explained in Section II-B. It can be seen in figure 1d, by directly optimizing MEE, JGSP-MEE obtains the highest MEE value compared to JGSP-EE and JGSP-SUM. The linear increase in MEE with respect to N can be observed in JSP-MEE and JSP-SUM as the number of scheduled users linearly with N in both the methods. However, as JSP-SUM utilizes the power inefficiently, it results in poorer MEE overall compared to JSP-MEE. Unlike JSP-MEE and JSP-SUM, the improvement in MEE obtained by JSP-EE is negligible as it does not gain in scheduled users and the increase in throughput is comparatively negligible.
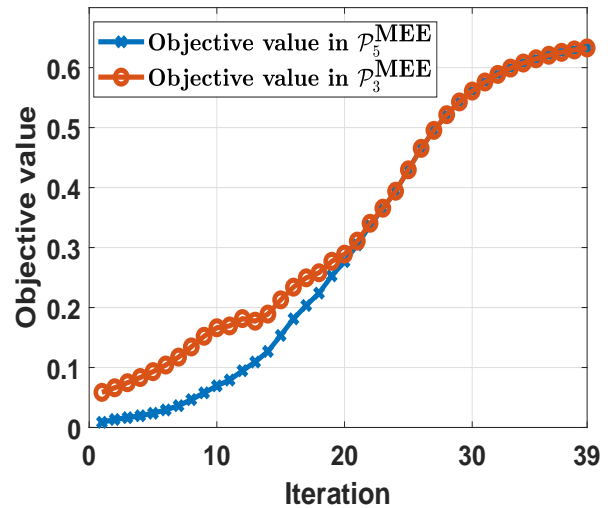

Fig. 2: Convergence of JSP-MEE as a function of iterations for $M = 5, G = 8, \{\epsilon_i = 1\text{bps/Hz}\}_{j=1}^N$, $N = 40$ and $P_T = 30$ dBW

5) Convergence of JSP-MEE versus iterations: Figure 2 illustrates the convergence behavior of the proposed algorithm. The objective value in $\mathcal{P}_3^{\text{MEE}}$ is simply the MEE value i.e., $\sum_{i=1}^N \sum_{j=1}^G f(\eta_{ij}, \Theta_j, t)$ and the objective value in $\mathcal{P}_5^{\text{MEE}}$ contains the sum of MEE and the penalty values added to ensure GSC constraint $\sum_{i=1}^N \sum_{j=1}^G f(\eta_{ij}, \Theta_j, t) + \lambda_2 \sum_{j=1}^G \mathbb{P}(\delta_j) + \sum_{j=1}^G \sum_{i=1}^N \lambda_1 \mathbb{P}(\eta_{ij}) - \Omega_1 \left\| \sum_{j=1}^G \delta_j - M \right\|^2$ and binary nature of $\boldsymbol{\eta}, \boldsymbol{\delta}$. In the initial iterations, $\boldsymbol{\eta}, \boldsymbol{\delta}$ and GSC constraints are not satisfied, hence, $\mathcal{P}_5^{\text{MEE}}$ has higher objective value than $\mathcal{P}_3^{\text{MEE}}$ which can be observed in Figure 2 until iteration 18. However, from iteration 19 the objective value of $\mathcal{P}_5^{\text{MEE}}$ and $\mathcal{P}_3^{\text{MEE}}$ almost same. This is because the additional penalty objective in $\mathcal{P}_5^{\text{MEE}}$ becomes zero i.e., $\lambda_2 \sum_{j=1}^G \mathbb{P}(\delta_j) + \sum_{j=1}^G \sum_{i=1}^N \lambda_1 \mathbb{P}(\eta_{ij}) - \Omega_1 \left\| \sum_{j=1}^G \delta_j - M \right\|^2 = 0$ as the

binary nature of $\boldsymbol{\eta}, \boldsymbol{\delta}$ and GSC constraints are satisfied by iteration 19. The proposed algorithm converges in 39 iterations for the system with $N = 40$, $M = 5$ and $G = 8$.

6) Number of scheduled users versus versus MEE: In figure 3, the performance of JSP-MEE is plotted as function of the number of users scheduled by JSP-MEE. Notice that the x-axis represents the average number of scheduled users. This corroborates the fact that the key factor contributing to the MEE is the number of scheduled users.
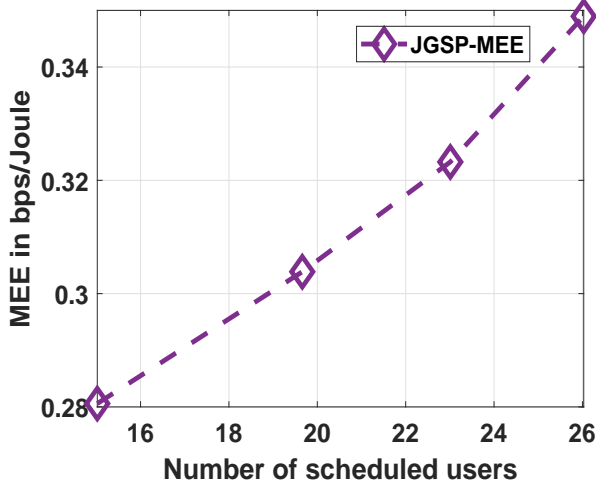
Fig. 3: Performance of JSP-MEE as function of Number of scheduled users versus MEE for N varying from 20 to 32, $M = 5, G = 8$, $\{\epsilon_i = 1\text{bps/Hz}\}_{j=1}^{N}$, and $P_T = 30$ dBW

C. Performance as a function of total power P_T

In figure 4, the performance of the proposed algorithms i.e., JGSP-MEE, JGSP-EE, and JGSP-SUM is illustrated as a function of P_T varying from 6 to 12 in steps of 2 dBW for $M = 5$, $G = 8$, $N = 15$ and $\epsilon_j = 1$ bps/Hz, $\forall j$. Fixed static power i.e., P_0 is assumed to be 14 dBW, $\mu = 0\text{Watts}/(\text{bits/sec})^2$.

1) Number of scheduled users versus P_T: In figure 4b, number of scheduled users is illustrated as function of P_T. By directly maximizing the number of scheduled users, JSP-SUM schedules the maximum number of users compared to JGSP-MEE and JGSP-EE. In the low-available power regime i.e., $P_T = 6$ dBW and 8 dBW, JSP-SUM schedules only few users. However, in the high-available power regime, due to the sufficient power, JSP-SUM schedules almost all the users i.e., 15 users by utilizing all of the power. Unlike JSP-SUM, despite the increased available power, the number of scheduled users in JSP-MEE is saturated to 13 users. This is because scheduling those extra users results in the consumption of huge power which decreases the overall MEE. Moreover, for the available for 8 dBW, JSP-SUM and JSP-MEE schedules almost equal number of users, however, JSP-MEE consumes almost 6.8 dBW less than JSP-SUM. On the contrary, for the reasons mentioned in section V-B, JSP-EE schedules only M despite the availability of power.
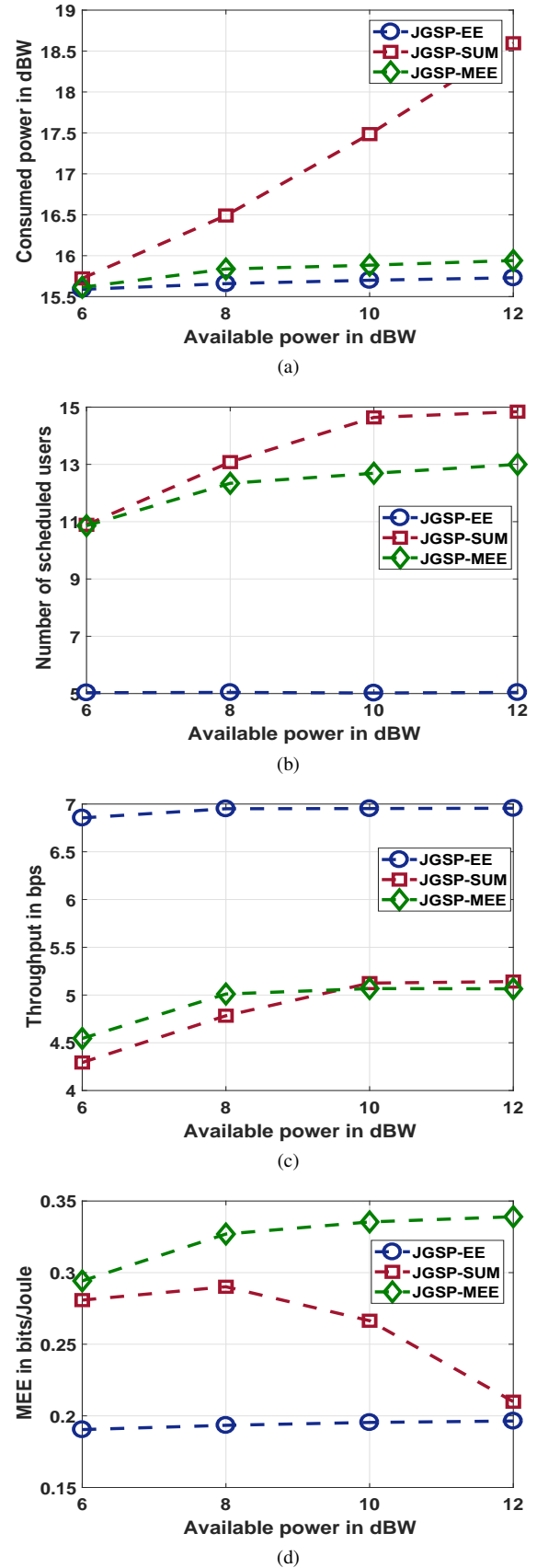

(a)

(b)

(c)

(d)

Fig. 4: Comparison of proposed algorithms as a function of P_T varying from 6 to 12 for $M = 5, G = 8$, $\{\epsilon_i = 1\text{bps/Hz}\}_{j=1}^{N}$, and $N = 15$ (a) consumed power in dBW (b) number of scheduled users (c) throughput in bps (d) MEE in bits/Joule

2) Throughput versus P_T: In figure 4c, the throughput in bps obtained by JGSP-MEE, JGSP-EE, and JGSP-SUM is illustrated as a function of P_T. Since JSP-MEE and JSP-SUM sacrifice in throughput to schedule more users for the available power, the lower throughput of JSP-MEE and JSP-SUM compared to JSP-EE can be observed in figure 4c. Moreover, the throughput of JSP-EE saturates to 8 bps for available power of 8 dBW as improving throughput further results in the consumption of huge power which results in overall lower EE. On the other hand, in the low-available power regime $P_T = 6$ and 8 dBW, the number of scheduled users (by JSP-SUM) are around 11 and 13 which less than total number of users $N = 15$. In other words, scheduling a higher number of users than 13 requires higher available power than 8 dBW. Therefore, the available-power in this regime is used to improve the minimum throughput of scheduled groups by JSP-MEE which can be observed in figure 4d. In this high-available power regime i.e., $P_T \geq 10$ dBW, JSP-SUM uses all of the available power to schedule almost all the users as shown in figure 4d. On the contrary, despite the availability of the power to schedule all users and/or to improve the throughput, JSP-MEE relatively maintains the same throughput as for the case of $P_T = 8$ dBW since the improvement in throughput leads to consumption of huge power.

3) Consumed power versus P_T: In figure 4a, the consumed power in Watts by JSP-SUM, JSP-EE and JSP-MEE is illustrated as a function of N. As the JSP-SUM does not optimize power, the inefficient utilization of the available power of JSP-SUM can be observed in figure 4a. On the contrary, as the EE and MEE objectives include the consumed power in the denominator, the objective values of EE and MEE are decreases inversely for a linear increase in consumed power. Hence, JSP-EE and JSP-MEE utilize power efficiently as shown in figure 4a.

4) MEE versus P_T: In figure 4d, the MEE in bits/Joule obtained by JGSP-MEE, JGSP-EE, and JGSP-SUM is illustrated as a function of P_T. By striking the trade-off among optimizing the number of scheduled users, throughput, and consumed power, JSP-MEE obtains higher MEE compared to JSP-EE and JSP-SUM as shown in figure 4d. Although JSP-SUM schedules more users than JSP-MEE, it does so by consuming huge power and also by inefficiently utilizing the available power. This results in decreasing in MEE of JSP-SUM with the increase of available power. On the other hand, JSP-MEE schedules also increase the number of users while simultaneously optimizing power and throughput. This results in an overall better MEE of JSP-MEE. On the contrary, JSP-EE schedules only $M = 5$ users despite the opportunity to schedule more users. Hence, JSP-EE results in the lowest MEE. However, JSP-MEE schedules more users while efficiently utilizing power and throughput.

VI. Conclusions

In this paper, the joint design of user grouping, scheduling, and precoding problem was considered for the message-based multigroup multicast scenario in multiuser MISO downlink channels. In this context, to fully leverage the multicast potential, a novel metric called multicast energy efficiency is considered as a performance metric. Further, this joint design problem is formulated as a structured MINLP problem with the help of Boolean variables addressing the scheduling and grouping, and linear variables addressing the precoding aspect of the design. Noticing the structure in MINLP to be difference-convex/concave, this paper proposed efficient reformulations and relaxations to transform it into structured DC programming problems. Subsequently, the paper proposed CCP based algorithms for MEE and its variants i.e., EE and SUM problems (JSP-MEE, JSP-EE, and JSP-SUM) which are guaranteed to converge to a stationary point for the aforementioned DC problems. Finally, the paper proposed low-complexity procedures to obtain good feasible initial points, critical to the implementation of CCP based algorithms. Through simulations, the paper established the efficacy of the proposed joint techniques and studied the influence of the algorithms on the different parameters namely scheduled users, multicast throughput, and consumed power. This framework can be applied to systems with any number of antennas as such. A smaller number is considered in the simulations for representative purposes. Further investigation is needed on simplification (e.g., MUI cancellation, MRT transmissions, SNR based scheduling) brought in the case of larger number of antennas than groups. This is left for future works.

Appendix I

$$\tilde{\mathbb{P}}^k\left(\eta_{ij}\right) \triangleq \eta_{ij}\nabla\mathbb{P}\left(\eta_{ij}^{k-1}\right); \ \tilde{\mathbb{P}}^k\left(\delta_j\right) \triangleq \delta_j\nabla\mathbb{P}\left(\delta_j^{k-1}\right),$$

$$f^k\left(\eta_{ij},\Theta_j,t\right) \triangleq B\left(2\frac{\left(\eta_{ij}^{k-1}+\Theta_j^{k-1}\right)\left(\eta_{ij}+\Theta_j\right)}{t^{k-1}} - \left(\frac{\eta_{ij}^{k-1}+\Theta_j^{k-1}}{t^{k-1}}\right)^2 t - \frac{\eta_{ij}^2}{t} - \frac{\Theta_j^2}{t}\right).$$

Similarly, linearization of the concave part of C_6, C_7 and C_9 in $\mathcal{P}_5^{\text{MEE}}$ is given by

$$\tilde{\mathcal{G}}_{ij}^k(\eta_{ij},\Theta_j) = -\left[\eta_{ij}^{k-1}\right]^2 - \left[\Theta_j^{k-1}\right]^2 + 2\eta_{ij}^{k-1}\eta_{ij} + 2\Theta_j^{k-1}\Theta_j,$$

$$\tilde{\mathcal{K}}_{ij}^k(\delta_j,\Theta_j) \triangleq \left(\delta_j^{k-1}+\Theta_j^{k-1}\right)^2 - 2\begin{bmatrix}\delta_j^{k-1}+\Theta_j^{k-1}\\\delta_j^{k-1}+\Theta_j^{k-1}\end{bmatrix}^H\begin{bmatrix}\left(\delta_j-\delta_j^{k-1}\right)\\\Theta_j-\Theta_j^{k-1}\end{bmatrix}$$

$$\tilde{\mathcal{J}}_{ij}^k(\mathbf{W},\alpha_{ij}) \triangleq -\mathcal{J}_{ij}(\mathbf{W}^{k-1},\alpha_{ij}^{k-1}) - \Re\left\{\nabla^H\mathcal{J}_{ij}(\mathbf{W}^{k-1},\alpha_{ij}^{k-1})\begin{bmatrix}\{\mathbf{w}_l-\mathbf{w}_l^{k-1}\}_{l=1}^G\\\alpha_{ij}-\alpha_{ij}^{k-1}\end{bmatrix}\right\},$$

where

$$\nabla\mathcal{J}_{ij}(\mathbf{W}^{k-1},\alpha_{ij}^{k-1}) = \begin{bmatrix}\{\frac{2\mathbf{h}_i\mathbf{h}_i^H\mathbf{w}_l^{k-1}}{\alpha_{ij}^{k-1}}\}_{l=1}^G,\\-\frac{\sigma^2+\sum_{l=1}^G|\mathbf{h}_i^H\mathbf{w}_l^{k-1}|^2}{\alpha_{ij}^{k-1^2}}\end{bmatrix}. \quad (14)$$

APPENDIX II

DC formulation and CCP based algorithm: SUM

Applying reformulations and relaxations proposed in Section III, the problem $\mathcal{P}_1^{\text{SUM}}$ is reformulated into DC problem as,

$$\mathcal{P}_2^{\text{SUM}} : \max_{\mathbf{W},\mathbf{\Theta},\boldsymbol{\eta},\boldsymbol{\delta}} \sum_{i=1}^{N} \sum_{j=1}^{G} \eta_{ij} - \Omega_3 \left\| \sum_{j=1}^{G} \delta_j - M \right\|^2$$

$$+ \lambda_4 \sum_{j=1}^{G} \sum_{i=1}^{N} \mathbb{P}\left(\eta_{ij}\right) + \sum_{j=1}^{G} \lambda_5 \mathbb{P}\left(\delta_j\right) \qquad (15)$$

s.t. $C_1 : 0 \leq \eta_{ij} \leq 1, \ \forall i, \ \forall j,$

$\quad C_4 : 0 \leq \delta_j \leq 1, \forall j,$

$\quad C_7 : \sum_{l \neq i} |\mathbf{h}_i^H \mathbf{w}_l|^2 + \sigma^2 \leq \dfrac{\sum_{l=1}^{G} |\mathbf{h}_i^H \mathbf{w}_l|^2 + \sigma^2}{1 + \eta_{ij} \epsilon_j}, \ \forall i, \ \forall j,$

$\quad C_2, C_3, C_5$ and C_6 in (13),

where λ_5, λ_6 and Ω_3 are the penalty parameters.

The convexified problem to be solved as part of JGSP-SUM (CCP based algorithm applied to the DC problem $\mathcal{P}_2^{\text{SUM}}$) algorithm at iteration k is:

$$\mathcal{P}_3^{\text{SUM}} : \max_{\mathbf{W},\mathbf{\Theta},\boldsymbol{\eta},\boldsymbol{\delta}} \sum_{i=1}^{N} \sum_{j=1}^{G} \eta_{ij} - \Omega_3 \left\| \sum_{j=1}^{G} \delta_j - M \right\|^2$$

$$+ \lambda_4 \sum_{j=1}^{G} \sum_{i=1}^{N} \tilde{\mathbb{P}}^k\left(\eta_{ij}\right) + \sum_{j=1}^{G} \lambda_5 \tilde{\mathbb{P}}^k\left(\delta_j\right) \qquad (16)$$

$\quad$ s.t. C_1, C_2, C_3, C_5 to C_6 in (15),

$\quad C_7 : \sum_{l \neq i} |\mathbf{h}_i^H \mathbf{w}_l|^2 + \sigma^2 \leq \mathcal{I}^k\left(\mathbf{W}, \eta_{ij}\right), \ \forall i, \ \forall j,$

where $\mathcal{I}^k\left(\mathbf{W}, \eta_{ij}\right) = \dfrac{\sum_{l=1}^{G} |\mathbf{h}_i^H \mathbf{w}_l^{k-1}|^2 + \sigma^2}{1 + \eta_{ij}^{k-1} \epsilon_j} +$

$$\Re \left\{ \left[\{ \dfrac{2\mathbf{h}_i \mathbf{h}_i^H \mathbf{w}_l^{k-1}}{1 + \eta_{ij}^{k-1} \epsilon_j} \}_{l=1}^{G} - \epsilon_j \dfrac{\sigma^2 + \sum_{l=1}^{G} |\mathbf{h}_i^H \mathbf{w}_l^{k-1}|^2}{\left(1 + \eta_{ij}^{k-1} \epsilon_j\right)^2} \right]^H \right.$$
$$\left. \begin{bmatrix} \{ \mathbf{w}_l - \mathbf{w}_l^{k-1} \}_{l=1}^{G} \\ \eta_{ij} - \eta_{ij}^{k-1} \end{bmatrix} \right\}.$$

Letting $\mathcal{P}_3^{\text{SUM}}(k)+$ be the objective value of the problem $\mathcal{P}_3^{\text{SUM}}$ at iteration k, the pseudo code of JGSP-EE-SR for the joint design problem is given in algorithm 3.

Algorithm 3 JGSP-SUM

Input: $\mathbf{H}, [\epsilon_1, \ldots, \epsilon_N], P_T, \Delta, \mathbf{W}^0, \boldsymbol{\delta}^0, \mathbf{\Theta}^0, \boldsymbol{\eta}^0, \ \lambda_4 = 0, \lambda_5 = 0, \Omega_3 = 0, k = 1;$
Output: $\mathbf{W}, \boldsymbol{\eta}$
while $|\mathcal{P}_3^{\text{SUM}}(k) - \mathcal{P}_3^{\text{SUM}}(k-1)| \geq \Delta$ **do**
$\quad$ **Convexification:** Convexify the problem (16)
$\quad$ **Optimization**: Update $(\mathbf{W}, \boldsymbol{\eta}, \boldsymbol{\delta}, \mathbf{\Theta})^k$ by solving $\mathcal{P}_3^{\text{SUM}}$
$\quad$ **Update :** $\mathcal{P}_3^{\text{SUM}}(k), \lambda_4, \lambda_5, \Omega_3, k$
end while

REFERENCES

[1] A. Bandi, M. R. Bhavani Shankar, S. Chatzinotas, and B. Ottersten, "Joint scheduling and precoding for frame-based multigroup multicasting in satellite communications," in *IEEE Global Communications Conference*, December 2019.

[2] ——, "Joint user scheduling, and precoding for multicast spectral efficiency in multigroup multicast systems," in *2020 IEEE International Conference on Signal Processing and Communications (SPCOM)*, July 2020.

[3] ITU-R, "IMT vision framework and overall objectives of the future development of IMT for 2020 and beyond," *Tech. Rep.*, 2015.

[4] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental trade-offs on green wireless networks," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 30–37, June 2011.

[5] F. Meshkati, H. V. Poor, and S. C. Schwartz, "Energy-efficient resource allocation in wireless networks," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 58–68, 2007.

[6] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental trade-offs on green wireless networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 30–37, 2011.

[7] E. Karipidis, N. D. Sidiropoulos, and Z. Luo, "Quality of service and max-min fair transmit beamforming to multiple cochannel multicast groups," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1268–1279, March 2008.

[8] M. Alodeh, D. Spano, A. Kalantari, C. G. Tsinos, D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Symbol-level and multicast precoding for multiuser multiantenna downlink: A state-of-the-art, classification, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 1733–1757, thirdquarter 2018.

[9] D. Lecompte and F. Gabin, "Evolved multimedia broadcast/multicast service (eMBMS) in LTE-advanced: overview and Rel-11 enhancements," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 68–74, November 2012.

[10] Z. Xiang, M. Tao, and X. Wang, "Coordinated multicast beamforming in multicell networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 12–21, January 2013.

[11] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Multicast multigroup precoding and user scheduling for frame-based satellite communications," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 4695–4707, Sept 2015.

[12] B. Hu, C. Hua, C. Chen, and X. Guan, "User grouping and admission control for multi-group multicast beamforming in MIMO systems," *Wireless Networks*, vol. 24, no. 8, pp. 2851–2866, Nov 2018. [Online]. Available: https://doi.org/10.1007/s11276-017-1510-5

[13] E. Matskani, N. D. Sidiropoulos, Z. Luo, and L. Tassiulas, "Efficient batch and adaptive approximation algorithms for joint multicast beamforming and admission control," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4882–4894, Dec 2009.

[14] H. Zhou and M. Tao, "Joint multicast beamforming and user grouping in massive MIMO systems," in *2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 1770–1775.

[15] S. He, Y. Huang, S. Jin, and L. Yang, "Energy efficient coordinated beamforming design in multi-cell multicast networks," *IEEE Commun. Lett.*, vol. 19, no. 6, pp. 985–988, June 2015.

[16] J. Denis, S. Smirani, B. Diomande, T. Ghariani, and B. Jouaber, "Energy-efficient coordinated beamforming for multi-cell multicast networks under statistical CSI," in *2017 IEEE 18th International Workshop SPAWC*, July 2017, pp. 1–5.

[17] O. Tervo, L. Tran, H. Pennanen, S. Chatzinotas, B. Ottersten, and M. Juntti, "Energy-efficient multicell multigroup multicasting with joint beamforming and antenna selection," *IEEE Trans. Signal Process.*, vol. 66, no. 18, pp. 4904–4919, Sep. 2018.

[18] Y. Li, M. Xia, and Y. Wu, "Energy-efficient precoding for non-orthogonal multicast and unicast transmission via first-order algorithm," *IEEE Trans. Wireless Commun.*, vol. 18, no. 9, pp. 4590–4604, Sep. 2019.

[19] O. Tervo, A. Tlli, M. Juntti, and L. Tran, "Energy-efficient beam coordination strategies with rate-dependent processing power," *IEEE Trans. Signal Process.*, vol. 65, no. 22, pp. 6097–6112, Nov 2017.

[20] Y. Yang, M. Pesavento, S. Chatzinotas, and B. Ottersten, "Energy efficiency optimization in mimo interference channels: A successive pseudoconvex approximation approach," *IEEE Trans. Signal Process.*, vol. 67, no. 15, pp. 4107–4121, 2019.

[21] Y. Zhang, J. An, K. Yang, X. Gao, and J. Wu, "Energy-Efficient User Scheduling and Power Control for Multi-Cell OFDMA Networks Based

on Channel Distribution Information," *IEEE Trans. Signal Process.*, vol. 66, no. 22, pp. 5848–5861, 2018.

[22] A. Zappone, L. Sanguinetti, G. Bacci, E. Jorswieck, and M. Debbah, "Energy-efficient power control: A look at 5g wireless technologies," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1668–1683, 2016.

[23] M. Kaliszan, E. Pollakis, and S. Staczak, "Multigroup multicast with application-layer coding: Beamforming for maximum weighted sum rate," in *2012 IEEE WCNC conference*, April 2012, pp. 2270–2275.

[24] T. X. Tran and G. Yue, "Grab: Joint adaptive grouping and beamforming for multi-group multicast with massive mimo," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[25] A. Bandi, B. S. M. R, S. Chatzinotas, and B. Ottersten, "A joint solution for scheduling and precoding in multiuser MISO downlink channels," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2019.

[26] A. L. Yuille and A. Rangarajan, "The concave-convex procedure (CCCP)," in *NIPS*, 2001.

[27] B. R. Marks and G. P. Wright, "Technical note-a general inner approximation algorithm for nonconvex mathematical programs," *Oper. Res.*, vol. 26, no. 4, p. 681683, Aug. 1978. [Online]. Available: https://doi.org/10.1287/opre.26.4.681

[28] B. K. Sriperumbudur and G. R. G. Lanckriet, "On the convergence of the concave-convex procedure," in *Neural Inf. Proc. Syst.*, Feb. 2009, pp. 1–9.

[29] P. Gahinet, A. Nemirovski, A. J. Laub, and M. Chilali, *LMI Control Toolbox Users Guide.* USA: MathWorks,, 1995.

Symeon Chatzinotas (S'06-M'09-SM'13) is currently Assistant Professor and Deputy Head of the SIGCOM Research Group at SnT University of Luxembourg. In the past, he has been a Visiting Professor at the University of Parma, Italy and he was involved in numerous Research and Development projects for the National Center for Scientific Research Demokritos, the Center of Research and Technology Hellas and the Center of Communication Systems Research, University of Surrey. He received the M.Eng. degree in telecommunications from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2003, and the M.Sc. and Ph.D. degrees in electronic engineering from the University of Surrey, Surrey, U.K., in 2006 and 2009, respectively. He was a co-recipient of the 2014 IEEE Distinguished Contributions to Satellite Communications Award, the CROWNCOM 2015 Best Paper Award and the 2018 EURASIC JWCN Best Paper Award. He has authored more than 350 technical papers in refereed international journals, conferences and scientific books. He has been the PI or Co-PI of more than 30 projects funded by FNR, ESA and H2020.

Ashok BANDI (S18) was born in Kunkalagunta, Andhra Pradesh, India, in 1988. He received the M.Tech. degree in electronics and communication engineering from the National Institute of Technology (NIT), Tiruchirappalli, India, in 2012. He has worked on Physical layer design and development for WLAN 802.11a/n/ac at Imagination Technologies, Hyderabad, India from 2012 to 2015 and at National Instruments, Bangalore, India (2015-2016). He was worked as a project associate in the department of ECE from 2016 to May 2017. He is currently pursuing the Ph.D. degree in electrical engineering with the University of Luxembourg. He joined the Interdisciplinary Centre for Security, Reliability, and Trust, University of Luxembourg, Luxembourg, in June 2017. He is working on sparse signal Recovery and joint update of integer and non-linear variables in MINLP problems that apprear in for Wireless communications within the Project PROSAT(on-board PROcessing techniques for high throughput SATellites), funded under FNR CORE-PPP Framework.

Björn Ottersten (S'87-M'89-SM'99-F'04) was born in Stockholm, Sweden, in 1961. He received the M.S. degree in electrical engineering and applied physics from Linkping University, Linköping, Sweden, in 1986, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 1990. He has held research positions with the Department of Electrical Engineering, Linköping University, the Information Systems Laboratory, Stanford University, the Katholieke Universiteit Leuven, Leuven, Belgium, and the University of Luxembourg, Luxembourg. From 1996 to 1997, he was the Director of Research with ArrayComm, Inc., a start-up in San Jose, CA, USA, based on his patented technology. In 1991, he was appointed a Professor of signal processing with the Royal Institute of Technology (KTH), Stockholm, Sweden. From 1992 to 2004, he was the Head of the Department for Signals, Sensors, and Systems, KTH, and from 2004 to 2008, he was the Dean of the School of Electrical Engineering, KTH. He is currently the Director for the Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg. As Digital Champion of Luxembourg, he acts as an Adviser to the European Commission.

He was a recipient of the IEEE Signal Processing Society Technical Achievement Award in 2011, EURASIP Meritorious Service Award 2010, and the European Research Council advanced research grant twice, in 2009-2013 and in 2017-2022. He has co-authored journal papers that received the IEEE Signal Processing Society Best Paper Award in 1993, 2001, 2006, and 2013, and seven other IEEE conference papers best paper awards. He has served as Editor-in-Chief of EURASIP Signal Processing Journal, Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the Editorial Board of the IEEE Signal Processing Magazine. He is currently a member of the editorial boards of EURASIP Signal Processing Journal, EURASIP Journal of Advances Signal Processing and Foundations and Trends of Signal Processing. He is a fellow of IEEE and EURASIP.

M. R. Bhavani Shankar (M11SM15) received Masters and Ph. D in Electrical Communication Engineering from Indian Institute of Science, Bangalore in 2000 and 2007 respectively. He was a Post Doc at the ACCESS Linnaeus Centre, Signal Processing Lab, Royal Institute of Technology (KTH), Sweden from 2007 to September 2009. He joined SnT in October 2009 as a Research Associate and is currently a Research Scientist at SnT. He was with Beceem Communications, Bangalore from 2006 to 2007 as a Staff Design Engineer working on Physical Layer algorithms for WiMAX compliant chipsets. He was a visiting student at the Communication Theory Group, ETH Zurich, headed by Prof. Helmut Blcskei during 2004. Prior to joining Ph. D, he worked on Audio Coding algorithms in Sasken Communications, Bangalore as a Design Engineer from 2000 to 2001. His research interests include Design and Optimization of MIMO Communication Systems, Automotive Radar and Array Processing, polynomial signal processing, Satellite communication systems, Resource Allocation, Game Theory and Fast Algorithms for Structured Matrices. He is currently on the Executive Committee of the IEEE Benelux joint chapter on communications and vehicular technology, member of the EURASIP Technical Area Committee on Theoretical and Methodological Trends in Signal Processing and serves as handling editor for Elsevier Signal Processing. He was a co-recipient of the 2014 Distinguished Contributions to Satellite Communications Award, from the Satellite and Space Communications Technical Committee of the IEEE Communications Society.