Multigateway Precoded NOMA in Multibeam Satellite Multicast Systems

Dong-Hyoun Na, Ki-Hong Park, Young-Chai Ko, and Mohamed-Slim Alouini

Abstract—In this paper, we consider multigateway-based multibeam satellite non-orthogonal multiple access (NOMA) multicast systems. We investigate the improvement in spectral efficiency and accessibility by superimposing multiple signals on each beam at the same frequency and time resource employing NOMA, where multiple gateways transmit precoded signals to alleviate inter-beam interference caused by full frequency reuse. We formulate an optimization problem of maximizing the sum rate while satisfying the gateways and satellite transmission power constraints, where the precoding vector and power allocation for superposition coding as well as the decoding order for successive interference cancellation are optimized. This optimization problem is challenging to solve due to its non-convex mixed-integer nonlinear programming. However, a suboptimal solution can be obtained using a block coordinate descent algorithm. The simulation results of the proposed NOMA technique are compared with those of the orthogonal multiple access (OMA) technique. The proposed technique outperforms the OMA technique. We also investigate the impact of channel imperfection and decoding capability of the proposed algorithm through some selected simulation results.

Index Terms—Multibeam satellite systems, non-orthogonal multiple access, precoding, sum-rate maximization.

I. INTRODUCTION

SATELLITE communication (SatCom) is expected to satisfy future communication requirements such as seamless access and high-speed broadband communication. Moreover, SatCom can be suitably exploited to ensure communication fairness in areas where communication demands are high and the deployment of terrestrial communication facilities is difficult or impossible [1]. In 6G communication, the SatCom infrastructure should be taken into account because three-dimensional scenarios with flying nodes are suggested for ubiquitous connectivity [2], [3]. In addition, the development of non-terrestrial communication (NTC) utilizing satellites has been studied, and the performance of NTC with millimeter-wave bands has been investigated in [4].

Conventional geostationary (GEO) satellites employ a single beam to cover a large area. However, it has been recently considered that GEO satellites take advantage of multiple beams with frequency reuse to increase throughput. Four-color frequency reuse has been mainly implemented. Four-color frequency reuse means dividing the frequency band into four so that adjacent beams use different frequency bands. By using four-color frequency reuse, multiple beams can be employed with lower inter-beam interference and different signals can be transmitted to the beams simultaneously. In order to use the limited spectrum more efficiently, full frequency reuse has been considered [5]. However, such aggressive frequency reuse significantly increases inter-beam interference. Accordingly, a joint use of zero-forcing (ZF) precoding and beam-hopping has been proposed to reduce inter-cluster, inter-beam, and intra-beam interference, where the beams operate according to the users' requirements in these beams [6]. In [7], the authors presented the concept of precoded cluster-hopping with minimum mean square error (MMSE). Compared with the four-color frequency reuse, these methods efficiently alleviate interference and match the capacity required by each beam.

A. Related Works

Obtaining channel state information (CSI) from ground users is necessary to transmit precoded data at gateways. In a fixed GEO SatCom, slowly varying channel enables gateways to obtain CSI with training symbols [8]. Even if outdated CSI is acquired due to the long propagation delay, employing precoding techniques is a pragmatic way to deal with interference [9]. In [10], nonlinear precoding techniques were adopted to mitigate interference. However, linear precoding techniques have primarily been studied because nonlinear precoding has a high implementation complexity. The authors in [11] designed an energy-efficient power allocation technique with MMSE precoding, which performs better than the four-color frequency reuse scheme in terms of energy efficiency (EE). In [12], a precoding technique for improving EE was proposed, where successive convex approximation (SCA)-based and ZF-based algorithms were compared. It was shown that the SCA-based algorithm surpassed the ZF-based one. The authors in [13] presented a linear precoding technique taking both linear and

nonlinear power constraints into account due to high power amplifiers. Their proposed technique achieved higher average throughput than ZF and MMSE precoding and showed stable performance with increasing the number of beams. In [14], a frame-based precoding technique and user scheduling policy were presented under DVB-S2X, which is the latest SatCom standard. The performance was investigated with respect to the objective functions and constraints. The combination of user scheduling and precoding scheme significantly alleviated the system performance degradation due to the increased number of users.

All the above studies have considered a single gateway. As the number of beams activated in the user link increases, the bandwidth required in the feeder link also expands. However, as the available bandwidth is limited, many works on exploiting multiple gateways have recently been suggested. Since the gateways utilize a very directional antenna, the bandwidth can be reused entirely in the feeder link. In [15], precoding based on regularized singular value decomposition was proposed. The performance of the cooperative schemes between gateways was investigated to reduce the amount of CSI exchanged. The authors in [16] calculated the precoding matrix with channel statistics to reduce the amount of channel information shared between different gateways. Since it is possible to increase the efficiency of frames and cover a large number of users with broad beams, multicast methods in the latest SatCom have been widely studied, where precoding is carried out for each frame transmitted to each beam. In [17] and [18], frame-based precoding was investigated at multiple gateways in a multicast scenario to support multiple users simultaneously.

Many works have been conducted on non-orthogonal multiple access (NOMA) in terrestrial communication to improve bandwidth efficiency. In general, NOMA refers to superimposing signals with different power levels in the same frequency and time resource and adopting successive interference cancellation (SIC) at the receivers to remove interference [19]. In fact, it is known that, in many respects, NOMA can achieve higher performance than OMA. NOMA improves spectral efficiency and user fairness because it can cover more users simultaneously than OMA [19], [20]. Since SatCom is fundamentally used to serve many users with wide coverage, NOMA can be one of the promising strategies in SatCom. In that sense, it is worth applying NOMA to SatCom, but it is not appropriate to apply the NOMA solutions designed for terrestrial communications to SatCom intactly. In terrestrial communication, a base station is located in each cell. It is obvious that users undergo different path loss attenuation depending on the distance from the base station, which leads to allocating the power resource for NOMA. However, in SatCom, since all beams are transmitted from one satellite, the complexity increases according to the number of beams. Moreover, users in the beams suffer similar path loss attenuation [21]. Therefore, algorithms for proper power allocation for superposition coding (SC) should be proposed in consideration of increasing complexity due to an increase in the number of beams. In [22], the applicability of NOMA to Satcom was shown and various application scenarios were presented. In particular, the introduction of NOMA in cognitive networks and integrated satellite-terrestrial networks (ISTN) has been described. In [23] and [24], NOMA was introduced for two users in a single beam, and the performance was analyzed in the downlink and uplink of SatCom, respectively, showing that NOMA can achieve a higher performance than OMA. While NOMA has also been introduced to enhance the capacity of ISTN, it has been applied to base stations rather than satellites [25], [26]. Moreover, in [27], several architectures that enabled NOMA to be implemented in SatCom were presented. Similar to our paper, NOMA schemes for each beam in multibeam satellite systems were suggested in [21], [28], and [29]. The conventional ZF precoding was adopted at a single gateway, and the performance was compared with regard to the users' decoding and scheduling methods in [21]. Up to recently, NOMA and multicast schemes were taken into account separately in SatCom. For the first time in [28], they were studied together for SatCom. The authors in [28] optimized power allocation and user clustering, and they found that the NOMA multicast proposed for a single gateway surpassed the OMA multicast. In [29], precoding was considered in the full-frequency reuse system. When calculating precoding in a multicast system, performance was shown depending on which channel vector was selected among multiple user channels. However, in [21], multicast was not handled, and in [28], four-frequency reuse was used, so inter-beam interference and precoding were not taken into consideration. Even if, in [29], full-frequency reuse was considered, only two groups were treated and precoding was performed simply using the MMSE technique. Although the NOMA systems for multiple users in multiple clusters have been considered in [30] and references therein, in this paper, NOMA is applied to multiple user groups in multiple beams. In [30], there are several users in multiple clusters, and each user receives superposed signals for users in the same cluster, where the clusters can be consistent with the beams of our system.

B. Contributions

In this paper, we study the application of NOMA to multigroup multicast systems in multibeam SatCom to employ the frequency band more efficiently and enhance the overall spectral efficiency. Even in terrestrial communication systems, NOMA has not been applied to multigroup multicast systems. Only one signal is generally transmitted per beam in multicast systems, although multiple signals can be transmitted to multiple groups in each beam simultaneously because of SC in the present work. Furthermore, we optimize the precoding for interference mitigation, power allocation coefficient for SC, and decoding order for SIC in multibeam SatCom.

- In our system, instead of simply distributing multiple users in the beams corresponding to the clusters in [30], multiple groups including multiple users exist in each beam. All the users in each group desire the same information and receive signals superposed by the number of groups with interference signals for other beams. While the users in the same beam receive identical superimposed signals, they can decode their own desired signals with

SIC. Even if there is a loss concerning the signal-to-noise-plus-interference ratio (SINR), it is possible to transmit multiple signals simultaneously, resulting in improved spectral efficiency and user accessibility. In this paper, unlike the previous papers on NOMA for SatCom, since full-frequency reuse is considered, precoding is essential to mitigate inter-beam interference. Moreover, SC and SIC should be optimized appropriately to make the best use of NOMA.

- To perform SIC, the aforementioned studies on NOMA SatCom fix the decoding order by limiting the condition of effective channel gain that the users receive when solving an optimization problem. It may not lead to the optimal precoding vector and power allocation to maximize the sum rate. Hence, it is necessary to take the decoding order into account. In [31], an exhaustive search was used to find the optimal decoding order among all the possible decoding orders. The exhaustive search is computationally efficient only when the search space is small. However, in our problem, exhaustive search is cumbersome to obtain the optimal decoding order in all beams because an optimization problem must be solved with many variables. Therefore, we solve the problem of optimizing the decoding order along with other variables such as precoding vectors and power allocation for SC to maximize the minimum rate for each signal under the satellite and gateway power constraints.

- Considering all the variables, the formulated problem is a non-convex mixed-integer nonlinear programming (MINLP) that is difficult to solve. Therefore, we adopt a block coordinate descent (BCD) method to handle this problem, which is efficient by updating each block variable [32]. This method sequentially approximates the non-convex function to the convex function at each iteration and updates each block variable. We show that it is more efficient in terms of convergence than solving all variables jointly.

C. Organization and Notations

The remainder of this paper is organized as follows. We first describe the system model under consideration in Section II. Section III provides the problem formulation. In Section IV, we present a solution to the problem. In Section V, the proposed technique is compared with the conventional technique through simulation. Finally, we conclude this paper in Section VI.

Throughout this paper, lowercase and uppercase letters in boldface represent vectors and matrices, respectively. The operators $|x|$, $\mathbb{E}[x]$, and $\Re\{x\}$ denote the absolute value, average value, and the real part of x, respectively. Further, $||\mathbf{x}||$ and $\mathbf{x}^H$ represent the Euclidean norm and the Hermitian transpose of vector $\mathbf{x}$, respectively. Finally, $x \sim \mathcal{N}\left(\mu, \sigma^2\right)$ denotes the Gaussian random variable x with mean μ and variance σ^2.
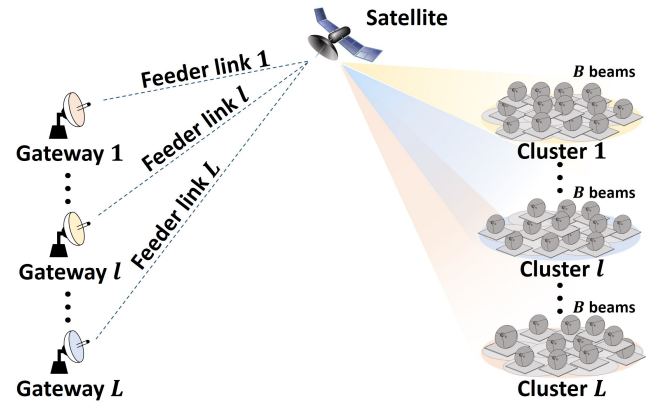

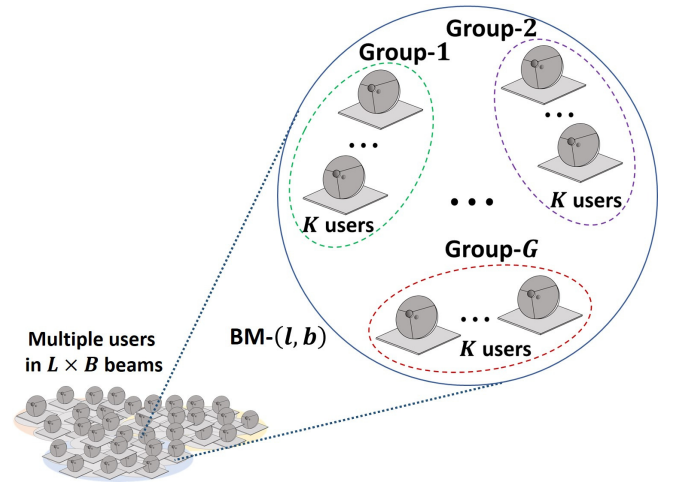Fig. 1. System model of the NOMA multigroup multicast with multiple gateways in SatCom.

Fig. 2. Each beam has G groups of K users.

II. SYSTEM MODEL

We consider a multibeam SatCom system in which a single satellite provides services to numerous users with full frequency reuse as shown in Fig. 1. The satellite receives signals from L gateways through feeder links and transfers them to the users through user links. There are L feeder link receivers in the satellite, and each gateway is assigned to each receiver to control a cluster of B beams. Besides, to transmit signals, the satellite is equipped with L antenna feed clusters. Each of them comprises B antenna feed elements so that each cluster serves a beam cluster, and each feed element creates one beam (i.e., single feed per beam). In each beam, GK users with a single antenna form G multicast groups of K users each as shown in Fig. 2. The users within a group desire the same multicast message. Since we take the NOMA system into consideration, the SC scheme for the power domain multiplexing is implemented to serve G multicast groups concurrently in each beam. At the users' end, accordingly, SIC is conducted to decode signals in decoding order.

On denoting $\mathcal{L} = \{1, 2, \cdots, L\}$, $\mathcal{B} = \{1, 2, \cdots, B\}$, $\mathcal{G} = \{1, 2, \cdots, G\}$, and $\mathcal{K} = \{1, 2, \cdots, K\}$, the beam from

Notation	Definition
L	Number of gateways / feeder link receivers / antenna feed clusters / beam clusters
B	Number of antenna feed elements / beams in a beam cluster
G	Number of multicast groups in a beam
K	Number of users in a group
BM-(l, b)	Beam from the bth feed element of the lth antenna feed cluster
UE-(l, b, g, k)	kth user of the gth multicast group in BM-(l, b)
$\mathbf{h}_{i,lbgk}^{H}$	Channel coefficient from the ith feed cluster to UE-(l, b, g, k)
G_T, G_R	Maximum satellite antenna gain and user receiver antenna gain
r_{lbgk}	Atmospheric fading of UE-(l, b, g, k)
$\mathbf{q}_{i,lbgk}$	Vector including the beam pattern and path loss
ξ_{lbgk}, ϕ_{lbgk}	Power gain and phase of rain attenuation
μ, σ	Lognormal location and scale of rain attenuation
λ	Carrier wavelength
$g_{ij,lbgk}$	Beam pattern
d_{lbgk}	Distance between the satellite and UE-(l, b, g, k)
$\kappa, T_{\mathrm{R}}, B_{\mathrm{ul}}$	Boltzmann constant, receiver noise temperature, and beamwidth
$\eta_{ij,lbgk}$	Phase variation
$\theta_{ij,lgbk}$	Angle between the center of BM-(i, j) and the user position
$\theta_{3\mathrm{dB}}$	Angle between 3 dB power loss and the beam center
s_{lbg}	Signal for the gth multicast group users in BM-(l, b)
s_{lb}	Signal for users in BM-(l, b)
α_{lbg}	Power allocation coefficient
$\mathbf{w}_{lb}$	Precoding vector for users in BM-(l, b)
n_{lbgk}	Additive white Gaussian noise
$a_{lb,ij}$	Binary variable for decoding order
SINR_{lbgk}^{m}	SINR of UE-(l, b, g, k) to decode the message for the mth group
P_l^{GW}	Maximum power at the lth gateway
P_{lv}^{Feed}	Maximum power at the vth feed element of lth antenna feed cluster

the bth feed element of the lth antenna feed cluster is denoted as BM-(l, b) and the kth user of the gth multicast group in BM-(l, b) is represented by UE-(l, b, g, k), where $l \in \mathcal{L}$, $b \in \mathcal{B}$, $g \in \mathcal{G}$, and $k \in \mathcal{K}$. The lth gateway transmits signals to the lth feeder link receiver connected to the lth antenna feed cluster in the satellite. We assume that perfect CSI is available at the gateways [13], [14]. However, in Section V, we carry out simulations on the effect of imperfect CSI as well. The main variables used in this paper are given in Table I for ease of reading.

A. Channel Model

We assume that the feeder links between the gateways and satellite are ideal because the gateways exploit very directive antennas to transmit signals and the SNR of the feeder link is much greater than that of the user link [15], [33]. Therefore, the noise and interference in the feeder links can be negligible.

Thus, $\mathbf{h}_{i,lbgk}^{H} \in \mathbb{C}^{1 \times B}$ represents the channel coefficient from the ith feed cluster to UE-(l, b, g, k) and can be decomposed as

$$\mathbf{h}_{i,lbgk} = \sqrt{G_T G_R} r_{lbgk} \mathbf{q}_{i,lbgk}, \tag{1}$$

where G_T and G_R are the maximum satellite antenna gain and user receiver antenna gain, respectively, and $\mathbf{q}_{i,lbgk}$ is a vector including the beam pattern and path loss. The parameter r_{lbgk} denotes the atmospheric fading of UE-(l, b, g, k) because the satellite channels using high-frequency bands, such as Ka-band, are significantly affected by atmospheric fading effects. Among them, rain attenuation is the most dominant factor. The atmospheric fading of UE-(l, b, g, k) due to rain attenuation is given as [34] and [35]

$$r_{lbgk} = \xi_{lbgk}^{-\frac{1}{2}} e^{j\phi_{lbgk}}, \tag{2}$$

where ξ_{lbgk} denotes the power gain of rain attenuation. The power gain in dB, $\xi_{lbgk}^{\mathrm{dB}} = 20 \log_{10}(\xi_{lbgk})$, follows a lognormal distribution, i.e., $\ln\left(\xi_{lbgk}^{\mathrm{dB}}\right) \sim \mathcal{N}(\mu, \sigma^2)$, where μ and σ are the lognormal location and scale parameters, respectively, besides ϕ_{lbgk} is the uniformly distributed phase over $[0, 2\pi)$. Furthermore, $\mathbf{q}_{i,lbgk}$ is defined by $\mathbf{q}_{i,lbgk} = [q_{i1,lbgk} \cdots q_{ij,lbgk} \cdots q_{iB,lbgk}]^T$, where $q_{ij,lbgk}$ can be written as [35] and [18]

$$q_{ij,lbgk} = \frac{\lambda}{4\pi d_{lbgk}} g_{ij,lbgk} e^{j\eta_{ij,lbgk}}, \tag{3}$$

where λ is the carrier wavelength and d_{lbgk} is the distance between the satellite and UE-(l, b, g, k). In (3), $\eta_{ij,lbgk}$ denotes phase variation due to the beam radiation pattern and radio wave propagation, where we assume that the satellite has stable oscillators [33]. The beam pattern $g_{ij,lbgk}$ depends on the user position and satellite antenna pattern. When we define $\theta_{ij,lgbk}$ as the angle between the center of BM-(i, j) and the user position and $\theta_{3\mathrm{dB}}$ as the angle 3 dB power loss from the beam center, $g_{ij,lbgk}$ can be approximated as [13] and [34]

$$g_{ij,lbgk} = \left| \frac{J_1(u_{ij,lgbk})}{2u_{ij,lgbk}} + 36 \frac{J_3(u_{ij,lgbk})}{u_{ij,lgbk}^3} \right|, \tag{4}$$

where $J_1(\cdot)$ and $J_3(\cdot)$ are the first-kind Bessel functions of order 1 and 3, respectively, and $u_{ij,lgbk} = 2.07123 \sin(\theta_{ij,lgbk}) / \sin(\theta_{3\mathrm{dB}})$.

B. Signal Model

Assuming that the message s_{lbg} is intended for the gth multicast group users in BM-(l, b), the signals for BM-(l, b) are superimposed using the NOMA scheme at the lth gateway. It can be written as $s_{lb} = \sum_{g=1}^{G} \sqrt{\alpha_{lbg}} s_{lbg}$, where $0 \leq \alpha_{lbg} \leq 1$ is the power allocation coefficient and $\sum_{g=1}^{G} \alpha_{lbg} = 1$. The signal s_{lbg} has unit average power $\left(\text{i.e., } \mathbb{E}\left[|s_{lbg}|^2\right] = 1\right)$. Thus, the signal for the lth cluster $\mathbf{s}_l$ can be expressed as $\mathbf{s}_l = [s_{l1} \; s_{l2} \; \cdots \; s_{lB}]^T$. Before transmission, $\mathbf{s}_l$ is precoded with matrix $\mathbf{W}_l \in \mathbb{C}^{B \times B}$ at the lth gateway to mitigate intra-cluster and inter-cluster interference. The transmitted signal $\mathbf{x}_l$ can be written as

$$\mathbf{x}_l = \mathbf{W}_l \mathbf{s}_l = \sum_{b=1}^{B} \mathbf{w}_{lb} \sum_{g=1}^{G} \sqrt{\alpha_{lbg}} s_{lbg}, \tag{5}$$

where $\mathbf{w}_{lb} \in \mathbb{C}^{B \times 1}$ is a precoding vector for users in M-(l, b).

The lth feeder link receiver receives the precoded signals in the satellite via the feeder link. Since we consider a transparent payload as in [13]–[15] and [33], the satellite transfers the signals via the antenna feed clusters as a relay. Therefore, the received signal at UE-(l, b, g, k) can be written as

$$y_{lbgk} = \sum_{i=1}^{L} \mathbf{h}_{i,lbgk}^{H} \mathbf{x}_i + n_{lbgk}, \tag{6}$$

where n_{lbgk} denotes the additive white Gaussian noise (AWGN) with zero mean and variance σ_{lbgk}^2. To clarify the interference due to the SC scheme, intra-cluster interference, and inter-cluster interference, y_{lbgk} can be rewritten as

$$y_{lbgk} = \underbrace{\mathbf{h}_{l,lbgk}^{H} \mathbf{w}_{lb} \sqrt{\alpha_{lbg}} s_{lbg}}_{\text{desired signal}} + \underbrace{\mathbf{h}_{l,lbgk}^{H} \mathbf{w}_{lb} \sum_{m \neq g}^{G} \sqrt{\alpha_{lbm}} s_{lbm}}_{\text{intra-beam (inter-group) interference}}$$

$$+ \underbrace{\mathbf{h}_{l,lbgk}^{H} \sum_{j \neq b}^{B} \mathbf{w}_{lj} s_{lj}}_{\text{intra-cluster (inter-beam) interference}} + \underbrace{\sum_{i \neq l}^{L} \sum_{j=1}^{B} \mathbf{h}_{i,lbgk}^{H} \mathbf{w}_{ij} s_{ij}}_{\text{inter-cluster interference}} + n_{lbgk}. \tag{7}$$

For each beam, some multicast group users adopt SIC to detect the signals. When users in a multicast group directly decode the desired signal without SIC, users in the other multicast groups exploit SIC. The users employing SIC decode the signals intended for other groups, remove them from interference, and obtain the signal desired for themselves. To determine the multicast groups using SIC (i.e., decoding order) in each beam, we introduce binary variable $a_{lb,ij} \in \{0,1\}$, where $a_{lb,ij} = 0$ indicates that the jth multicast group users in BM-(l, b) take advantage of SIC and decode the signal for the ith group, otherwise $a_{lb,ij} = 1$. Note that $a_{lb,ij} + a_{lb,ji} = 1$ and $a_{lb,ii} = 0, i \neq j, \forall i, j$. Hence, the SINR of UE-$(l, b, g, k)$ to decode the message for the mth group in BM-(l, b) can be expressed as

$$\text{SINR}_{lbgk}^m (\mathcal{W}, \mathcal{P}_{lb}, \mathcal{A}_{lb})$$

$$= \frac{\left| \mathbf{h}_{l,lbgk}^{H} \mathbf{w}_{lb} \right|^2 \alpha_{lbm} + a_{lb,mg} M}{\sum_{u \neq m}^{G} a_{lb,um} \left| \mathbf{h}_{l,lbgk}^{H} \mathbf{w}_{lb} \right|^2 \alpha_{lbu} + I_{lbgk}(\mathcal{W}) + \sigma_{lbgk}^2}, \tag{8}$$

where $\mathcal{W} \triangleq \{\mathbf{w}_{lb}, \forall l \in \mathcal{L}, \forall b \in \mathcal{B}\}$, $\mathcal{P}_{lb} \triangleq \{\alpha_{lbg}, \forall g \in \mathcal{G}\}$, and $\mathcal{A}_{lb} \triangleq \{a_{lb,ij}, \forall i, j \in \mathcal{G}\}$. An arbitrary large positive number M is introduced to exclude the SINR that does not match the decoding order. In other words, if $a_{lb,mg} = 1$, SINR_{lbgk}^m is not calculated according to the decoding order and is not selected in multicast systems. In (8), $I_{lbgk}(\mathcal{W})$ is the sum of intra-cluster and inter-cluster interference at UE-(l, b, g, k), which can be written as $I_{lbgk}(\mathcal{W}) = \sum_{j \neq b}^{B} \left| \mathbf{h}_{l,lbgk}^{H} \mathbf{w}_{lj} \right|^2 + \sum_{i \neq l}^{L} \sum_{j=1}^{B} \left| \mathbf{h}_{i,lbgk}^{H} \mathbf{w}_{ij} \right|^2$.

III. PROBLEM FORMULATION

Our objective is to develop a transmission scheme that maximizes the sum rate of multigroup multicast transmission in the SatCom system with the NOMA scheme

while satisfying perfect SIC and power constraints. Accordingly, we propose a method to optimize the precoding vectors $\mathcal{W}$, the power allocation coefficients of SC $\mathcal{P} \triangleq \{\mathcal{P}_{lb}, \forall l \in \mathcal{L}, \forall b \in \mathcal{B}\}$, and the selection of multicast groups to employ SIC $\mathcal{A} \triangleq \{\mathcal{A}_{lb}, \forall l \in \mathcal{L}, \forall b \in \mathcal{B}\}$. To successfully perform SIC in the NOMA multigroup multicast system [36], the optimization problem can be formulated as

$$\max_{\mathcal{W}, \mathcal{P}, \mathcal{A}} \sum_{l=1}^{L} \sum_{b=1}^{B} \sum_{m=1}^{G} \log_2 \left(1 + \min_{g \in \mathcal{G}, k \in \mathcal{K}} \text{SINR}_{lbgk}^m (\mathcal{W}, \mathcal{P}_{lb}, \mathcal{A}_{lb}) \right) \tag{9a}$$

$$\text{s.t.} \quad \sum_{b=1}^{B} \mathbf{w}_{lb}^H \mathbf{w}_{lb} \leq P_l^{\text{GW}}, \tag{9b}$$

$$\sum_{b=1}^{B} \mathbf{w}_{lb}^H \mathbf{Q}_v \mathbf{w}_{lb} \leq P_{lv}^{\text{Feed}}, \tag{9c}$$

$$0 \leq \alpha_{lbg} \leq 1, \ \sum_{g=1}^{G} \alpha_{lbg} = 1, \tag{9d}$$

$$a_{lb,ij} \in \{0,1\}, \ a_{lb,ij} + a_{lb,ji} = 1, \ a_{lb,ii} = 0, \tag{9e}$$

for $\forall l \in \mathcal{L}$, $\forall b, v \in \mathcal{B}$, and $\forall g, i, j \in \mathcal{G}$. P_l^{GW} and P_{ln}^{Feed} are the maximum power at the lth gateway and the nth feed element of lth antenna feed cluster, respectively. In (9c), $\mathbf{Q}_v$ is a matrix whose vth diagonal element is 1 and the other elements are 0. Problem (9) is a non-convex problem due to (9a) and the binary constraints of (9e) although the others are convex constraints. It is difficult to find a globally optimal solution for this problem. Therefore, in the next section, we propose a method to obtain a local optimum by decomposing (9) and approximating it to convex problems.

IV. PROPOSED SOLUTION

Since (9) is a non-convex MINLP, we suggest designing an algorithm based on BCD to solve it efficiently. To approximate the problem in each iteration, we need to transform it into a more amenable form. Thus, we introduce two sets of variables: $\Gamma \triangleq \{\gamma_{lb}^m, \forall l \in \mathcal{L}, \forall b \in \mathcal{B}, \forall m \in \mathcal{G}\}$ and $\mathcal{R} \triangleq \{R_{lb}^m, \forall l \in \mathcal{L}, \forall b \in \mathcal{B}, \forall m \in \mathcal{G}\}$, and then (9) can be transformed into

$$\max_{\mathcal{W}, \mathcal{P}, \mathcal{A}, \Gamma, \mathcal{R}} \sum_{l=1}^{L} \sum_{b=1}^{B} \sum_{m=1}^{G} R_{lb}^m \tag{10a}$$

$$\text{s.t.} \quad \sum_{b=1}^{B} \mathbf{w}_{lb}^H \mathbf{w}_{lb} \leq P_l^{\text{GW}}, \tag{10b}$$

$$\sum_{b=1}^{B} \mathbf{w}_{lb}^H \mathbf{Q}_v \mathbf{w}_{lb} \leq P_{lv}^{\text{Feed}}, \tag{10c}$$

$$0 \leq \alpha_{lbg} \leq 1, \ \sum_{g=1}^{G} \alpha_{lbg} = 1, \tag{10d}$$

$$a_{lb,ij} \in \{0,1\}, \ a_{lb,ij} + a_{lb,ji} = 1, a_{lb,ii} = 0, \tag{10e}$$

$$\text{SINR}_{lbgk}^m (\mathbf{w}, \boldsymbol{\alpha}_{lb}, \mathbf{a}_{lb}) \geq \gamma_{lb}^m, \tag{10f}$$

$$\log_2 (1 + \gamma_{lb}^m) \geq R_{lb}^m, \tag{10g}$$

for $\forall l \in \mathcal{L}$, $\forall b, v \in \mathcal{B}$, $\forall g, m, i, j \in \mathcal{G}$, and $\forall k \in \mathcal{K}$. Since (10f) and (10g) hold equality at the optimal condition, the equivalence of (9) and (10) is guaranteed and can be easily proved by contradiction.

Even if the (10a) is relaxed to a convex function, (10) is still non-convex owing to (10e) and (10f). To solve the problem effectively, we propose decomposing (10) into three subproblems and solving them by approximating the non-convex constraints iteratively with the help of the BCD algorithm. In other words, it is divided into (i) a problem for precoding vector with power allocation and decoding order fixed, (ii) a problem for power allocation with precoding vector and decoding order fixed, and (iii) a problem for decoding order with precoding vector and power allocation fixed. The three subproblems are alternately solved utilizing SCA until they converge.

A. Precoding Optimization

Given $\mathcal{P}$ and $\mathcal{A}$, the optimization problem for $\mathcal{W}$ can be written as

$$\max_{\mathcal{W},\Gamma,\mathcal{R}} \sum_{l=1}^{L} \sum_{b=1}^{B} \sum_{m=1}^{G} R_{lb}^m \tag{11a}$$

$$\text{s.t. SINR}_{lbgk}^m \left(\mathcal{W}, \mathcal{P}_{lb}^{(n)}, \mathcal{A}_{lb}^{(n)}\right) \geq \gamma_{lb}^m, \tag{11b}$$

$$(10\text{b}), (10\text{c}), (10\text{g}), \tag{11c}$$

$$\forall l \in \mathcal{L}, \ \forall b, v \in \mathcal{B}, \ \forall g, m \in \mathcal{G}, \ \forall k \in \mathcal{K},$$

where $\mathcal{P}_{lb}^{(n)}$ and $\mathcal{A}_{lb}^{(n)}$ denote the solutions obtained at the nth iteration. The non-convex constraint (11b) can be expressed as

$$\frac{\left|\mathbf{h}_{l,lbgk}^H \mathbf{w}_{lb}\right|^2 \alpha_{lbm}^{(n)}}{\gamma_{lb}^m} + \frac{a_{lb,mg}^{(n)} M}{\gamma_{lb}^m}$$

$$\geq \sum_{u \neq m}^{G} a_{lb,um}^{(n)} \left|\mathbf{h}_{l,lbgk}^H \mathbf{w}_{lb}\right|^2 \alpha_{lbu}^{(n)} + I_{lbgk}(\mathcal{W}) + \sigma_{lbgk}^2, \tag{12}$$

where the right-hand side is a convex function for $\mathbf{w}$. The left-hand side is jointly convex for $\mathbf{w}_{lb}$ and γ_{lb}^m because it is a quadratic-over-linear function [37]. When we define $f_{lbgk,m}\left(\mathbf{w}_{lb}, \gamma_{lb}^m\right) \triangleq \left|\mathbf{h}_{l,lbgk}^H \mathbf{w}_{lb}\right|^2 \alpha_{lbm}^{(n)} / \gamma_{lb}^m$ and $g_{lb,mg}\left(\gamma_{lb}^m\right) \triangleq a_{lb,mg}^{(n)} M / \gamma_{lb}^m$, they can be lower bounded with the first-order Taylor series expansion around $\mathbf{w}_{lb}^{(n)}$ and $(\gamma_{lb}^m)^{(n)}$, which can be expressed as

$$f_{lbgk,m}\left(\mathbf{w}_{lb}, \gamma_{lb}^m\right) \geq F_{lbgk,m}\left(\mathbf{w}_{lb}, \gamma_{lb}^m; \mathbf{w}_{lb}^{(n)}, (\gamma_{lb}^m)^{(n)}\right)$$

$$\triangleq \alpha_{lbm}^{(n)} \left[\frac{2\Re\left\{\left(\mathbf{w}_{lb}^{(n)}\right)^H \mathbf{h}_{l,lbgk}\mathbf{h}_{l,lbgk}^H \mathbf{w}_{lb}\right\}}{(\gamma_{lb}^m)^{(n)}} \right.$$

$$\left. - \frac{\left(\mathbf{w}_{lb}^{(n)}\right)^H \mathbf{h}_{l,lbgk}\mathbf{h}_{l,lbgk}^H \mathbf{w}_{lb}^{(n)}}{\left\{(\gamma_{lb}^m)^{(n)}\right\}^2} \gamma_{lb}^m \right], \tag{13}$$

$$g_{lb,mg}\left(\gamma_{lb}^m\right) \geq G_{lb,mg}\left(\gamma_{lb}^m; (\gamma_{lb}^m)^{(n)}\right)$$

$$\triangleq a_{lb,mg}^{(n)} M \left[\frac{2}{(\gamma_{lb}^m)^{(n)}} - \frac{\gamma_{lb}^m}{\left((\gamma_{lb}^m)^{(n)}\right)^2} \right], \tag{14}$$

where $\mathbf{w}_{lb}^{(n)}$ and $(\gamma_{lb}^m)^{(n)}$ are the solutions from the nth iteration. Hence, (12) can be convexified as

$$F_{lbgk,m}\left(\mathbf{w}_{lb}, \gamma_{lb}^m; \mathbf{w}_{lb}^{(n)}, (\gamma_{lb}^m)^{(n)}\right) + G_{lb,mg}\left(\gamma_{lb}^m; (\gamma_{lb}^m)^{(n)}\right)$$

$$\geq \sum_{u \neq m}^{G} a_{lb,um}^{(n)} \left|\mathbf{h}_{l,lbgk}^H \mathbf{w}_{lb}\right|^2 \alpha_{lbm}^{(n)} + I_{lbgk}(\mathcal{W}) + \sigma_{lbgk}^2. \tag{15}$$

Although (10g) is a convex constraint, because log functions are not preferred in optimization solvers (e.g., MOSEK), we change it to a more computationally efficient form [38]. First, we have

$$\gamma_{lb}^m \log_2\left(1 + \gamma_{lb}^m\right) \geq \gamma_{lb}^m R_{lb}^m. \tag{16}$$

Then, by applying the first-order Taylor series expansion, the convex function on the left-hand side can be lower-bounded as

$$\gamma_{lb}^m \log_2\left(1 + \gamma_{lb}^m\right) \geq (\tau_{lb}^m)^{(n)} \gamma_{lb}^m - (\upsilon_{lb}^m)^{(n)}, \tag{17}$$

where $(\tau_{lb}^m)^{(n)}$ and $(\upsilon_{lb}^m)^{(n)}$ represent

$$(\tau_{lb}^m)^{(n)} \triangleq \frac{1}{\ln 2} \left\{ \ln\left(1 + (\gamma_{lb}^m)^{(n)}\right) + \frac{(\gamma_{lb}^m)^{(n)}}{\left(1 + (\gamma_{lb}^m)^{(n)}\right)} \right\}, \tag{18a}$$

$$(\upsilon_{lb}^m)^{(n)} \triangleq \frac{\left((\gamma_{lb}^m)^{(n)}\right)^2}{\ln 2 \left(1 + (\gamma_{lb}^m)^{(n)}\right)}. \tag{18b}$$

Although the left-hand side in (16) is lower bounded as (17), (16) is non-convex owing to the right-hand side. However, we can convert (16) equivalently into a second-order cone constraint, which can be expressed as

$$\gamma_{lb}^m - R_{lb}^m + (\tau_{lb}^m)^{(n)} \geq \left\| \left[\gamma_{lb}^m + R_{lb}^m - (\tau_{lb}^m)^{(n)} \quad 2\sqrt{(\upsilon_{lb}^m)^{(n)}} \right] \right\|_2. \tag{19}$$

This constraint can be utilized equally in the other two divided optimization problems. Therefore, (11) can be rewritten as

$$\max_{\mathcal{W},\Gamma,\mathcal{R}} \sum_{l=1}^{L} \sum_{b=1}^{B} \sum_{m=1}^{G} R_{lb}^m \tag{20a}$$

$$\text{s.t. } (10\text{b}), (10\text{c}), (15), (19), \tag{20b}$$

$$\forall l \in \mathcal{L}, \ \forall b, v \in \mathcal{B}, \ \forall g, m \in \mathcal{G}, \ \forall k \in \mathcal{K},$$

which is a second-order cone programming (SOCP) and can be solved by standard convex optimization software.

B. Power Allocation Optimization

Given $\mathcal{W}$ and $\mathcal{A}$, the optimization problem to obtain $\mathcal{P}$ can be written as

$$\max_{\mathcal{P},\Gamma,\mathcal{R}} \sum_{l=1}^{L}\sum_{b=1}^{B}\sum_{m=1}^{G} R_{lb}^{m} \tag{21a}$$

$$\text{s.t. } \mathrm{SINR}_{lbgk}^{m}\left(\mathcal{W}^{(n)},\mathcal{P}_{lb},\mathcal{A}_{lb}^{(n)}\right) \geq \gamma_{lb}^{m}, \tag{21b}$$

$$(10d),(19), \tag{21c}$$

$$\forall l \in \mathcal{L},\ \forall b,v \in \mathcal{B},\ \forall g,m \in \mathcal{G},\ \forall k \in \mathcal{K},$$

where $\mathcal{W}^{(n)}$ represents the solutions obtained at the nth iteration. We can rewrite (21b) as

$$\left|\mathbf{h}_{l,lbgk}^{H}\mathbf{w}_{lb}^{(n)}\right|^{2}\alpha_{lbm} + a_{lb,mg}^{(n)}M$$
$$\geq \sum_{u\neq m}^{G} a_{lb,um}^{(n)}\left|\mathbf{h}_{l,lbgk}^{H}\mathbf{w}_{lb}^{(n)}\right|^{2}\gamma_{lb}^{m}\alpha_{lbu} + \gamma_{lb}^{m}\left(I_{lbgk}\left(\mathcal{W}^{(n)}\right)+\sigma_{lbgk}^{2}\right), \tag{22}$$

which is non-convex owing to $\gamma_{lb}^{m}\alpha_{lbu}$ in the first term on the right hand side. When we take advantage of simple algebraic operations and first-order Taylor series expansion around $\left((\gamma_{lb}^{m})^{(n)},\alpha_{lbu}^{(n)}\right)$, it can be upper bounded as

$$\gamma_{lb}^{m}\alpha_{lbu} \leq H_{1,lbu}^{m}\left(\gamma_{lb}^{m},\alpha_{lbu};(\gamma_{lb}^{m})^{(n)},\alpha_{lbu}^{(n)}\right)$$
$$\triangleq \frac{1}{4}\left(\gamma_{lb}^{m}+\alpha_{lbu}\right)^{2} - \frac{1}{4}\left((\gamma_{lb}^{m})^{(n)}-\alpha_{lbu}^{(n)}\right)^{2}$$
$$- \frac{1}{2}\left((\gamma_{lb}^{m})^{(n)}-\alpha_{lbu}^{(n)}\right)\left(\gamma_{lb}^{m}-\alpha_{lbu}-(\gamma_{lb}^{m})^{(n)}+\alpha_{lbu}^{(n)}\right). \tag{23}$$

Therefore, (21b) can be convexified as

$$\left|\mathbf{h}_{l,lbgk}^{H}\mathbf{w}_{lb}^{(n)}\right|^{2}\alpha_{lbm} + a_{lb,mg}^{(n)}M$$
$$\geq \sum_{u\neq m}^{G} a_{lb,um}^{(n)}\left|\mathbf{h}_{l,lbgk}^{H}\mathbf{w}_{lb}^{(n)}\right|^{2}H_{1,lbu}^{m}\left(\gamma_{lb}^{m},\alpha_{lbu};(\gamma_{lb}^{m})^{(n)},\alpha_{lbu}^{(n)}\right)$$
$$+ \gamma_{lb}^{m}\left(I_{lbgk}\left(\mathcal{W}^{(n)}\right)+\sigma_{lbgk}^{2}\right), \tag{24}$$

The relaxed problem for $\mathcal{P}$ can be rewritten as

$$\max_{\mathcal{P},\Gamma,\mathcal{R}} \sum_{l=1}^{L}\sum_{b=1}^{B}\sum_{m=1}^{G} R_{lb}^{m} \tag{25a}$$

$$\text{s.t. } (10d),(19),(24), \tag{25b}$$

$$\forall l \in \mathcal{L},\ \forall b,v \in \mathcal{B},\ \forall g,m \in \mathcal{G},\ \forall k \in \mathcal{K}, \tag{25c}$$

which is also an SOCP.

C. Decoding Order Optimization

Finally, given $\mathcal{W}$ and $\mathcal{P}$, the optimization problem for $\mathcal{A}$ can be formulated as

$$\max_{\mathcal{A},\Gamma,\mathcal{R}} \sum_{l=1}^{L}\sum_{b=1}^{B}\sum_{m=1}^{G} R_{lb}^{m} \tag{26a}$$

$$\text{s.t. } \mathrm{SINR}_{lbgk}^{m}\left(\mathcal{W}^{(n)},\mathcal{P}_{lb}^{(n)},\mathcal{A}_{lb}\right) \geq \gamma_{lb}^{m}, \tag{26b}$$

$$(10e),(19), \tag{26c}$$

$$\forall l \in \mathcal{L},\ \forall b,v \in \mathcal{B},\ \forall g,m,i,j \in \mathcal{G},\ \forall k \in \mathcal{K},$$

where (26b) and (10e) are non-convex constraints. The constraint (26b) can be expressed as

$$\left|\mathbf{h}_{l,lbgk}^{H}\mathbf{w}_{lb}^{(n)}\right|^{2}\alpha_{lbm}^{(n)} + a_{lb,mg}M$$
$$\geq \sum_{u\neq m}^{G}\gamma_{lb}^{m}a_{lb,um}\left|\mathbf{h}_{l,lbgk}^{H}\mathbf{w}_{lb}^{(n)}\right|^{2}\alpha_{lbu}^{(n)} + \gamma_{lb}^{m}\left(I_{lbgk}\left(\mathcal{W}^{(n)}\right)+\sigma_{lbgk}^{2}\right), \tag{27}$$

which can be relaxed similar to (24) as

$$\left|\mathbf{h}_{l,lbgk}^{H}\mathbf{w}_{lb}^{(n)}\right|^{2}\alpha_{lbm}^{(n)} + a_{lb,mg}M$$
$$\geq \sum_{u\neq m}^{G} H_{2,lbu}^{m}\left(\gamma_{lb}^{m},a_{lb,um};(\gamma_{lb}^{m})^{(n)},a_{lb,um}^{(n)}\right)\left|\mathbf{h}_{l,lbgk}^{H}\mathbf{w}_{lb}^{(n)}\right|^{2}\alpha_{lbu}^{(n)}$$
$$+ \gamma_{lb}^{m}\left(I_{lbgk}\left(\mathcal{W}^{(n)}\right)+\sigma_{lbgk}^{2}\right), \tag{28}$$

where $H_{2,lbu}^{m}\left(\gamma_{lb}^{m},a_{lb,um};(\gamma_{lb}^{m})^{(n)},a_{lb,um}^{(n)}\right)$ is represented as

$$\gamma_{lb}^{m}a_{lb,um} \leq H_{2,lbu}^{m}\left(\gamma_{lb}^{m},a_{lb,um};(\gamma_{lb}^{m})^{(n)},a_{lb,um}^{(n)}\right)$$
$$\triangleq \frac{1}{4}\left(\gamma_{lb}^{m}+a_{lb,um}\right)^{2} - \frac{1}{4}\left((\gamma_{lb}^{m})^{(n)}-a_{lb,um}^{(n)}\right)^{2}$$
$$- \frac{1}{2}\left((\gamma_{lb}^{m})^{(n)}-a_{lb,um}^{(n)}\right)\left(\gamma_{lb}^{m}-a_{lb,um}-(\gamma_{lb}^{m})^{(n)}+a_{lb,um}^{(n)}\right). \tag{29}$$

In (10e), the binary constraint $a_{lb,ij} \in \{0,1\}$ can be transformed into an equivalently continuous form as follows [39]:

$$0 \leq a_{lb,ij} \leq 1, \tag{30a}$$

$$a_{lb,ij} - a_{lb,ij}^{2} \leq 0. \tag{30b}$$

Since (30b) is a form of difference-convex function, we can take advantage of the first-order Taylor series expansion around $a_{lb,ij}^{(n)}$ to obtain convex approximation. It is given as

$$a_{lb,ij} - 2a_{lb,ij}^{(n)}a_{lb,ij} + \left(a_{lb,ij}^{(n)}\right)^{2} \leq 0. \tag{31}$$

However, if (31) is deployed directly, the problem becomes infeasible and fails to be optimized. Hence, a penalty parameter $\rho > 0$ and slack variable $\Lambda = \{\lambda_{lb,ij} \geq 0, \forall l \in \mathcal{L}, \forall b \in \mathcal{B}, \forall i,j \in \mathcal{G}\}$ are introduced [40]. Therefore, the approximated problem for $\mathbf{a}$ can be expressed as

$$\max_{\mathcal{A},\Gamma,\mathcal{R}} \sum_{l=1}^{L}\sum_{b=1}^{B}\sum_{m=1}^{G} R_{lb}^{m} - \rho\sum_{l=1}^{L}\sum_{b=1}^{B}\sum_{i=1}^{G}\sum_{j=1}^{G} \lambda_{lb,ij} \tag{32a}$$

$$\text{s.t. } a_{lb,ij} - 2a_{lb,ij}^{(n)}a_{lb,ij} + \left(a_{lb,ij}^{(n)}\right)^{2} \leq \lambda_{lb,ij}, \tag{32b}$$

$$a_{lb,ij} + a_{lb,ji} = 1,\ a_{lb,ii} = 0, \tag{32c}$$

$$(19),(28),(30a), \tag{32d}$$

$$\forall l \in \mathcal{L},\ \forall b \in \mathcal{B},\ \forall g,m,i,j \in \mathcal{G},\ \forall k \in \mathcal{K},$$

where (32a) and (32b) are equivalent to (26a) and (31), respectively, when $\lambda_{lb,ij} = 0$. (32) is also an SOCP as (20) and (25).

The algorithm based on BCD for solving (10) is summarized in Algorithm 1. For the relaxation of (32b), the initial value of

Algorithm 1 The proposed iterative algorithm to solve (10)

1: Set $n := 0$;
2: Initialization: Feasible starting points of $\mathcal{W}^{(n)}$, $\mathcal{P}^{(n)}$, $\mathcal{A}^{(n)}$, and $\Gamma^{(n)}$ and penalty parameter $\rho^{(n)}$;
3: **repeat**
4:     Solve (20) to obtain optimal solution $\mathcal{W}^*$ and Γ^* for given $\mathcal{P}^{(n)}$, $\mathcal{A}^{(n)}$, and $\Gamma^{(n)}$;
5:     Update $\mathcal{W}^{(n)} := \mathcal{W}^*$ and $\Gamma^{(n)} := \Gamma^*$;
6:     Solve (25) to obtain optimal solution $\mathcal{P}^*$ and Γ^* for given $\mathcal{W}^{(n)}$, $\mathcal{A}^{(n)}$, and $\Gamma^{(n)}$;
7:     Update $\mathcal{P}^{(n)} := \mathcal{P}^*$ and $\Gamma^{(n)} := \Gamma^*$;
8:     Solve (32) to obtain optimal solution $\mathcal{A}^*$ and Γ^* for given $\mathcal{W}^{(n)}$, $\mathcal{P}^{(n)}$, and $\Gamma^{(n)}$;
9:     Set $n := n + 1$;
10:    Update $\mathcal{W}^{(n)} := \mathcal{W}^*$, $\mathcal{P}^{(n)} := \mathcal{P}^*$, $\mathcal{A}^{(n)} := \mathcal{A}^*$, and $\Gamma^{(n)} := \Gamma^*$;
11:    Update $\rho^{(n)} := \min\left(\epsilon\rho^{(n-1)}, \rho_{\max}\right)$;
12: **until** Convergence;
13: Output $\mathcal{W}^{(n)}$, $\mathcal{P}^{(n)}$, and $\mathcal{A}^{(n)}$.

Fig. 3. Beam topology with multiple clusters and beams for $L = 3$ and $B = 7$.

ρ is set small, but as in [40], it increases by the constant $\epsilon > 1$ until $\sum_{l=1}^{L}\sum_{b=1}^{B}\sum_{i=1}^{G}\sum_{j=1}^{G}\lambda_{lb,ij} \approx 0$. The convergence of Algorithm 1 including $\rho_{\max}$ can be proven in a similar to [40]. If there are $\rho_{\max}$ and n_1 that satisfy $\sum_{l,b,i,j}\left|\lambda_{lb,ij}^{(n)}\right| = 0$ at the nth iteration larger than n_1, then $\sum_{l,b,m} R_{lb}^m$ converges. Assuming $p_{lb,ij}^{(n)}$ and $q_{lb,ij}^{(n)}$ are Lagrangian multipliers for (32b) and $\lambda_{lb,ij} \geq 0$, the following expressions can be obtained by the Karush-Kuhn-Tucker (KKT) conditions [37], $\rho^{(n)} - p_{lb,ij}^{(n)} - q_{lb,ij}^{(n)} = 0$ and $q_{lb,ij}^{(n)}\lambda_{lb,ij}^{(n)} = 0$. Besides, $\sum_{l,b,i,j}\left|p_{lb,ij}^{(n)}\right|$ can be upper bounded [41] and if $\rho_{\max}$ is greater than the upper bound, n_1 satisfying $p_{lb,ij}^{(n)} < \rho^{(n)}$ exists for $\forall n > n_1$. Therefore, $q_{lb,ij}^{(n)} > 0$ and $\lambda_{lb,ij}^{(n)} = 0$ for $\forall n > n_1$. When $n < n_1$, $\rho^{(n)}$ increases at each iteration. Accordingly, (32) converges and the minimum possible $\lambda^{(n)}$ is determined. The solution to (32) is also feasible in the next iteration for (20). When $n > n_1$, since $\rho^{(n)}$ is fixed and $\lambda^{(n)} = 0$, it is a general BCD-based algorithm. Also, since it is bounded by power constraints, Algorithm 1 converges after more than n_1 iterations.

The complexity of Algorithm 1 based on the BCD method lies in the complexity of subproblems. We note that the subproblems are SOCPs with lower computational complexity compared to other nonlinear problems. The complexity of SOCP is related to the number of variables and the number of constraints [42]. By using an interior-point method, the numbers of iterations for the subproblems in Section IV-A, IV-B, and IV-C to reach an acceptable duality gap are bounded by $\mathcal{O}\left(\sqrt{LBG^2K}\right)$, $\mathcal{O}\left(\sqrt{LBG^2K}\right)$, and $\mathcal{O}\left(\sqrt{LBG^2K}\right)$, respectively. In addition, the computational complexity of each iteration for the subproblems is $\mathcal{O}\left(L^3B^3G^4K\right)$, $\mathcal{O}\left(L^3B^3G^4K\right)$, and $\mathcal{O}\left(L^3B^3G^6K\right)$, respectively.

V. NUMERICAL RESULTS

In this section, we present the simulation results to demonstrate the superiority of the proposed NOMA multicast scheme. There is a GEO satellite, $L = 3$ gateways and beam clusters, and each beam cluster has $B = 7$ beams. In terms of the number of users per beam, we simulated a reasonable and sufficient number of users, which is comparable with other multicast multibeam satellite works [5], [18], [33]. As shown in Fig. 3, it is assumed that users are uniformly distributed in the beams and utilize the beam topology considered in [18]. Furthermore, we establish that the maximum power of all antenna feed elements is identical and the maximum power of all gateways is equal to B times of P_{lv}^{Feed} (i.e., $P_{ln}^{\text{Feed}} = P$, $P_l^{\text{GW}} = BP, \forall l \in \mathcal{L}, \forall n \in \mathcal{B}$) [18]. Since noise variance does not affect the algorithm and its results, we suppose that all users have the same noise variance $\sigma_{lbgk}^2 = \kappa T_{\text{R}} B_{\text{ul}}$, where κ, T_{R}, and B_{ul} are the Boltzmann constant, receiver noise temperature, and beam bandwidth, respectively [13], [15]. The system parameters are listed in Table II. The average user throughput per beam R_{avg} used for a multibeam satellite system is given as [14]

$$R_{\text{avg}} = \frac{2B_{\text{ul}}}{(1 + \beta) LB}$$
$$\times \sum_{l=1}^{L}\sum_{b=1}^{B}\sum_{m=1}^{G}\log_2\left(1 + \min_{g\in\mathcal{G}, k\in\mathcal{K}}\text{SINR}_{lbgk}^m(\mathcal{W}, \mathcal{P}_{lb}, \mathcal{A}_{lb})\right),$$
(33)

where β is the roll-off factor.

We compare the proposed multicast scheme with another baseline scheme, the general OMA scheme. The OMA method transmits signals to multiple user groups by dividing the time according to the number of groups within the beam. Similar to the OMA multicast technique, the SCA-SOCP algorithm proposed in [18] is employed in accordance with the situation

TABLE II
SYSTEM PARAMETERS.

PARAMETER	VALUE
Satellite height	35,786 km (GEO orbit)
Link frequency band	f_c = 20 GHz (Ka-band)
Boltzmann constant	$\kappa = 1.3807 \times 10^{-23}$
Noise temperature	T_R = 517 K
User link bandwidth	B_{ul} = 500 MHz
Satellite antenna gain	52 dBi
User antenna gain	41.7 dBi
Beam diameter	250 km
3 dB angle	0.4°
Rain attenuation	$(\mu, \sigma) = (-3.125, 1.591)$
Roll-off factor	$\beta = 0.20$

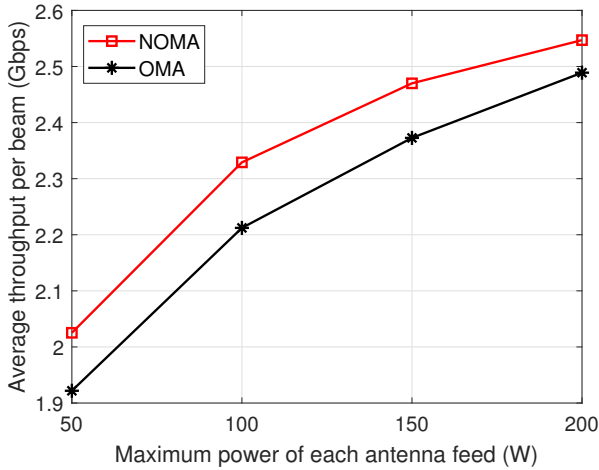Fig. 5. Average user throughput vs. the number of users per group for $G = 2$ and $P = 100$ W [1].

Fig. 4. Average user throughput vs. the maximum power of antenna feed element for $G = 2$ and $K = 2$.

considered in this paper. The SCA-SOCP algorithm results in higher throughput performance than those suggested in [14] and [17].

Fig. 4 plots the average user throughput per beam vs. the maximum power of the antenna feed element. In each beam, there are two groups with two users per group, i.e., $G = 2$ and $K = 2$. For all two techniques, the throughput increases as a higher power is available. The proposed NOMA scheme has higher performance than the OMA scheme. The main reason for this performance difference is the number of users that can be covered simultaneously. NOMA employs SC to cover G user groups in each beam at a time, whereas OMA covers one group at a time in each beam by dividing time with the TDMA technique. That is, during a one-time slot, NOMA covers GK users at the same time, while OMA covers K users. OMA needs G times more timeslots to cover the same number of users as NOMA. For the NOMA scheme, each beamforming vector needs to be optimized for multiple groups of users because one beamforming vector is used for one beam. On the other hand, for OMA scheme, it needs to be optimized for only one group of users. In other words, the beamforming vector is obtained for G times more user channel information than OMA in the case of NOMA. Nevertheless, NOMA still outperforms OMA.
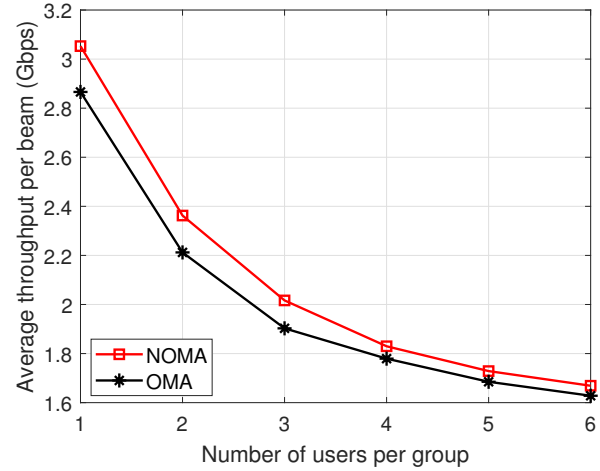
As shown in Fig. 5, we set the number of groups in each beam as $G = 2$ and show the average user throughput per beam for the number of users per group (K) in the beams, where the antenna feed element power is fixed at 100 W. As K increases, R_{avg} decreases in all two techniques because the precoding vector needs to be adjusted for more multigroup multicast channels. The gap between NOMA and OMA schemes also decreases. However, the NOMA scheme still outperforms the OMA scheme. This is because NOMA covers more users simultaneously, even if the difference in the number of users sharing the same precoding vector between NOMA and OMA gradually increases.

Furthermore, the average user throughput per beam with respect to the number of groups per beam is investigated. Each group has two users and the antenna feed element power is set to 100 W. In Fig. 6, as the number of groups increases from 1 to 4, the NOMA scheme can transmit more signals simultaneously, resulting in an increase in the throughput. However, in the case of the OMA scheme, signals can be transmitted to only one group per time slot, so the number of time slots required increases in proportion to the number of groups. Thus, its throughput remains constant regardless of the number of groups.

The accurate channel information may not be obtained due to limited feedback or channel estimation error. It is important to investigate the impact of imperfection in our proposed scheme since the channel imperfection might degrade the performance more severely in our system requiring channel information of more multigroup multicast users. To consider those situations, we examine the performance even in the

---

[1]The number of users per group was set by referring to other papers on multicast multibeam SatCom [14], [17], [18], [43], [44], and other technical reports on DVB-S2X such as [45]. According to [14], if the number of users per beam to be covered at once is too large, lower performance can be obtained than the conventional frequency reuse techniques without precoding. Therefore, if the number of earth stations that can be covered by a beam at once is exceeded, the users are divided so that the appropriate number of users is covered for every time slot. By properly dividing the users, the number of users per frame is not increased too much, and the number of users to be covered at one time per beam can be kept moderate.
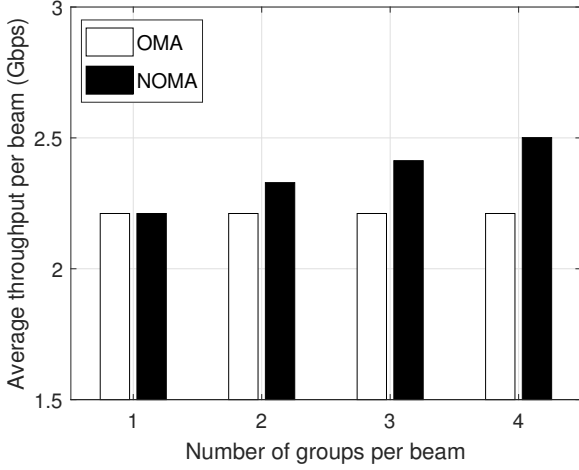
Fig. 6. Average user throughput vs. the number of groups per beam for $K = 2$ and $P = 100$ W.
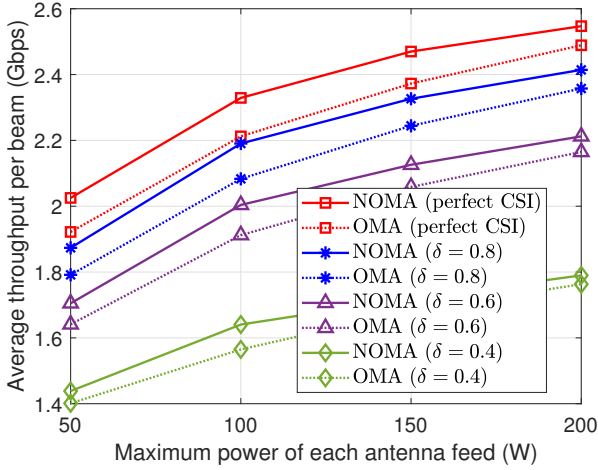

Fig. 8. Average user throughput vs. the maximum power of antenna feed element for different ε.

Fig. 7. Average user throughput vs. the maximum power of antenna feed element for different δ.

Fig. 9. Convergence of Algorithm 1 over a random channel generation for $G = 2$, $K = 2$, and $P = 50$ W.

cases of obtaining channel information of various qualities. We model the imperfect CSI at the gateways as follows, $\mathbf{h} = \hat{\mathbf{h}} + \tilde{\mathbf{h}}$, where $\tilde{\mathbf{h}}$ and $\hat{\mathbf{h}}$ are the channel estimate of the channel vector $\mathbf{h}$ and the channel estimation errors, respectively. The channel estimation error is independent and identically distributed and independent of the channel estimate. It follows a complex Gaussian distribution $\mathcal{CN}\left(0, \sigma_e^2\right)$, where $\sigma_e^2 = (BP)^{-\delta}$ [43], [44]. A larger δ value means more accurate channel information and a value of δ between 0 and 1 is mainly selected. Additionally, we also investigate the effect of imperfect SIC. In order to calculate the residual interference due to the imperfect SIC, the error ratio $\varepsilon > 0$ is introduced and the performance is calculated for several error ratio values $\varepsilon \in [0, 1]$, where $\varepsilon = 0$ denotes perfect SIC [46]. The SINR to which the imperfect SIC is applied can be written as

$$\frac{\left|\mathbf{h}_{l,lbgk}^H \mathbf{w}_{lb}\right|^2 \alpha_{lbm} + a_{lb,mg} M}{\sum_{u \neq m}^G \bar{a}_{lb,um} \left|\mathbf{h}_{l,lbgk}^H \mathbf{w}_{lb}\right|^2 \alpha_{lbu} + I_{lbgk}(\mathcal{W}) + \sigma_{lbgk}^2}, \quad (34)$$
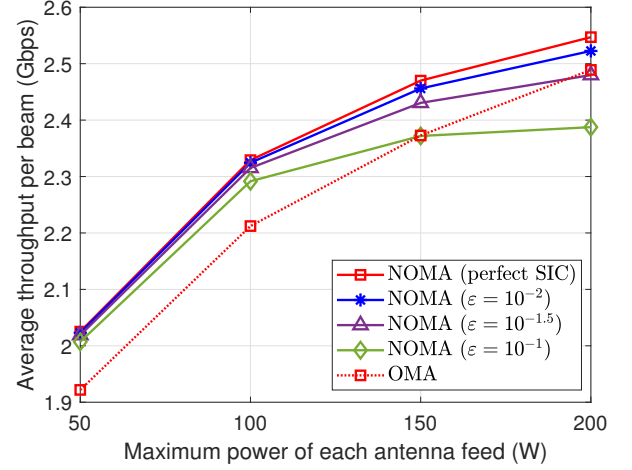
where $\bar{a}_{lb,um} = \varepsilon$ if $a_{lb,um} = 0$, $\bar{a}_{lb,um} = 1$ otherwise. In Figs. 7 and 8, we compare the proposed NOMA scheme with the OMA scheme in the context of considering the imperfect CSI and imperfect SIC. Fig. 7 shows the average user throughput per beam according to the channel quality, where $\delta = [0.4, 0.6, 0.8]$. As the quality of channel information decreases, the average throughput of both schemes decreases, but the proposed NOMA scheme still shows higher performance than the OMA scheme. Fig. 8 plots the effect on the imperfect SIC, where $\varepsilon = \left[10^{-2}, 10^{-1.5}, 10^{-1}\right]$. As the error ratio increases, the performance of the proposed NOMA scheme decreases. In the high-power region, the proposed NOMA scheme shows lower throughput than the OMA scheme. This is because residual interference due to the imperfect SIC also increases as the transmit power increases. However, the proposed method can be said to be robust to SIC capability because the antenna feed element power is mainly considered around 100 W in realistic SatCom systems.

Finally, we examine the convergence of Algorithm 1 over a channel realization where we set $G = 2$, $K = 2$, and

$P = 50$ W. As shown in Fig. 9, we can observe that the objective function value increases and converges after some iterations.

VI. CONCLUSION AND REMARKS

In this work, we investigated the application of NOMA to a multibeam multicast satellite system with multiple gateways. The precoding vector, power allocation factor, and decoding order were optimized to reduce inter-beam interference due to full frequency reuse and improve spectral efficiency. We formulated the problem to maximize the sum rate under power constraints. After the optimization problem was converted into a tractable form, we could find the local optimal solution using the developed algorithm. Simulation results have demonstrated that the proposed NOMA scheme outperforms the OMA method.

When the system of this paper is handled in low earth orbit (LEO) SatCom, the channel model should be considered differently from GEO SatCom. Although LEO SatCom has a shorter propagation delay than GEO SatCom, it experiences a Doppler shift due to a different orbital period than the Earth's rotation period. If the Doppler shift is predicted and compensated effectively as in [47]–[49], the technique presented in this paper can also be applied to LEO SatCom.

REFERENCES

[1] G. Giambene, S. Kota, and P. Pillai, "Satellite-5G integration: A network perspective," *IEEE Netw.*, vol. 32, no. 5, pp. 25–31, Sep. 2018.

[2] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini, "What should 6G be?" *Nat. Electron.*, vol. 3, no. 1, pp. 20–29, Jan. 2020.

[3] E. Yaacoub and M.-S. Alouini, "A key 6G challenge and opportunity—Connecting the base of the pyramid: A survey on rural connectivity," *Proc. IEEE*, vol. 108, no. 4, pp. 533–582, Apr. 2020.

[4] M. Giordani and M. Zorzi, "Non-terrestrial networks in the 6G era: Challenges and opportunities," *IEEE Netw.*, vol. 35, no. 2, pp. 244–251, Mar. 2021.

[5] M. A. Vazquez *et al.*, "Precoding in multibeam satellite communications: Present and future challenges," *IEEE Wireless Commun.*, vol. 23, no. 6, pp. 88–95, Dec. 2016.

[6] V. Joroughi *et al.*, "Deploying joint beam hopping and precoding in multibeam satellite networks with time variant traffic," in *Proc. GlobalSIP*, 2018.

[7] M. G. Kibria, E. Lagunas, N. Maturo, D. Spano, and S. Chatzinotas, "Precoded cluster hopping in multi-beam high throughput satellite systems," in *Proc. IEEE GLOBECOM*, 2019.

[8] A. M. K., "Channel estimation and detection in hybrid satellite–terrestrial communication systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5764–5771, Jul. 2016.

[9] G. Taricco, "Linear precoding methods for multi-beam broadband satellite systems," in *Proc. European Wireless Conf.*, 2014.

[10] M. A. Diaz, N. Courville, C. Mosquera, G. Liva, and G. E. Corazza, "Non-linear interference mitigation for broadband multimedia satellite systems," in *Proc. IWSSC*, 2007.

[11] S. Chatzinotas, G. Zheng, and B. Ottersten, "Energy-efficient MMSE beamforming and power allocation in multibeam satellite systems," in *Proc. ASILOMAR*, 2011.

[12] C. Qi and X. Wang, "Precoding design for energy efficiency of multibeam satellite communications," *IEEE Commun. Lett.*, vol. 22, no. 9, pp. 1826–1829, Sep. 2018.

[13] G. Zheng, S. Chatzinotas, and B. Ottersten, "Generic optimization of linear precoding in multibeam satellite systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2308–2320, Jun. 2012.

[14] 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, Sep. 2015.

[15] V. Joroughi, M. Á. Vázquez, and A. I. Pérez-Neira, "Precoding in multigateway multibeam satellite systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4944–4956, Jul. 2016.

[16] C. Mosquera, R. López-Valcarce, T. Ramírez, and V. Joroughi, "Distributed precoding systems in multi-gateway multibeam satellites: Regularization and coarse beamforming," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6389–6403, Oct. 2018.

[17] D. Christopoulos, H. Pennanen, S. Chatzinotas, and B. Ottersten, "Multicast multigroup precoding for frame-based multi-gateway satellite communications," in *Proc. ASMS/SPSC*, 2016.

[18] J. Wang, L. Zhou, K. Yang, X. Wang, and Y. Liu, "Multicast precoding for multigateway multibeam satellite systems with feeder link interference," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1637–1650, Mar. 2019.

[19] S. M. R. Islam, N. Avazov, O. A. Dobre, and K. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 721–742, 2nd Quart. 2017.

[20] Y. Liu *et al.*, "Nonorthogonal multiple access for 5G and beyond," *Proc. IEEE*, vol. 105, no. 12, pp. 2347–2381, Dec. 2017.

[21] M. Caus, M. A. Vázquez, and A. Pérez-Neira, "NOMA and interference limited satellite scenarios," in *Proc. ACSSC*, 2016.

[22] X. Yan *et al.*, "The application of power-domain non-orthogonal multiple access in satellite communication networks," *IEEE Access*, vol. 7, pp. 63531–63539, May 2019.

[23] X. Yan *et al.*, "Performance analysis of NOMA-based land mobile satellite networks," *IEEE Access*, vol. 6, pp. 31327–31339, Jun. 2018.

[24] S. A. Tegos, P. D. Diamantoulakis, J. Xia, L. Fan, and G. K. Karagiannidis, "Outage performance of uplink NOMA in land mobile satellite communications," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1710–1714, Oct. 2020.

[25] X. Zhu, C. Jiang, L. Kuang, N. Ge, and J. Lu, "Non-orthogonal multiple access based integrated terrestrial-satellite networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2253–2267, Oct. 2017.

[26] Z. Lin, M. Lin, J. Wang, T. de Cola, and J. Wang, "Joint beamforming and power allocation for satellite-terrestrial integrated networks with non-orthogonal multiple access," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 657–670, Jun. 2019.

[27] A. I. Perez-Neira, M. Caus, and M. A. Vazquez, "Non-orthogonal transmission techniques for multibeam satellite systems," *IEEE Commun. Mag.*, vol. 57, no. 12, pp. 58–63, Dec. 2019.

[28] S. M. Ivari *et al.*, "Power allocation and user clustering in multicast NOMA based satellite communication systems," in *Proc. IEEE ICC*, 2020.

[29] S. M. Ivari, *et al.*, "Precoding and scheduling in multibeam multicast NOMA based satellite communication systems," in *Proc. IEEE ICC Workshops*, 2021.

[30] Y. Huang, *et al.*, "Signal processing for MIMO-NOMA: Present and future challenges," *IEEE Wirel. Commun.*, vol. 25, no. 2, pp. 32–38, Apr. 2018.

[31] D. Hu, Q. Zhang, Q. Li, and J. Qin, "Joint position, decoding order, and power allocation optimization in UAV-based NOMA downlink communications," *IEEE Syst. J.*, vol. 14, no. 2, pp. 2949–2960, Jun. 2020.

[32] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *J. Optim. Theory Appl.*, vol. 109, no. 3, pp. 475–494, Jun. 2001.

[33] V. Joroughi, M. Á. Vázquez, and A. I. Pérez-Neira, "Generalized multicast multibeam precoding for satellite communications," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 952–966, Feb. 2017.

[34] G. Zheng, P. Arapoglou, and B. Ottersten, "Physical layer security in multibeam satellite systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 852–863, Feb. 2012.

[35] M. Lin, Z. Lin, W. Zhu, and J. Wang, "Joint beamforming for secure communication in cognitive satellite terrestrial networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 5, pp. 1017–1029, May 2018.

[36] M. F. Hanif, Z. Ding, T. Ratnarajah, and G. K. Karagiannidis, "A minorization-maximization method for optimizing sum rate in the downlink of non-orthogonal multiple access systems," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 76–88, Jan. 2016.

[37] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[38] K. Nguyen, Q. Vu, M. Juntti, and L. Tran, "Distributed solutions for energy efficiency fairness in multicell MISO downlink," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6232–6247, Jul. 2017.

[39] H. Tuy, *Convex Analysis and Global Optimization*. Springer, 2016.

[40] Q. Vu, K. Nguyen, and M. Juntti, "Max-min fairness for multicast multigroup multicell transmission under backhaul constraints," in *Proc. IEEE GLOBECOM*, 2016.

[41] T. Van Do, H. A. L. Thi, and N. T. Nguyen, *Advanced Computational Methods for Knowledge Engineering*. Springer, 2014.

[42] M. S. Lobo, l. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra Appl.*, vol. 284, no. 1-3, pp. 193–228, Nov. 1998.

[43] L. Yin and B. Clerckx, "Rate-splitting multiple access for multigroup multicast and multibeam satellite systems," *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 976–990, Feb. 2021.

[44] Z. W. Si, L. Yin, and B. Clerckx, "Rate-splitting multiple access for multigateway multibeam satellite systems with feeder link interference," *IEEE Trans. Commun.*, vol. 70, no. 3, pp. 2147–2162, Mar. 2022.

[45] *Digital video broadcasting (DVB); Implementation guidelines for the second generation system for broadcasting, interactive services, news gathering and other broadband satellite applications. part 2: S2 extensions (DVB-S2X)*. ETSI TR 102 376-2 V1.1.1, European Broadcasting Union, Nov. 2015.

[46] Y. Fu, Y. Chen, and C. W. Sung, "Distributed power control for the downlink of multi-cell NOMA systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6207–6220, Sep. 2017.

[47] L. You *et al.*, "Massive MIMO transmission for LEO satellite communications," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1851–1865, Aug. 2020.

[48] W. Wang *et al.*, "Near optimal timing and frequency offset estimation for 5G integrated LEO satellite communication system," *IEEE Access*, vol. 7, pp. 113298–113310, Aug. 2019.

[49] S. Amiri and M. Mehdipour, "Accurate doppler frequency shift estimation for any satellite orbit," in *Proc. RAST*, 2007.
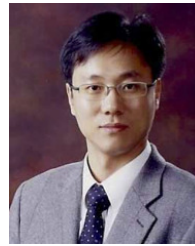
Young-Chai Ko (S'97–M'01-SM'06) received the B.Sc. degree in Electrical and Telecommunication Engineering from the Hanyang University, Seoul, Korea and the M.S.E.E. and Ph.D. degrees in Electrical Engineering from the University of Minnesota, Minneapolis, MN in 1999 and 2001, respectively. He was with Novatel Wireless as a Research Scientist from January 2001 to March 2001. In March 2001, he joined the Texas Instruments, Inc., wireless center, San Diego, CA, as a Senior Engineer. He is now with the school of Electrical Engineering at Korea University as a Professor. His current research interests include the design and evaluations of multi-user cellular system, MODEM architecture, mm-wave and Tera Hz wireless systems.

Dong-Hyoun Na (S'21) received the B.S. and M.S. degrees in Electrical Engineering from Korea University, Seoul, South Korea, in 2017 and 2019, respectively, where he is currently pursuing the Ph.D. degree with the School of Electrical Engineering. He is also pursuing the Ph.D. degree with the Division of Computer, Electrical, Mathematical Science and Engineering (CEMSE), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. His current research interests include performance analysis and optimization of satellite communication systems.

Mohamed-Slim Alouini (S'94-M'98-SM'03-F'09) was born in Tunis, Tunisia. He received the Ph.D. degree in Electrical Engineering from the California Institute of Technology (Caltech), Pasadena, CA, USA, in 1998. He served as a Faculty member at the University of Minnesota, Minneapolis, MN, USA, then in the Texas A & M University at Qatar, Education City, Doha, Qatar before joining King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia as a Professor of Electrical Engineering in 2009. His current research interests include modeling, design, and performance analysis of wireless communication systems.

Ki-Hong Park (S'06–M'11-SM'19) received the B.Sc. degree in Electrical, Electronic, and Radio Engineering from Korea University, Seoul, South Korea, in 2005, and the joint M.S. and Ph.D. degrees from the School of Electrical Engineering, Korea University, in 2011. He joined the King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia, in 2011, as a Post-Doctoral Fellow. Since 2014, he has been a Research Scientist of Electrical and Computer Engineering with the Division of Computer, Electrical, Mathematical Science and Engineering (CEMSE), KAUST. His research interests include communication theory and its application to the design and performance evaluation of wireless communication systems and networks. His research interests include the application to optical wireless communications, non-terrestrial communications, and physical layer secrecy.