Radio Resource Management Design for RSMA: Optimization of Beamforming, User Admission, and Discrete/Continuous Rates With Imperfect SIC

Luis F. Abanto-Leon, Aravindh Krishnamoorthy, *Graduate Student Member, IEEE*, Andres Garcia-Saavedra, Gek Hong Sim, Robert Schober, *Fellow, IEEE*, and Matthias Hollick

***Abstract*—This paper investigates the radio resource management (RRM) design for multiuser rate-splitting multiple access (RSMA), accounting for various characteristics of practical wireless systems, such as the use of discrete rates, the inability to serve all users, and the imperfect successive interference cancellation (SIC). Specifically, failure to consider these characteristics in RRM design may lead to inefficient use of radio resources. Therefore, we formulate the RRM of RSMA as optimization problems to maximize respectively the weighted sum rate (WSR) and weighted energy efficiency (WEE), and jointly optimize the beamforming, user admission, discrete/continuous rates, accounting for imperfect SIC, which result in nonconvex mixed-integer nonlinear programs that are challenging to solve. Despite the difficulty of the optimization problems, we develop algorithms that can find high-quality solutions. We show via simulations that carefully accounting for the aforementioned characteristics, can lead to significant gains. Precisely, by considering that transmission rates are discrete, the transmit power can be utilized more intelligently, allocating just enough power to guarantee a given discrete rate. Additionally, we reveal that user admission plays a crucial role in RSMA, enabling additional gains compared to random admission by facilitating the servicing of selected users with mutually beneficial channel characteristics. Furthermore, provisioning for possibly imperfect SIC makes RSMA more robust and reliable.**

***Index Terms*—Beamforming, user admission, discrete rates, rate splitting, imperfect SIC, spectral efficiency, energy efficiency.**

NOMENCLATURE

BnB	Branch-and-bound
EE	Energy efficiency
IPM	Interior-point method
MCS	Modulation and coding scheme
MINLP	Mixed-integer nonlinear program
MISOCP	Mixed-integer second-order cone program
NOMA	Non-orthogonal multiple access
NOUM	Non-orthogonal unicast and multicast
RRM	Radio resource management
RSMA	Rate-splitting multiple access
SCA	Successive convex approximation
SDMA	Space-division multiple access
SDR	Semidefinite relaxation
SE	Spectral efficiency
SIC	Successive interference cancellation
SINR	Signal-to-interference-plus-noise ratio
SOCP	Second-order cone program
SR	Sum rate
SSR	Sum secrecy rate
WEE	Weighted energy efficiency
WSR	Weighted sum rate

Manuscript received 18 July 2023; revised 6 March 2024; accepted 29 April 2024. Date of publication 2 May 2024; date of current version 5 November 2024. This work was supported in part by the State of Hesse through LOEWE emergenCITY under Grant LOEWE/1/12/519/03/05.001(0016)/72, in part by German Research Foundation through HyRIS under Grant 455077022, in part by B5G-Cell under Grant 210487104, and in part by the European Commission through DAEMON under Grant 101017109. Recommended for acceptance by E. Aryafar. *(Corresponding author: Luis F. Abanto-Leon.)*

Luis F. Abanto-Leon, Gek Hong Sim, and Matthias Hollick are with the Secure Mobile Networking Lab, Technische Universität Darmstadt, 64289 Darmstadt, Germany (e-mail: labanto@seemoo.tu-darmstadt.de).

Aravindh Krishnamoorthy and Robert Schober are with the Friedrich-Alexander-Universität Erlangen-Nürnberg, 91054 Erlangen, Germany.

Andres Garcia-Saavedra is with the NEC Laboratories Europe, 69115 Heidelberg, Germany.

This article has supplementary material provided by the authors available at https://doi.org/10.1109/TMC.2024.3396389.

Digital Object Identifier 10.1109/TMC.2024.3396389

I. INTRODUCTION

RATE-SPLITTING multiple access (RSMA) has emerged as a promising technology capable of outperforming non-orthogonal multiple access (NOMA) and space-division multiple access (SDMA), owing to its superior ability to cope with multiuser interference [1], [2], [3]. RSMA is a power-domain non-orthogonal technology that relies on multi-antenna rate-splitting at the transmitter and successive interference cancellation (SIC) at the user side. Specifically, the transmitter partitions the message for each user into a common and a private portion. Then, it encodes the common portions of all users into a common stream and each of the private portions into an independent stream. Prior to over-the-air transmission, the common and private streams are precoded by the transmitter. Upon reception of the streams, each user employs SIC to decode and remove the common stream, which carries the common portions of other users and its own, before accessing the private streams to decode its private portion. With both the common and private portions available, the user can reassemble its message. By adjusting the partitioning of messages into common and private portions, RSMA flexibly controls the level of interference that each user can cancel, thus bridging smoothly between the two extreme strategies of fully decoding interference (as in NOMA) and fully treating it as noise (as in SDMA), leading to further performance gains [3], [4], [5].

To date, several studies have demonstrated in various use cases [6] the higher capabilities of RSMA compared to SDMA and NOMA, thus positioning RSMA as a formidable multiple access candidate with enormous potential to meet the stringent connectivity requirements of next-generation wireless communications systems [7]. Still, a key element that needs further advances to ensure high RSMA performance is the radio resource management (RRM) design. The RRM for RSMA systems has focused chiefly on the beamforming and power allocation design, which have been investigated for a plethora of use cases and different design goals, such as fairness, sum secrecy rate (SSR), sum rate (SR), weighted SR (WSR), and weighted energy efficiency (WEE)[1] optimization, as summarized in the following.

The authors of [9] studied the beamforming design for WSR maximization in joint radar and communications (JRC), whereas the authors of [10] developed cooperative beamforming strategies to maximize the WSR in visible light communications (VLC). The authors of [11], [12] investigated the beamforming design for WSR maximization in non-orthogonal unicast and multicast (NOUM) systems. In contrast, the security aspect was investigated in [13], where the beamforming and artificial noise were designed to maximize the secrecy rate fairness. Besides, the beamforming design for rate fairness maximization was investigated in [14] in the context of cooperative relaying networks. Driven by the same goal, the joint design of beamforming and the phase shifts of an intelligent reflecting surface (IRS) was researched in [15]. In addition, the beamforming design for WEE maximization was studied for unmanned aerial vehicle networks in [16], for VLC in [17], for semantic communications in [18], and for JRC in [19]. The authors of [20], [21] investigated the beamforming design for maximization of respectively the WSR and WEE, whereas the authors of [22] designed the beamforming for simultaneous WSR and WEE maximization. Power allocation was investigated for SR, SSR, and rate fairness maximization in [23], [24], [25], [26], and [27], respectively. *The body of work on beamforming and power allocation design of RSMA continues to grow and show promising results. However, the literature has so far ignored critical characteristics that are inherent to practical wireless systems, i.e., rate discretization, user admission, and imperfect SIC, which merit investigation.*

Concerning the first characteristic, most of the literature, assumed continuous rates modeled by Shannon's capacity formula, e.g., [1], [2], [3], [6], [9], [10], [13], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27]. This assumption contrasts with the predominant use of discrete rates in practical wireless systems and raises questions as to whether RSMA's gains will hold when discrete rates are accounted for. Transmission rates in wireless systems are determined by the choice of a modulation and coding scheme (MCS), as specified by 3GPP [28], leading to a finite set of discrete rates. Employing Shannon's capacity formula is mathematically more tractable than assuming discrete rates, hence its wide adoption. However, it renders continuous-valued rate upper bounds that are unachievable in practice. A naive solution to this problem is to project the continuous rates onto the discrete rate set, i.e., round them to the closest feasible discrete rate [29]. However, this may lead to performance degradation. Therefore, rate discretization must be properly accounted for in the RRM design to exploit the full potential of RSMA. *To date, the incorporation of discrete rates into the RRM design of RSMA remains to be investigated.* An early study in [30] showed that RSMA outperforms SDMA for discrete rates. The proposed design, however, did not account for predefined MCSs, as the authors assumed continuous rates and tailored the MCSs to achieve a SR close to the ensemble average obtained over multiple channel realizations. A few works investigated the impact of beamforming and discrete rates on the performance of SDMA. For instance, joint beamforming and discrete rate selection design was investigated in [31], [32] for SR maximization and in [33] for WSR maximization. However, these results are not applicable to RSMA, since RSMA is a more general framework that includes SDMA as a special case.

Concerning the second characteristic, access policies in wireless systems typically restrict the number of users served per time slot, e.g., due to the availability of a limited number of radio frequency (RF) chains. This characteristic is especially limiting for SDMA, which requires one RF chain per served user. In contrast, RSMA can support more users than RF chains are available since RSMA can exploit the multicast signal to aggregate information for several users. However, in this case, RSMA may degrade severely as the number of users increases since the multicast signal must be delivered to all users. This peculiarity becomes a limitation in RSMA and raises the need for selective user admission. *The impact of user admission on performance has been studied for SDMA and NOMA, but has yet to be researched for RSMA.* For instance, the joint beamforming and user admission design for SDMA was investigated in [34] to minimize the transmit power, and in [35] to maximize the SR and rate fairness. The authors of [36] developed joint beamforming and user admission strategies for SDMA to maximize the number of users served. With the same goal, the authors of [37] designed the power and subchannel allocation for NOMA. Besides, the joint design of beamforming, user admission, and discrete rate selection for SR maximization of SDMA was investigated in [38]. However, the solutions developed in the preceding studies are not valid for RSMA. Specifically, user admission for RSMA differs significantly from that for SDMA and NOMA, as RSMA delivers information to users via superimposed multicast and unicast precoders. Unicast precoders benefit from users with uncorrelated channels because interference is easier to mitigate. In contrast, the multicast precoder benefits from users with correlated channels as this facilitates transmitting shared information with less power. Therefore, given these conflicting objectives inherent to RSMA, including user admission in the RRM design is essential.

Finally, as the last key characteristic, it is important to account for imperfect SIC. Specifically, the performance of RSMA is highly dependent on the success of SIC. In practice, SIC is generally not perfect, which causes unmanaged self-interference

[1]In our work, EE (WEE) is defined as the ratio of SR (WSR) to total power consumption, with the goal being its maximization. An alternative but less commonly used definition of EE is expressed as the total energy expenditure, with the goal being its minimization [8].

that can compromise performance. *Despite the importance of accounting for imperfect SIC in the RRM design of RSMA, SIC has been assumed to be perfect in most of the RSMA literature, with few exceptions.* For instance, the authors of [39] investigated the beamforming and subcarrier allocation design for SR maximization taking into account imperfect SIC. However, the proposed design assumed continuous rates and did not consider user admission. The impact of imperfect SIC on the SR of RSMA was also investigated in [40], where the authors derived bounds for power allocation but did not take user admission and discrete rates into account. NOMA can also be affected by imperfect SIC and, therefore, a number of works have investigated its impact. In particular, for NOMA, the impact of imperfect SIC and power allocation was studied in [41], [42], [43] for SR maximization, and in [44] for EE maximization.

Motivated by the above discussion, the performance of RSMA systems in real-world deployments is anticipated to improve significantly if the characteristics above are taken into account. Thus, we propose to account for discrete rates, user admission, and imperfect SIC during RRM design. To the best of the authors' knowledge, RRM design for RSMA systems considering these characteristics has not been investigated yet. Due to the widespread adoption of Shannon's capacity formula for RRM design, we also investigate the integration of continuous rates, user admission, and imperfect SIC, which has not been studied before for RSMA systems. We adopt the maximization of the WSR and WEE as design objectives. This paper makes the following contributions:

- We formulate two novel RRM problems to maximize respectively the WSR and WEE of RSMA by jointly optimizing the beamforming, user admission, and private and common discrete rates, while accounting for imperfect SIC. The resulting WSR and WEE problems, denoted by $\mathcal{P}'_{\text{DWSR}}$ and $\mathcal{P}'_{\text{DWEE}}$ in Section III-A, are nonconvex mixed-integer nonlinear programs (MINLPs) and are difficult to solve. In addition, we formulate problems $\mathcal{Q}'_{\text{CWSR}}$ and $\mathcal{Q}'_{\text{CWEE}}$ in Section IV-A, which represent the continuous-rate counterparts of $\mathcal{P}'_{\text{DWSR}}$ and $\mathcal{P}'_{\text{DWEE}}$, and are also nonconvex MINLPs.
- In Section III-B, we propose an optimal mixed-integer second-order cone program (`OPT-MISOCP`) algorithm, which tackles the nonconvexities of $\mathcal{P}'_{\text{DWSR}}$ and $\mathcal{P}'_{\text{DWEE}}$ through a series of convex transformations. Specifically, instead of treating $\mathcal{P}'_{\text{DWSR}}$ and $\mathcal{P}'_{\text{DWEE}}$ as general nonconvex MINLPs, `OPT-MISOCP` approximates $\mathcal{P}'_{\text{DWSR}}$ and $\mathcal{P}'_{\text{DWEE}}$ as convex MINLPs $\mathcal{P}_{\text{DWSR}}$ and $\mathcal{P}_{\text{DWEE}}$, respectively, which can be solved in a globally optimal manner. However, circumventing the nonconvexities of $\mathcal{P}'_{\text{DWSR}}$ and $\mathcal{P}'_{\text{DWEE}}$ poses the risk of shrinking their feasible set, possibly resulting in a loss of optimality. Therefore, we derive an upper bound to evaluate the corresponding loss in performance and show via simulations that the globally optimal solutions for $\mathcal{P}_{\text{DWSR}}$ and $\mathcal{P}_{\text{DWEE}}$ result in near-optimal solutions for $\mathcal{P}'_{\text{DWSR}}$ and $\mathcal{P}'_{\text{DWEE}}$ with negligible degradation. In addition, the proposed `OPT-MISOCP` algorithm features customized cutting planes that reduce the runtime by a factor of $3-20$ without impacting performance.
- In Section IV-B, we solve $\mathcal{Q}'_{\text{CWSR}}$ and $\mathcal{Q}'_{\text{CWEE}}$ based on binary enumeration and convex transformations. In particular, we employ enumeration to list all possible user admission combinations, thus resulting in multiple subproblems. To solve each subproblem, we devise an optimal successive convex approximation with semidefinite relaxation (`OPT-SCA-SDR`) algorithm, which finds a Karush-Kuhn-Tucker (KKT) point. In addition, to comply with the finite set of discrete rates, the continuous rates obtained by `OPT-SCA-SDR` are projected, i.e., rounded to the closest feasible discrete rates.
- Our simulations show that RSMA designed for discrete rates achieves gains of up to 89.7% (WSR) and 21.5% (WEE) compared to projecting continuous rates. Also, user admission proves crucial for RSMA as it yields additional gains of up to 15.3% (WSR) and 11.4% (WEE) compared to random user admission when discrete rates are considered. Furthermore, accounting for imperfect SIC prevents severe performance degradation by mitigating the impact of self-interference. Overall, our simulation results reveal that accounting for characteristics of practical wireless systems in RRM of RSMA leads to improved exploitation of the radio resources, and therefore to higher spectral efficiency (SE) and energy efficiency (EE).

The remainder of this paper is organized as follows. In Section II, we present the system model. In Section III, considering discrete rates, we formulate and solve the RRM as optimization problems for maximization of the WSR and WEE, respectively, while in Section IV, we do the same for continuous rates. Simulation results are presented in Section V. Finally, we summarize our conclusions in Section VII.

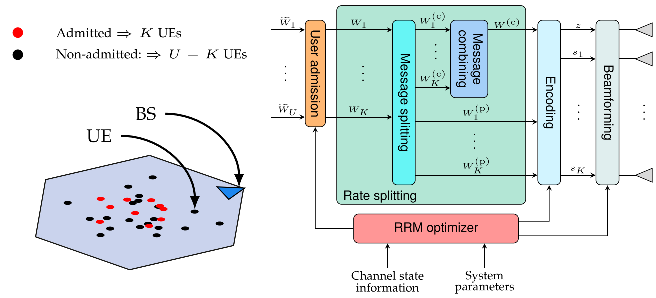
Notation: In this paper, $|a|$ and the $\|\mathbf{a}\|_2$ represent the absolute value of scalar a and the ℓ_2-norm of vector $\mathbf{a}$, respectively. $\mathbf{A}^{\text{T}}$, $\mathbf{A}^{\text{H}}$, $\text{Rank}(\mathbf{A})$, $\text{Tr}(\mathbf{A})$, and $\mathfrak{Re}\{\mathbf{A}\}$ and $\mathfrak{Im}\{\mathbf{A}\}$ denote the transpose, Hermitian transpose, rank, trace, and real and imaginary part of matrix $\mathbf{A}$, respectively. $\mathbf{A} \succcurlyeq \mathbf{0}$ indicates that $\mathbf{A}$ is a positive semidefinite matrix. $\mathbb{C}^{N\times M}$ denotes the space of $N \times M$ complex-valued matrices. $\mathbf{I}$ is the identity matrix, $j \triangleq \sqrt{-1}$ is the imaginary unit, and $\mathbb{E}\{\cdot\}$ denotes statistical expectation.

II. System Model

In this section, we present the system model for the considered RRM optimization problems.

A. System Architecture

We consider the downlink cellular system shown in Fig. 1(a), where a base station (BS) is equipped with an antenna array with N_{tx} elements, which can consume a maximum transmit power of $P_{\text{tx}}^{\max}$. There are U single-antenna user equipments (UEs), and the BS admits only K UEs, where $K \leq U$. We index the UEs with the elements of set $\mathcal{U}$, such that $|\mathcal{U}| = U$. The UEs are distributed within a 120° sector and are located at a maximum distance D_{BS} from the BS. The BS estimates the channel state information (CSI) using uplink pilots exploiting channel reciprocity. The RRM optimizer at the BS uses the CSI



(a) System model (b) RSMA with integrated RRM optimizer

Fig. 1. System model and RSMA with integrated RRM optimizer. In the system, K out of U UEs are admitted for downlink transmission. The messages for the admitted UEs are precoded via rate splitting and transmitted over the air.

and other system parameters, such as the maximum transmit power, to maximize the WSR or WEE.

B. RSMA Principle

RSMA allows for partial interference decoding, which facilitates the selective cancellation of interference components. Precisely, UEs focus on decoding their desired signals while coping only with specific interference components from other UEs. This strategy allows RSMA to achieve additional gains over NOMA, which attempts to decode interference entirely, and SDMA, which treats interference as noise and hence does attempt to decode it at all. Partial interference decoding in RSMA is accomplished by splitting the UEs' messages at the BS into common and private portions and sending them via multicast and unicast signals, respectively. As all UEs receive the multicast signal containing the common portions, each UE decodes information intended for other UEs. Precisely, by optimizing the partitioning of messages into common and private portions, RSMA dynamically adjusts the level of interference for each UE. This adaptive capability allows RSMA to seamlessly transition between fully treating interference as noise, as in SDMA, and fully decoding it, as in NOMA [3], [4], [5]. Although NOMA and RSMA both employ SIC, the former requires UEs to employ multiple SIC stages to sequentially decode interference from other UEs and thus recover their desired signals. Particularly, the number of SIC stages required increases with the number of UEs served. Furthermore, the order in which signals are decoded and removed can affect the performance and complexity of NOMA receivers. In contrast, regardless of the number of UEs, RSMA requires only one SIC stage as we consider single-layer RSMA [1].

In the following, we explain the technical aspects of RSMA but exclude UE admission for ease of presentation. In Fig. 1(b), every UE has a corresponding message denoted by $\widetilde{W}_u, u \in \mathcal{U}$, but only K UEs out of U are served. Thus, we assume that K UEs are pre-selected for downlink transmission, and denote this set of UEs by $\mathcal{U}'$, such that $|\mathcal{U}'| = K$, and by UE_u the u-th UE in $\mathcal{U}'$. Now, every UE in $\mathcal{U}'$ is served with a message denoted by $W_u, u \in \mathcal{U}'$, which is decomposed into two parts as $W_u \triangleq (W_u^{(\mathrm{p})}, W_u^{(\mathrm{c})})$, where $W_u^{(\mathrm{p})}$ and $W_u^{(\mathrm{c})}$ are respectively referred to as the private and common portions of W_u. The private portion of UE_u is encoded into a symbol $s_u \in \mathbb{C}$ that is transmitted at a rate $R_u^{(\mathrm{p})}$ in an unicast manner. On the other hand, the common portions $W_u^{(\mathrm{c})}$ of all UEs are combined and encoded into a symbol $s_0 \in \mathbb{C}$, which is transmitted at a rate $R^{(\mathrm{c})}$ in a multicast manner to all UEs. The symbols are assumed to be statistically independent, such that $\mathbb{E}\{\mathbf{s}^{\mathrm{H}}\mathbf{s}\} = \mathbf{I}$ and $\mathbf{s} = [s_0, s_1, \ldots, s_K]^{\mathrm{T}} \in \mathbb{C}^{(K+1)\times 1}$. The rate portion of $R^{(\mathrm{c})}$ corresponding to UE_u is denoted by C_u, such that $R^{(\mathrm{c})} = \sum_u C_u$. As a result, UE_u is served with an overall rate of $R_u^{(\mathrm{p})} + C_u$. Each UE_u recovers $W_u^{(\mathrm{c})}$ first by decoding s_0 and then recovers $W_u^{(\mathrm{p})}$ by decoding s_u with the assistance of SIC. With both portions $W_u^{(\mathrm{c})}$ and $W_u^{(\mathrm{p})}$ available at UE_u, the original message W_u can be reassembled. In addition, each UE_u acquires the common portions $W_{i\neq u}^{(\mathrm{c})}$, corresponding to the other UEs, which are used for interference decoding and cancellation. In particular, the size of the common portions $W_u^{(\mathrm{c})}$ are adjusted according to the level of interference that can be canceled by the UEs [2] and represents the amount of decodable interference, which is removed using SIC.

C. Discrete and Continuous Rates

Practical wireless systems, as defined, e.g., in cellular standards, support only a finite set of data rates [28, p. 64]. These predefined rates are identified by their channel quality indicator (CQI) and correspond to specific MCSs. For each rate, a target SINR is required to ensure a given block error rate (BLER) [45]. The rates and MCSs are typically standardized, e.g., by 3GPP. However, the target SINRs are specific to the equipment in use. We denote with J the total number of available MCSs supported by the system and with $\mathcal{J} = \{1, \ldots, J\}$ the set that indexes them. Hence, for a given discrete rate R_j, $j \in \mathcal{J}$, there is a corresponding target SINR Γ_j that must be met to guarantee that rate. In addition, we assume that $\mathcal{J}$ is an ordered set, such that $R_{j+1} > R_j$ and $\Gamma_{j+1} > \Gamma_j$. Thus, if an UE achieves an SINR of $\bar{\Gamma}$, the BS can allocate any discrete rate $\bar{R}_{\text{dis}} \triangleq \{R_j \mid \Gamma_j \leq \bar{\Gamma}, j \in \mathcal{J}\}$ to the UE. On the contrary, when Shannon's capacity formula is used for rate allocation, the BS assigns continuous rate $\bar{R}_{\text{con}} \triangleq \log_2(1 + \bar{\Gamma})$.

III. Problem Formulation and Proposed Algorithm for Discrete-Rate RSMA

In this section, we formulate and solve the WSR and WEE maximization problems to optimize the beamforming, user admission, and discrete rates for imperfect SIC. For ease of presentation, we summarize the most important parameters and decision variables in Table I.

A. Problem Formulation

We consider two objectives, namely, WSR and WEE maximization, and define the corresponding optimization problems $\mathcal{P}'_{\text{DWSR}}$ and $\mathcal{P}'_{\text{DWEE}}$, shown at the bottom of next page. Specifically, ω_u is the weight associated with UE_u, which can be

TABLE I
PARAMETERS AND DECISION VARIABLES

Parameters and Decision Variables	Notation
Number of antennas at the BS	N_{tx}
Number of UEs	U
Number of admitted UEs	K
Number of discrete rates	J
Channel between the BS and UE_u	$\mathbf{h}_u$
Noise power	σ^2
Weight of UE_u	ω_u
Dynamic power consumption of the circuitry	P_{dyn}
Static power consumption of the circuitry	P_{sta}
Conversion efficiency of the power amplifier	η_{eff}
Common rate of UE_u	C_u
Private precoder for UE_u	$\mathbf{w}_u$
Common precoder for all admitted UEs	$\mathbf{m}$
Binary variable for private rate selection	$\alpha_{u,j}$
Binary variable for common rate selection	κ_j
Binary variable for UE admission	χ_u
Binary variable for private precoder design	μ_u
Binary variable for common precoder design	ψ

set by the network operator, e.g., to improve rate fairness. Besides, we define $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_U]$, $\mathbf{c} = [C_1, \ldots, C_U]$, $\boldsymbol{\chi} = [\chi_1, \ldots, \chi_U]$, $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_U]$, $\boldsymbol{\alpha} = [\alpha_{1,1}, \ldots, \alpha_{U,J}]$, and $\boldsymbol{\kappa} = [\kappa_1, \ldots, \kappa_J]$. In addition, η_{eff} represents the amplifier efficiency and $P_{\text{cir}} = N_{\text{tx}} P_{\text{dyn}} + P_{\text{sta}}$ is the power consumed by the circuitry at the BS, where P_{dyn} and P_{sta} denote the dynamic and static parts, respectively [22]. Next, we discuss the constraints of the optimization problems.

1) User Admission: To indicate whether a given UE_u is admitted, we introduce constraint $\bar{\mathrm{C}}_1 : \chi_u \in \{0, 1\}, \forall u \in \mathcal{U}$, i.e., $\chi_u = 1$ indicates that the BS serves UE_u, and $\chi_u = 0$ otherwise. Further, we have $\bar{\mathrm{C}}_2 : \sum_{u\in\mathcal{U}} \chi_u = K$ as the number of admitted UEs is K. An admitted UE can receive its message via the common signal only, the private signal only, or both. To indicate whether an admitted UE_u is served via the private signal, we introduce $\bar{\mathrm{C}}_3 : \mu_u \in \{0, 1\}, \forall u \in \mathcal{U}$, i.e., $\mu_u = 1$ indicates that UE_u is served via the private signal, and $\mu_u = 0$ otherwise. We also include $\bar{\mathrm{C}}_4 : \mu_u \le \chi_u, \forall u \in \mathcal{U}$, to ensure that non-admitted UEs are not served by a private signal. Naturally, non-admitted UEs are also not served by the common signal but this is handled by constraint $\bar{\mathrm{C}}_{12}$, discussed in Section III-A5. Lastly, we introduce $\bar{\mathrm{C}}_5 : \psi \in \{0, 1\}$ to denote whether the common signal is used.

2) Beamforming: The BS employs a private precoder $\mathbf{w}_u\mu_u \in \mathbb{C}^{N_{\text{tx}}\times 1}$ to precode symbol s_u, and a common precoder $\mathbf{m}\psi \in \mathbb{C}^{N_{\text{tx}}\times 1}$ to precode symbol s_0. The private precoder is $\mathbf{0}$ when UE_u is not admitted. Thus, the downlink signal of the BS is given by $\mathbf{x} = \sum_{u\in\mathcal{U}} \mathbf{w}_u\mu_u s_u + \mathbf{m}\psi s_0$. To account for the transmit power limitation of the BS, the precoders must satisfy $\bar{\mathrm{C}}_6 : \sum_{u\in\mathcal{U}} \|\mathbf{w}_u\mu_u\|_2^2 + \|\mathbf{m}\psi\|_2^2 \le P_{\text{tx}}^{\max}$.

3) Imperfect SIC: The signal received by UE_u is expressed as $y_u = \mathbf{h}_u^{\mathrm{H}}\mathbf{x} + \eta_u$, which is equivalent to

$$y_u = \underbrace{\mathbf{h}_u^{\mathrm{H}}\mathbf{m}\psi s_0}_{\substack{\text{common signal}\\ y_u^{(\mathrm{c})}}} + \underbrace{\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_u\mu_u s_u}_{\substack{\text{private signal}\\ y_u^{(\mathrm{p})}}} + \underbrace{\sum\nolimits_{i\ne u}\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_i\mu_i s_i}_{\substack{\text{interference}\\ y_u^{(\mathrm{int})}}} + \underbrace{\eta_u}_{\text{noise}},$$

$$\begin{aligned}
&\begin{matrix}\mathcal{P}'_{\text{DWSR}}\\ \mathcal{P}'_{\text{DWEE}}\end{matrix} : \max_{\mathbf{W},\mathbf{m},\mathbf{c},\boldsymbol{\chi},\boldsymbol{\mu},\boldsymbol{\alpha},\boldsymbol{\kappa},\psi} \quad \begin{cases} f_{\text{DWSR}}(\mathbf{c},\boldsymbol{\alpha}) \triangleq \sum_{u\in\mathcal{U}}\omega_u\left(\sum_{j\in\mathcal{J}}\alpha_{u,j}R_j + C_u\right) & \text{(linear)}\\ f_{\text{DWEE}}(\mathbf{W},\mathbf{m},\mathbf{c},\boldsymbol{\mu},\boldsymbol{\alpha},\psi) \triangleq \frac{\sum_{u\in\mathcal{U}}\omega_u\left(\sum_{j\in\mathcal{J}}\alpha_{u,j}R_j + C_u\right)}{\frac{1}{\eta_{\text{eff}}}\left(\sum_{u\in\mathcal{U}}\|\mathbf{w}_u\mu_u\|_2^2 + \|\mathbf{m}\psi\|_2^2\right) + P_{\text{cir}}} & \text{(nonconvex)}\end{cases}\\
&\text{s.t.} \quad \bar{\mathrm{C}}_1 : \chi_u \in \{0,1\}, \forall u \in \mathcal{U}, && \text{(binary)}\\
&\quad \bar{\mathrm{C}}_2 : \textstyle\sum_{u\in\mathcal{U}}\chi_u = K, && \text{(linear)}\\
&\quad \bar{\mathrm{C}}_3 : \mu_u \in \{0,1\}, \forall u \in \mathcal{U}, && \text{(binary)}\\
&\quad \bar{\mathrm{C}}_4 : \mu_u \le \chi_u, \forall u \in \mathcal{U}, && \text{(linear)}\\
&\quad \bar{\mathrm{C}}_5 : \psi \in \{0,1\}, && \text{(binary)}\\
&\quad \bar{\mathrm{C}}_6 : \textstyle\sum_{u\in\mathcal{U}}\|\mathbf{w}_u\mu_u\|_2^2 + \|\mathbf{m}\psi\|_2^2 \le P_{\text{tx}}^{\max}, && \text{(nonconvex)}\\
&\quad \bar{\mathrm{C}}_7 : \alpha_{u,j} \in \{0,1\}, \forall u \in \mathcal{U}, j \in \mathcal{J}, && \text{(binary)}\\
&\quad \bar{\mathrm{C}}_8 : \textstyle\sum_{j\in\mathcal{J}}\alpha_{u,j} = \mu_u, \forall u \in \mathcal{U}, && \text{(linear)}\\
&\quad \bar{\mathrm{C}}_9 : \mathsf{SINR}_u^{(\mathrm{p})} \ge \textstyle\sum_{j\in\mathcal{J}}\alpha_{u,j}\Gamma_j, \forall u \in \mathcal{U}, && \text{(nonconvex)}\\
&\quad \bar{\mathrm{C}}_{10} : \kappa_j \in \{0,1\}, \forall j \in \mathcal{J}, && \text{(binary)}\\
&\quad \bar{\mathrm{C}}_{11} : \textstyle\sum_{j\in\mathcal{J}}\kappa_j = \psi, && \text{(linear)}\\
&\quad \bar{\mathrm{C}}_{12} : \mathsf{SINR}_u^{(\mathrm{c})} \ge \chi_u\textstyle\sum_{j\in\mathcal{J}}\kappa_j\Gamma_j, \forall u \in \mathcal{U}, && \text{(nonconvex)}\\
&\quad \bar{\mathrm{C}}_{13} : C_u \ge 0, \forall u \in \mathcal{U}, && \text{(linear)}\\
&\quad \bar{\mathrm{C}}_{14} : C_u \le \chi_u\textstyle\sum_{j\in\mathcal{J}}\kappa_j R_j, \forall u \in \mathcal{U}, && \text{(nonconvex)}\\
&\quad \bar{\mathrm{C}}_{15} : \textstyle\sum_{u\in\mathcal{U}}C_u = \sum_{j\in\mathcal{J}}\kappa_j R_j, && \text{(linear)}\\
&\quad \bar{\mathrm{C}}_{16} : \textstyle\sum_{j\in\mathcal{J}}\alpha_{u,j}R_j + C_u \ge R_{\min}\chi_u, \forall u \in \mathcal{U}, && \text{(linear)}
\end{aligned}$$

TABLE II
RATES AND TARGET SINRS FOR VARIOUS CQIS

CQI (j)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Modulation	QPSK						16QAM			64QAM					
Coding rate	0.0762	0.1172	0.1885	0.3008	0.4385	0.5879	0.3691	0.4785	0.6016	0.4551	0.5537	0.6504	0.7539	0.8525	0.9258
Rate (R_j) [bps/Hz]	0.1523	0.2344	0.3770	0.6016	0.8770	1.1758	1.4766	1.9141	2.4063	2.7305	3.3223	3.9023	4.5234	5.1152	5.5547
Target SINR (Γ_j)	0.1128	0.2159	0.3892	0.6610	1.0962	1.7474	2.8113	4.3321	7.0081	10.6316	16.6648	25.8345	38.4503	60.0620	95.6974

where $y_u^{(\mathrm{c})}$ is the received common signal, $y_u^{(\mathrm{p})}$ is the received private signal, and $y_u^{(\mathrm{int})}$ is the interference at UE_u. Further, $\eta_u \sim \mathcal{CN}(0, \sigma^2)$ denotes additive white Gaussian noise, and $\mathbf{h}_u \in \mathbb{C}^{N_{\mathrm{tx}} \times 1}$ represents the channel between the BS and UE_u. An admitted UE_u utilizes SIC in order to recover its message from y_u. Specifically, UE_u decodes first the common symbol s_0 by treating signals $y_u^{(\mathrm{p})}$ and $y_u^{(\mathrm{int})}$ as noise. Next, UE_u reconstructs the received common signal $y_u^{(\mathrm{c})}$ and subtracts it from y_u, yielding $y_u^{\mathrm{SIC}} = y_u^{(\mathrm{p})} + y_u^{(\mathrm{int})} + \eta_u$, based on which it decodes its private symbol s_u. However, in practice, the removal of $y_u^{(\mathrm{c})}$ is not perfect, which can be caused by, e.g., hardware impairments [40], [41]. Therefore, the signal after imperfect SIC can be expressed as $y_u^{\mathrm{iSIC}} = \Delta_{\mathrm{SIC}} y_u^{(\mathrm{c})} + y_u^{(\mathrm{p})} + y_u^{(\mathrm{int})} + \eta_u$, where $0 \leq \Delta_{\mathrm{SIC}} \leq 1$ is the percentage of the common signal that is not canceled, i.e., $\Delta_{\mathrm{SIC}} = 0$ implies perfect SIC. As a result, the SINRs of the common and private signals at UE_u are $\mathsf{SINR}_u^{(\mathrm{c})} = \frac{|\mathbf{h}_u^{\mathrm{H}} \mathbf{m} \psi|^2}{\sum_{i \in \mathcal{U}} |\mathbf{h}_u^{\mathrm{H}} \mathbf{w}_i \mu_i|^2 + \sigma^2}$ and $\mathsf{SINR}_u^{(\mathrm{p})} = \frac{|\mathbf{h}_u^{\mathrm{H}} \mathbf{w}_u \mu_u|^2}{|\Delta_{\mathrm{SIC}} \mathbf{h}_u^{\mathrm{H}} \mathbf{m} \psi|^2 + \sum_{i \neq u} |\mathbf{h}_u^{\mathrm{H}} \mathbf{w}_i \mu_i|^2 + \sigma^2}$, respectively. The exact value of Δ_{SIC} is usually not known by the BS. Therefore, it must be set properly to avoid performance degradation, and thus guarantee the target SINRs that enable the allocated rates. Typical values for Δ_{SIC} are in the range of 4% and 10% [41].

4) Rate Selection for the Private Signals: An UE_u receiving a private signal at a rate R_j, can only decode the message if $\mathsf{SINR}_u^{(\mathrm{p})} \geq \Gamma_j$, where Γ_j is the target SINR that guarantees R_j (for numerical values, see Table II in Section V). To depict the assignment of private rates, we introduce constraint $\bar{\mathrm{C}}_7 : \alpha_{u,j} \in \{0, 1\}, \forall u \in \mathcal{U}, j \in \mathcal{J}$, where $\alpha_{u,j} = 1$ indicates that UE_u is served by a private signal transmitted at rate R_j. In addition, we include $\bar{\mathrm{C}}_8 : \sum_{j \in \mathcal{J}} \alpha_{u,j} = \mu_u, \forall u \in \mathcal{U}$, to ensure that a rate is allocated to UE_u, if it is served by the private signal. Further, to associate the discrete rates and their corresponding target SINRs, we include $\bar{\mathrm{C}}_9 : \mathsf{SINR}_u^{(\mathrm{p})} \geq \sum_{j \in \mathcal{J}} \alpha_{u,j} \Gamma_j, \forall u \in \mathcal{U}$, which ensures for UE_u a private rate of $\sum_{j \in \mathcal{J}} \alpha_{u,j} R_j$ if $\mu_u = 1$. Note that $\mu_u = 0$ does not indicate that UE_u is not admitted since UE_u can also be served by the common signal if $C_u > 0$.

5) Rate Selection for the Common Signal: An admitted UE_u can only decode the common message transmitted at rate R_j, if $\mathsf{SINR}_u^{(\mathrm{c})} \geq \Gamma_j$. To this end, we introduce constraint $\bar{\mathrm{C}}_{10} : \kappa_j \in \{0, 1\}, j \in \mathcal{J}$, where $\kappa_j = 1$ indicates that rate R_j is selected. We include $\bar{\mathrm{C}}_{11} : \sum_{j \in \mathcal{J}} \kappa_j = \psi$ to allow for the possibility that the common rate is zero (see constraint $\bar{\mathrm{C}}_5$). To unify user admission and the allocation of the common rate, we add $\bar{\mathrm{C}}_{12} : \mathsf{SINR}_u^{(\mathrm{c})} \geq \chi_u \sum_{j \in \mathcal{J}} \kappa_j \Gamma_j, \forall u \in \mathcal{U}$, which results in the common rate $\sum_{j \in \mathcal{J}} \kappa_j R_j$ for all admitted UEs. Although the rate portions C_u are not continuous, they have very fine granularity because rate splitting is capable of dividing the messages W_u into portions of any size. Thus, we treat C_u as continuous-valued by adding $\bar{\mathrm{C}}_{13} : C_u \geq 0, \forall u \in \mathcal{U}$. To keep consistency with user admission, we include $\bar{\mathrm{C}}_{14} : C_u \leq \chi_u \sum_{j \in \mathcal{J}} \kappa_j R_j, \forall u \in \mathcal{U}$, to enforce $C_u = 0$ for non-admitted UEs or when the common rate is zero (see constraints $\bar{\mathrm{C}}_5, \bar{\mathrm{C}}_{11}$). Moreover, we add $\bar{\mathrm{C}}_{15} : \sum_u C_u = \sum_{j \in \mathcal{J}} \kappa_j R_j$ to guarantee that the sum of all common portions C_u is equal to the overall common rate. Finally, we enforce a minimum rate $R_{\min}$ per admitted UE by including constraint $\bar{\mathrm{C}}_{16} : \sum_{j \in \mathcal{J}} \alpha_{u,j} R_j + C_u \geq R_{\min} \chi_u$.

Remark 1: Problems $\mathcal{P}'_{\mathrm{DWSR}}$ and $\mathcal{P}'_{\mathrm{DWEE}}$ are nonconvex MINLPs, which are challenging to solve. Specifically, the nonconvexity is due to constraints $\bar{\mathrm{C}}_6, \bar{\mathrm{C}}_9, \bar{\mathrm{C}}_{12}, \bar{\mathrm{C}}_{14}$ and the objective function $f_{\mathrm{DWEE}}(\mathbf{W}, \mathbf{m}, \mathbf{c}, \boldsymbol{\mu}, \boldsymbol{\alpha}, \psi)$, which contain quotients of quadratic functions and multiplicative couplings. Further, a simple strategy to obtain the SDMA versions of $\mathcal{P}'_{\mathrm{DWSR}}$ and $\mathcal{P}'_{\mathrm{DWEE}}$ is to set $\psi = 0$ because RSMA includes SDMA as a special case.

B. Proposed Algorithm

We propose the `OPT-MISOCP` algorithm to solve the nonconvex MINLPs $\mathcal{P}'_{\mathrm{DWSR}}$ and $\mathcal{P}'_{\mathrm{DWEE}}$. Instead of treating these problems as general nonconvex MINLPs, we propose a sequence of transformations to overcome the nonconvexities, thereby transforming them into the convex MISOCPs $\mathcal{P}_{\mathrm{DWSR}}$ and $\mathcal{P}_{\mathrm{DWEE}}$, respectively, whose global optima can be found using branch-and-bound (BnB) and interior-point methods (IPMs). Specifically, BnB is used for decomposing the binary variables of the MISOCP, whereas IPMs are used for solving the underlying SOCPs. In the following, we describe the proposed algorithm for $\mathcal{P}'_{\mathrm{DWSR}}$ and then we extend it to $\mathcal{P}'_{\mathrm{DWEE}}$.

1) Circumventing Integer Multiplicative Couplings: To cope with the multiplicative coupling between binary variables in constraints $\bar{\mathrm{C}}_{12}$, $\bar{\mathrm{C}}_{14}$, appearing in the form of $\chi_u \kappa_j$, we transform such products into the intersection of linear combinations. We introduce new variables $\pi_{u,j} = \chi_u \kappa_j$, and equivalently rewrite constraints $\bar{\mathrm{C}}_{12}$, $\bar{\mathrm{C}}_{14}$ as constraints $\bar{\mathrm{D}}_1 - \bar{\mathrm{D}}_6$, shown at the bottom of the next page (cf. Appendix A, available online).

2) Circumventing Mixed-Integer Multiplicative Couplings: To deal with the mixed-integer multiplicative couplings in constraints $\bar{\mathrm{C}}_6$, $\bar{\mathrm{C}}_9$, $\bar{\mathrm{D}}_5$, appearing in the form of $\mathbf{w}_u \mu_u$ and $\mathbf{m}\psi$, we reformulate such products as linear relations without altering the nature of the problem. Thus, constraints $\bar{\mathrm{C}}_6, \bar{\mathrm{C}}_9, \bar{\mathrm{D}}_5$ are equivalently rewritten as constraints $\bar{\mathrm{E}}_1 - \bar{\mathrm{E}}_5$, shown at the bottom of the next page (cf. Appendix B, available online).

3) Circumventing Integer Additive Couplings: The additive couplings of binary variables in $\bar{\mathrm{E}}_4$, $\bar{\mathrm{E}}_5$, appearing in the form $\sum_{j \in \mathcal{J}} \alpha_{u,j} \Gamma_j$ and $\sum_{j \in \mathcal{J}} \pi_{u,j} \Gamma_j$, pose a difficulty for convexification since multiple binary variables and their corresponding target SINRs are combined. However, since the couplings are linear and sum to at most one, we can handle them by expanding them into several constraints (i.e., as multiple choice constraints), such that each of the resulting constraints depends on one binary variable only. Thus, constraints $\bar{\mathrm{E}}_4$, $\bar{\mathrm{E}}_5$ are equivalently recast as $\bar{\mathrm{F}}_1$, $\bar{\mathrm{F}}_2$, shown at the bottom of this page (cf. Appendix C, available online).

4) Reformulating the SINR Constraints Via the Big-M Method: To deal with the disjunctiveness caused by the binary variables, which lead to different SINR requirements for the admitted and non-admitted UEs, in $\bar{\mathrm{F}}_1, \bar{\mathrm{F}}_2$, we merge these two cases into a single one via the Big-M method. Thus, by defining $\bar{\mathbf{W}}_u = [\Delta_{\text{SIC}}\mathbf{m}, \mathbf{w}_1, \ldots, \mathbf{w}_{u-1}, \mathbf{w}_{u+1}, \ldots, \mathbf{w}_U]$ and $L^2_{\text{max},u} = \|\mathbf{h}_u\|_2^2 P_{\text{tx}}^{\text{max}} + \sigma^2$, we recast constraints $\bar{\mathrm{F}}_1, \bar{\mathrm{F}}_2$ as $\bar{\mathrm{G}}_1, \bar{\mathrm{G}}_2$, shown at the bottom of this page (cf. Appendix D, available online).

5) Convexifying the Private SINR Constraints: Although constraint $\bar{\mathrm{G}}_1$ is nonconvex, as it contains a difference of convex functions, it can be convexified without changing its feasible set. Thus, constraint $\bar{\mathrm{G}}_1$ can be expressed as $\bar{\mathrm{H}}_1$, $\bar{\mathrm{H}}_2$, $\bar{\mathrm{H}}_3$, shown at the bottom of this page (cf. Appendix E, available online).

6) Convexifying the Common SINR Constraints: The nonconvex constraint $\bar{\mathrm{G}}_2$ can be replaced by the inner convex approximations $\bar{\mathrm{I}}_1$, $\bar{\mathrm{I}}_2$, shown at the bottom of this page, which may shrink the original feasible set (cf. Appendix F, available online).

7) Adding Cutting Planes to Tighten the Feasible Domain: To reduce the search complexity caused by BnB branching for the binary variables, we add problem-specific cutting planes, which do not impact optimality (cf. Appendix G, available online). We add cuts $\bar{\mathrm{J}}_1$ to tighten the feasible set, which can help accelerating the optimization. In addition, we include $\bar{\mathrm{J}}_2$ as an upper bound of the sum-rate, which facilitates early stopping:

$$\bar{\mathrm{J}}_1 : \mathfrak{Re}\left\{\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_u\right\} \geq \sigma \textstyle\sum_{j\in\mathcal{J}} \alpha_{u,j}\sqrt{\Gamma_j}, \forall u \in \mathcal{U}, \qquad \text{(linear)}$$

$$\bar{\mathrm{J}}_2 : \textstyle\sum_{u\in\mathcal{U}} \left(\sum_{j\in\mathcal{J}} \alpha_{u,j}R_j + C_u\right) \leq (K+1)R_J. \qquad \text{(linear)}$$

Remark 2: In our simulations, a remarkable improvement in runtime was observed with the addition of $\bar{\mathrm{J}}_1$ and $\bar{\mathrm{J}}_2$, which accelerated the optimization $3-20$ times compared to the case without them.

8) Outlining the Algorithm and Its Extension to Solve $\mathcal{P}'_{DWEE}$: Recapitulating the results above, problem $\mathcal{P}'_{\text{DWSR}}$ is recast as

$$\mathcal{P}_{\text{DWSR}} : \max_{\boldsymbol{\Theta}} f_{\text{DWSR}}(\mathbf{c}, \boldsymbol{\alpha}) \ \text{ s.t. } \boldsymbol{\Theta} \in \mathcal{C},$$

where $\boldsymbol{\Theta} = (\mathbf{W}, \mathbf{m}, \mathbf{c}, \boldsymbol{\chi}, \boldsymbol{\mu}, \boldsymbol{\alpha}, \boldsymbol{\kappa}, \boldsymbol{\pi}, \psi)$ and $\mathcal{C}$ is the feasible set of $\boldsymbol{\Theta}$ defined by $\bar{\mathrm{C}}_1 - \bar{\mathrm{C}}_5, \bar{\mathrm{C}}_7, \bar{\mathrm{C}}_8, \bar{\mathrm{C}}_{10}, \bar{\mathrm{C}}_{11}, \bar{\mathrm{C}}_{13}, \bar{\mathrm{C}}_{15}, \bar{\mathrm{C}}_{16}, \bar{\mathrm{D}}_1 - \bar{\mathrm{D}}_4, \bar{\mathrm{D}}_6, \bar{\mathrm{E}}_1 - \bar{\mathrm{E}}_3, \bar{\mathrm{H}}_1 - \bar{\mathrm{H}}_3, \bar{\mathrm{I}}_1, \bar{\mathrm{I}}_2, \bar{\mathrm{J}}_1, \bar{\mathrm{J}}_2$.

$$\bar{\mathrm{C}}_{12}, \bar{\mathrm{C}}_{14} \Leftrightarrow \begin{cases} \bar{\mathrm{D}}_1 : \pi_{u,j} \in \{0,1\}, \forall u \in \mathcal{U}, j \in \mathcal{J}, & \text{(binary)} \\ \bar{\mathrm{D}}_2 : \pi_{u,j} \leq \chi_u, \forall u \in \mathcal{U}, j \in \mathcal{J}, \quad \bar{\mathrm{D}}_3 : \pi_{u,j} \leq \kappa_j, \forall u \in \mathcal{U}, j \in \mathcal{J}, & \text{(linear)} \\ \bar{\mathrm{D}}_4 : \pi_{u,j} \geq \chi_u + \kappa_j - 1, \forall u \in \mathcal{U}, j \in \mathcal{J}, & \text{(linear)} \\ \bar{\mathrm{D}}_5 : \frac{\left|\mathbf{h}_u^{\mathrm{H}}\mathbf{m}\psi\right|^2}{\sum_{i\in\mathcal{U}}\left|\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_i\mu_i\right|^2 + \sigma^2} \geq \sum_{j\in\mathcal{J}} \pi_{u,j}\Gamma_j, \forall u \in \mathcal{U}, & \text{(nonconvex)} \\ \bar{\mathrm{D}}_6 : C_u \leq \sum_{j\in\mathcal{J}} \pi_{u,j}R_j, \forall u \in \mathcal{U}. & \text{(linear)} \end{cases}$$

$$\bar{\mathrm{C}}_6, \bar{\mathrm{C}}_9, \bar{\mathrm{D}}_5 \Leftrightarrow \begin{cases} \bar{\mathrm{E}}_1 : \|\mathbf{w}_u\|_2^2 \leq \mu_u P_{\text{tx}}^{\text{max}}, \forall u \in \mathcal{U}, \quad \bar{\mathrm{E}}_2 : \|\mathbf{m}\|_2^2 \leq \psi P_{\text{tx}}^{\text{max}}, \quad \bar{\mathrm{E}}_3 : \sum_{u\in\mathcal{U}} \|\mathbf{w}_u\|_2^2 + \|\mathbf{m}\|_2^2 \leq P_{\text{tx}}^{\text{max}}, & \text{(convex)} \\ \bar{\mathrm{E}}_4 : \frac{\left|\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_u\right|^2}{\Delta_{\text{SIC}}^2\left|\mathbf{h}_u^{\mathrm{H}}\mathbf{m}\right|^2 + \sum_{i\neq u, i\in\mathcal{U}}\left|\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_i\right|^2 + \sigma^2} \geq \sum_{j\in\mathcal{J}} \alpha_{u,j}\Gamma_j, \forall u \in \mathcal{U}, & \text{(nonconvex)} \\ \bar{\mathrm{E}}_5 : \frac{\left|\mathbf{h}_u^{\mathrm{H}}\mathbf{m}\right|^2}{\sum_{i\in\mathcal{U}}\left|\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_i\right|^2 + \sigma^2} \geq \sum_{j\in\mathcal{J}} \pi_{u,j}\Gamma_j, \forall u \in \mathcal{U}. & \text{(nonconvex)} \end{cases}$$

$$\bar{\mathrm{E}}_4, \bar{\mathrm{E}}_4 \Leftrightarrow \begin{cases} \bar{\mathrm{F}}_1 : \frac{\left|\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_u\right|^2}{\Delta_{\text{SIC}}^2\left|\mathbf{h}_u^{\mathrm{H}}\mathbf{m}\right|^2 + \sum_{i\neq u, i\in\mathcal{U}}\left|\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_i\right|^2 + \sigma^2} \geq \alpha_{u,j}\Gamma_j, \forall u \in \mathcal{U}, j \in \mathcal{J}, & \text{(nonconvex)} \\ \bar{\mathrm{F}}_2 : \frac{\left|\mathbf{h}_u^{\mathrm{H}}\mathbf{m}\right|^2}{\sum_{i\in\mathcal{U}}\left|\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_i\right|^2 + \sigma^2} \geq \pi_{u,j}\Gamma_j, \forall u \in \mathcal{U}, j \in \mathcal{J}. & \text{(nonconvex)} \end{cases}$$

$$\bar{\mathrm{F}}_1, \bar{\mathrm{F}}_2 \Leftrightarrow \begin{cases} \bar{\mathrm{G}}_1 : \left\|\left[\mathbf{h}_u^{\mathrm{H}}\bar{\mathbf{W}}_u, \sigma\right]\right\|_2^2 \leq \frac{1}{\Gamma_j}\left|\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_u\right|^2 + (1-\alpha_{u,j})\,L_{\text{max},u}^2, \forall u \in \mathcal{U}, j \in \mathcal{J}, & \text{(nonconvex)} \\ \bar{\mathrm{G}}_2 : \left\|\left[\mathbf{h}_u^{\mathrm{H}}\mathbf{W}, \sigma\right]\right\|_2^2 \leq \frac{1}{\Gamma_j}\left|\mathbf{h}_u^{\mathrm{H}}\mathbf{m}\right|^2 + (1-\pi_{u,j})\,L_{\text{max},u}^2, \forall u \in \mathcal{U}, j \in \mathcal{J}. & \text{(nonconvex)} \end{cases}$$

$$\bar{\mathrm{G}}_1 \Leftrightarrow \begin{cases} \bar{\mathrm{H}}_1 : \mathfrak{Re}\left\{\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_u\right\} \geq 0, \forall u \in \mathcal{U}, \quad \bar{\mathrm{H}}_2 : \mathfrak{Im}\left\{\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_u\right\} = 0, \forall u \in \mathcal{U}, & \text{(linear)} \\ \bar{\mathrm{H}}_3 : \left\|\left[\mathbf{h}_u^{\mathrm{H}}\bar{\mathbf{W}}_u, \sigma\right]\right\|_2 \leq \frac{1}{\sqrt{\Gamma_j}}\mathfrak{Re}\left\{\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_u\right\} + (1-\alpha_{u,j})\,L_{\text{max},u}, \forall u \in \mathcal{U}, j \in \mathcal{J}. & \text{(convex)} \end{cases}$$

$$\bar{\mathrm{G}}_2 \Leftrightarrow \begin{cases} \bar{\mathrm{I}}_1 : \mathfrak{Re}\left\{\mathbf{h}_u^{\mathrm{H}}\mathbf{m}\right\} \geq 0, \forall u \in \mathcal{U}, & \text{(linear)} \\ \bar{\mathrm{I}}_2 : \left\|\left[\mathbf{h}_u^{\mathrm{H}}\mathbf{W}, \sigma\right]\right\|_2 \leq \frac{1}{\sqrt{\Gamma_j}}\mathfrak{Re}\left\{\mathbf{h}_u^{\mathrm{H}}\mathbf{m}\right\} + (1-\pi_{u,j})\,L_{\text{max},u}, \forall u \in \mathcal{U}, j \in \mathcal{J}. & \text{(convex)} \end{cases}$$

Analogous to $\mathcal{P}'_{\text{DWSR}}$, we recast problem $\mathcal{P}'_{\text{DWEE}}$ as

$$\mathcal{P}_{\text{DWEE}} : \min_{\boldsymbol{\Theta}} \frac{\frac{1}{\eta_{\text{eff}}}\left(\sum_{u\in\mathcal{U}}\|\mathbf{w}_u\|_2^2+\|\mathbf{m}\|_2^2\right)+P_{\text{cir}}}{\sum_{u\in\mathcal{U}}\omega_u\left(\sum_{j\in\mathcal{J}}\alpha_{u,j}R_j+C_u\right)} \text{ s.t. } \boldsymbol{\Theta}\in\mathcal{C},$$

where we have transformed the maximization of $f_{\text{DWEE}}(\mathbf{W},\mathbf{m},\mathbf{c},\boldsymbol{\mu},\boldsymbol{\alpha},\psi)$ into the minimization of its reciprocal $\frac{1}{f_{\text{DWEE}}(\mathbf{W},\mathbf{m},\mathbf{c},\boldsymbol{\mu},\boldsymbol{\alpha},\psi)}$. In addition, we have removed the mixed-integer couplings from the objective function, as described in Section III-B2. Hence, the objective function of $\mathcal{P}_{\text{DWEE}}$ is convex. Problems $\mathcal{P}_{\text{DWSR}}$ and $\mathcal{P}_{\text{DWEE}}$ are MISOCPs, which can be solved globally optimal via BnB and IPMs with off-the-shelf solvers, such as `MOSEK` and `GUROBI`, as the problems are convex in the continuous variables.

Remark 3: Due to the inner convexification of the feasible sets of $\mathcal{P}'_{\text{DWSR}}$ ($\mathcal{P}'_{\text{DWEE}}$) in Section III-B6, a globally optimal solution for $\mathcal{P}_{\text{DWSR}}$ ($\mathcal{P}_{\text{DWEE}}$) is feasible for $\mathcal{P}'_{\text{DWSR}}$ ($\mathcal{P}'_{\text{DWEE}}$) but not necessarily globally optimal for $\mathcal{P}'_{\text{DWSR}}$ ($\mathcal{P}'_{\text{DWEE}}$). However, such solution is found to be near-optimal for $\mathcal{P}'_{\text{DWSR}}$ ($\mathcal{P}'_{\text{DWEE}}$), as shown in Section V-A, where we compare $\mathcal{P}_{\text{DWSR}}$ against an upper bound of $\mathcal{P}'_{\text{DWSR}}$, showing a negligible performance loss.

Remark 4: Problems $\mathcal{P}_{\text{DWSR}}$ and $\mathcal{P}_{\text{DWEE}}$ employ discrete rates but allow dynamic rate allocation, i.e., the assigned rates can vary within the designated values of the set of discrete rates $\mathcal{J}$. Our formulation allows to finds optimal rates for unicast and multicast signals, adapting to the channel characteristics and aiming to maximize the objective function.

9) Computational Complexity: Problems $\mathcal{P}_{\text{DWSR}}$ and $\mathcal{P}_{\text{DWEE}}$ involve $N_v = (U+1)N_{\text{tx}} + U$ continuous variables and $N_c = 5UJ + 7U + 4$ linear and convex constraints. The dimension of the underlying SOCP is $N_d = 2JN_{\text{tx}}U^2 + U^2 + 3UJ + 7U + 2UN_{\text{tx}} + 2N_{\text{tx}}$ for any fixed values of binary variables. Therefore, the complexities of problems $\mathcal{P}_{\text{DWSR}}$ and $\mathcal{P}_{\text{DWEE}}$ is $\mathcal{C}_{\texttt{OPT-MISOCP}} = \mathcal{O}(N_p N_c^{0.5} N_v^2 N_d)$, where N_p is the number of solutions evaluated by the BnB solver. The worst-case for N_p is given by $N_p^{\text{all}} = \binom{U}{K}(\sum_{m=0}^{K}\binom{K}{m}J^{K+1-m})$. In practice, $N_p \ll N_p^{\text{all}}$ since BnB methods are capable of pruning infeasible and suboptimal branches, thus reducing the search complexity.

IV. Problem Formulation and Proposed Algorithm for Continuous-Rate RSMA

In this section, we formulate and solve the WSR and WEE maximization problems for RSMA to optimize the beamforming, user admission, and continuous rates, while accounting for imperfect SIC.

A. Problem Formulation

We consider again WSR and WEE maximization as objectives, as in Section III-A. Thus, we define the corresponding optimization problems $\mathcal{Q}'_{\text{CWSR}}$ and $\mathcal{Q}'_{\text{CWEE}}$, shown at the bottom of this page. To account for continuous rates, we have applied the following changes to problems $\mathcal{P}'_{\text{DWSR}}$ and $\mathcal{P}'_{\text{DWEE}}$ in Section III-A. First, we have eliminated binary variables $\boldsymbol{\alpha},\boldsymbol{\kappa}$ used for discrete rate selection. Second, we have reduced the number of binary variables by dropping $\boldsymbol{\mu}$ and only use $\boldsymbol{\chi}$ since $\boldsymbol{\mu}=\boldsymbol{\chi}$, as rates are continuous. Hence, we could remove constraints $\bar{\text{C}}_3, \bar{\text{C}}_4, \bar{\text{C}}_7 - \bar{\text{C}}_{12}$ and employ Shannon's capacity formula to redefine constraints $\bar{\text{C}}_{14}, \bar{\text{C}}_{15}, \bar{\text{C}}_{16}$. Specifically, we have replaced constraint $\bar{\text{C}}_{14}$ with $\bar{\text{C}}_{17}: C_u \leq \psi\chi_u S_{\max}, \forall u\in\mathcal{U}$, and constraint $\bar{\text{C}}_{15}$ with $\bar{\text{C}}_{18}: \sum_{i\in\mathcal{U}} C_i \leq \log_2(1+\mathsf{SINR}_u^{(\text{c})}) + (1-\chi_u)S_{\max}, \forall u\in\mathcal{U}$, where $S_{\max} = \max_{u\in\mathcal{U}}\log_2(1+\frac{P_{\text{tx}}^{\max}}{\sigma^2}\|\mathbf{h}_u\|_2^2)$ is an upper bound for the common rate. Note that $\bar{\text{C}}_{18}$ is tighter when $\chi_u=1$, and therefore the sum of the common rates is bounded by the minimum common rate of all admitted UEs. Finally, we have replaced $\bar{\text{C}}_{16}$ with $\bar{\text{C}}_{19}: \log_2(1+\mathsf{SINR}_u^{(\text{p})}) + C_u \geq R_{\min}\chi_u, \forall u\in\mathcal{U}$, and also redefined the objective functions using Shannon's capacity formula as $f'_{\text{CWSR}}(\mathbf{W},\mathbf{m},\mathbf{c}) \triangleq \sum_{u\in\mathcal{U}}\omega_u(\log_2(1+\mathsf{SINR}_u^{(\text{p})})+C_u)$ and $f'_{\text{CWEE}}(\mathbf{W},\mathbf{m},\mathbf{c}) \triangleq \frac{\sum_{u\in\mathcal{U}}\omega_u(\log_2(1+\mathsf{SINR}_u^{(\text{p})})+C_u)}{\frac{1}{\eta_{\text{eff}}}(\sum_{u\in\mathcal{U}}\|\mathbf{w}_u\chi_u\|_2^2+\|\mathbf{m}\psi\|_2^2)+P_{\text{cir}}}$.

Remark 5: Problems $\mathcal{Q}'_{\text{CWSR}}$, $\mathcal{Q}'_{\text{CWEE}}$ are nonconvex MINLPs, and compared to $\mathcal{P}'_{\text{CWSR}}$, $\mathcal{P}'_{\text{CWEE}}$, they assume continuous rates. In addition, their structure is more complex than that

$$\begin{array}{llll}
\begin{array}{l}\mathcal{Q}'_{\text{CWSR}}\\ \mathcal{Q}'_{\text{CWEE}}\end{array} : \max\limits_{\mathbf{W},\mathbf{m},\mathbf{c},\boldsymbol{\chi},\psi} & \left\{\begin{array}{l} f'_{\text{CWSR}}(\mathbf{W},\mathbf{m},\mathbf{c}) \triangleq \sum_{u\in\mathcal{U}}\omega_u\left(\log_2\left(1+\mathsf{SINR}_u^{(\text{p})}\right)+C_u\right) \\ f'_{\text{CWEE}}(\mathbf{W},\mathbf{m},\mathbf{c}) \triangleq \frac{\sum_{u\in\mathcal{U}}\omega_u\left(\log_2\left(1+\mathsf{SINR}_u^{(\text{p})}\right)+C_u\right)}{\frac{1}{\eta_{\text{eff}}}\left(\sum_{u\in\mathcal{U}}\|\mathbf{w}_u\chi_u\|_2^2+\|\mathbf{m}\psi\|_2^2\right)+P_{\text{cir}}}\end{array}\right. & & \begin{array}{r}(\text{nonconvex})\\(\text{nonconvex})\end{array}\\
\text{s.t.} & \bar{\text{C}}_1: \chi_u\in\{0,1\}, \forall u\in\mathcal{U}, & & (\text{binary})\\
& \bar{\text{C}}_2: \sum_{u\in\mathcal{U}}\chi_u = K, & & (\text{linear})\\
& \bar{\text{C}}_5: \psi\in\{0,1\}, & & (\text{binary})\\
& \bar{\text{C}}_6: \sum_{u\in\mathcal{U}}\|\mathbf{w}_u\chi_u\|_2^2+\|\mathbf{m}\psi\|_2^2 \leq P_{\text{tx}}^{\max}, & & (\text{nonconvex})\\
& \bar{\text{C}}_{13}: C_u\geq 0, \forall u\in\mathcal{U}, & & (\text{linear})\\
& \bar{\text{C}}_{17}: C_u\leq\psi\chi_u S_{\max}, \forall u\in\mathcal{U}, & & (\text{linear})\\
& \bar{\text{C}}_{18}: \sum_{i\in\mathcal{U}}C_i \leq \log_2\left(1+\mathsf{SINR}_u^{(\text{c})}\right)+(1-\chi_u)S_{\max}, \forall u\in\mathcal{U}, & & (\text{nonconvex})\\
& \bar{\text{C}}_{19}: \log_2\left(1+\mathsf{SINR}_u^{(\text{p})}\right)+C_u\geq R_{\min}\chi_u, \forall u\in\mathcal{U}. & & (\text{nonconvex})
\end{array}$$

of $\mathcal{P}'_{\text{CWSR}}, \mathcal{P}'_{\text{CWEE}}$, as they involve multiplicative couplings of continuous variables, which are not present in $\mathcal{P}'_{\text{CWSR}}, \mathcal{P}'_{\text{CWEE}}$.

B. Proposed Algorithm

To solve $\mathcal{Q}'_{\text{CWSR}}$ and $\mathcal{Q}'_{\text{CWEE}}$, we leverage successive convex approximation (SCA), semidefinite relaxation (SDR), and binary enumeration. In particular, we enumerate all combinations of admitted UEs and then solve the underlying nonconvex subproblem for each combination via the proposed `OPT-SCA-SDR` algorithm, which finds a KKT point by exploiting SCA and SDR. In the following, we describe the proposed algorithm by considering $\mathcal{Q}'_{\text{CWSR}}$ and then we extend it to $\mathcal{Q}'_{\text{CWEE}}$.

1) Enumerating the Binary Variables: Let N be the total number of combinations of admitted UEs and $\mathcal{N} = \{1, \ldots, N\}$ the set collecting them. Considering a given combination $n \in \mathcal{N}$, problem $\mathcal{Q}'_{\text{CWSR}}$ reduces to $\mathcal{Q}_{\text{CWSR}_n}$, shown at the bottom of this page. In particular, $f_{\text{CWSR}_n}(\mathbf{W}, \mathbf{m}, \mathbf{c}) \triangleq \sum_{u\in\mathcal{U}'_n} \omega_u(\log_2(1 + \frac{|\mathbf{h}_u^{\text{H}}\mathbf{w}_u|^2}{\Delta^2_{\text{SIC}}|\mathbf{h}_u^{\text{H}}\mathbf{m}|^2+\sum_{i\neq u, i\in\mathcal{U}'_n}|\mathbf{h}_u^{\text{H}}\mathbf{w}_i|^2+\sigma^2}) + C_u)$, and $\mathcal{U}'_n \subseteq \mathcal{U}$ denotes the set of admitted UEs in combination n, such that $\mu_u = 1$, $\forall u \in \mathcal{U}'_n$ and $|\mathcal{U}'_n| = K$. For notational simplicity, we reset the UE indices in $\mathcal{U}'_n$, such that $\mathcal{U}'_n = \{1, \ldots, K\}$. Here, constraint $\bar{\text{C}}_{20}$ is included to eliminate the coupling $\mathbf{m}\psi_0$ in an analogous manner as in Section III-B2 for constraint E_1. We have not included C_{17} because it is implied by $\bar{\text{C}}_{18}, \bar{\text{C}}_{20}$ when ψ_0 is given. We adopt $\psi_0 = 1$ for RSMA and $\psi_0 = 0$ for SDMA.

2) Transforming the Problem Via Sublevel and Superlevel Sets: We introduce nonnegative variables $\boldsymbol{\gamma} \in \mathbb{R}^K_+, \boldsymbol{\rho} \in \mathbb{R}^K_+, \boldsymbol{\lambda} \in \mathbb{R}^K_+, \boldsymbol{\tau} \in \mathbb{R}^K_+$, and $\beta \in \mathbb{R}_+$ to define sublevel and superlevel sets, thereby transforming problem $\mathcal{Q}_{\text{CWSR}_n}$ into $\widetilde{\mathcal{Q}}_{\text{CWSR}_n}$. In Appendix H, available online, we show that $\mathcal{Q}_{\text{CWSR}_n}$ and $\widetilde{\mathcal{Q}}_{\text{CWSR}_n}$ are equivalent. Specifically, we bound the private SINRs from below via $\frac{|\mathbf{h}_u^{\text{H}}\mathbf{w}_u|^2}{\Delta^2_{\text{SIC}}|\mathbf{h}_u^{\text{H}}\mathbf{m}|^2+\sum_{i\neq u, i\in\mathcal{U}'_n}|\mathbf{h}_u^{\text{H}}\mathbf{w}_i|^2+\sigma^2} \geq \gamma_u - 1$. Also, we bound the interference at each UE from above by including $\Delta^2_{\text{SIC}}|\mathbf{h}_u^{\text{H}}\mathbf{m}|^2 + \sum_{i\neq u, i\in\mathcal{U}'_n}|\mathbf{h}_u^{\text{H}}\mathbf{w}_i|^2 + \sigma^2 \leq \rho_u$. Following the same idea, we include $\frac{|\mathbf{h}_u^{\text{H}}\mathbf{m}|^2}{\sum_{i\in\mathcal{U}'_n}|\mathbf{h}_u^{\text{H}}\mathbf{w}_i|^2+\sigma^2} \geq \tau_u - 1$ and $\sum_{i\in\mathcal{U}'_n}|\mathbf{h}_u^{\text{H}}\mathbf{w}_i|^2 + \sigma^2 \leq \lambda_u$ to bound the common SINRs and the interference. Furthermore, we bound the objective function from below, such that $f_{\text{CWSR}_n}(\mathbf{W}, \mathbf{m}, \mathbf{c}) \geq \beta$, thus defining a new objective function $f_{\text{CWSR}_n}(\beta) \triangleq \beta$. Upon applying these transformations to $\mathcal{Q}_{\text{CWSR}_n}$, we obtain problem $\widetilde{\mathcal{Q}}_{\text{CWSR}_n}$, shown at the bottom of the next page.

3) Leveraging Semidefinite Programming: By employing semidefinite programming and introducing positive semidefinite variables $\mathbf{W}_u \in \mathbb{C}^{N_{\text{tx}}\times N_{\text{tx}}}$ and $\mathbf{M} \in \mathbb{C}^{N_{\text{tx}}\times N_{\text{tx}}}$, which replace $\mathbf{w}_u\mathbf{w}_u^{\text{H}}$ and $\mathbf{m}\mathbf{m}^{\text{H}}$, respectively, constraints $\bar{\text{C}}_6, \bar{\text{C}}_{20}, \bar{\text{K}}_1, \bar{\text{K}}_2, \bar{\text{K}}_4, \bar{\text{K}}_5$ can be equivalently reformulated to $\bar{\text{L}}_1 - \bar{\text{L}}_{10}$, shown at the bottom of the next page. In doing so, the nonconvexity of constraints $\bar{\text{K}}_1, \bar{\text{K}}_4$ are circumvented in part as the quadratic terms on the left-hand side are linearized. The newly introduced variables also affect constraints $\bar{\text{K}}_2, \bar{\text{K}}_5, \bar{\text{C}}_6, \bar{\text{C}}_{20}$. The positive semidefiniteness of $\mathbf{W}_u$ and $\mathbf{M}$ are specified by $\bar{\text{L}}_7, \bar{\text{L}}_8$, whereas $\bar{\text{L}}_9, \bar{\text{L}}_{10}$ allow for the private and common signals to be used or not. Considering the equivalence between $\bar{\text{C}}_6, \bar{\text{C}}_{20}, \bar{\text{K}}_1, \bar{\text{K}}_2, \bar{\text{K}}_4, \bar{\text{K}}_5$ and $\bar{\text{L}}_1 - \bar{\text{L}}_{10}$, we define problem

$$\widehat{\mathcal{Q}}_{\text{CWSR}_n}: \max_{\widehat{\mathbf{W}}, \mathbf{M}, \mathbf{c}, \boldsymbol{\gamma}, \boldsymbol{\rho}, \boldsymbol{\lambda}, \boldsymbol{\tau}, \beta} \quad f_{\text{CWSR}_n}(\beta)$$
$$\text{s.t.} \quad \bar{\text{C}}_{13}, \bar{\text{K}}_3, \bar{\text{K}}_6 - \bar{\text{K}}_8, \bar{\text{L}}_1 - \bar{\text{L}}_{10},$$

where $\widehat{\mathbf{W}} = (\mathbf{W}_1, \ldots, \mathbf{W}_K)$. We note that $\widehat{\mathcal{Q}}_{\text{CWSR}_n}$ is equivalent to $\widetilde{\mathcal{Q}}_{\text{CWSR}_n}$ and $\mathcal{Q}_{\text{CWSR}_n}$ since the feasible set and objective function are not affected by the applied transformation of the constraints.

4) Addressing the Nonconvex Constraints: To cope with the nonconvex constraints $\bar{\text{L}}_3, \bar{\text{L}}_5, \bar{\text{L}}_9, \bar{\text{L}}_{10}$, we adopt an iterative approach whereby we sequentially approximate these constraints by convex approximations.

• *Quasi-convex constraints:* To circumvent the quasi-convex constraints $\bar{\text{L}}_3, \bar{\text{L}}_5$, we replace them with the inner convex approximations $\bar{\text{M}}_1, \bar{\text{M}}_3$, shown at the bottom of the next page, where t is the iteration index, and $\bar{\Omega}^{(t)}_{1,u}$, $\bar{\Omega}^{(t)}_{2,u}$, $u \in \mathcal{U}'_n$, are parameters adapted iteratively. In recasting $\bar{\text{L}}_3, \bar{\text{L}}_5$ as $\bar{\text{M}}_1, \bar{\text{M}}_2$, we have employed the arithmetic-geometric mean inequality, which states that $ab \leq \frac{a^2}{2} + \frac{b^2}{2}$ for $a, b \in \mathbb{R}_+$. By introducing a new parameter $\Phi \in \mathbb{R}_+$ and applying transformations $a \leftarrow \sqrt{\Phi}a$ and $b \leftarrow \sqrt{\frac{1}{\Phi}}b$, we obtain inequality $ab \leq \frac{\Phi}{2}a^2 + \frac{1}{2\Phi}b^2$, which becomes tight when $\Phi = \frac{b}{a}$ [46]. Note that $\frac{\Phi}{2}a^2 + \frac{1}{2\Phi}b^2$ is a convex overestimate of ab, which decouples a and b, allowing to circumvent the nonconvexity of the product ab. Exploiting

$$\mathcal{Q}_{\text{CWSR}_n}: \max_{\mathbf{W},\mathbf{m},\mathbf{c}} \quad f_{\text{CWSR}_n}(\mathbf{W}, \mathbf{m}, \mathbf{c}) \qquad \text{(nonconvex)}$$

$$\text{s.t.} \quad \bar{\text{C}}_6: \sum_{u\in\mathcal{U}'_n}\|\mathbf{w}_u\|_2^2 + \|\mathbf{m}\|_2^2 \leq P_{\text{tx}}^{\max}, \qquad \text{(convex)}$$

$$\bar{\text{C}}_{13}: C_u \geq 0, \forall u \in \mathcal{U}'_n, \qquad \text{(linear)}$$

$$\bar{\text{C}}_{18}: \sum_{i\in\mathcal{U}'_n} C_i \leq \log_2\left(1 + \frac{\left|\mathbf{h}_u^{\text{H}}\mathbf{m}\right|^2}{\sum_{i\in\mathcal{U}'_n}\left|\mathbf{h}_u^{\text{H}}\mathbf{w}_i\right|^2+\sigma^2}\right), \forall u \in \mathcal{U}'_n, \qquad \text{(nonconvex)}$$

$$\bar{\text{C}}_{19}: \log_2\left(1 + \frac{\left|\mathbf{h}_u^{\text{H}}\mathbf{w}_u\right|^2}{\Delta^2_{\text{SIC}}|\mathbf{h}_u^{\text{H}}\mathbf{m}|^2+\sum_{i\neq u, i\in\mathcal{U}'_n}\left|\mathbf{h}_u^{\text{H}}\mathbf{w}_i\right|^2+\sigma^2}\right) + C_u \geq R_{\min}, \forall u \in \mathcal{U}'_n, \qquad \text{(nonconvex)}$$

$$\bar{\text{C}}_{20}: \|\mathbf{m}\|_2^2 \leq \psi_0 P_{\text{tx}}^{\max}, \qquad \text{(convex)}$$

this observation, we introduce parameter $\bar{\Omega}_{1,u}^{(t)}$ and apply the parameterized inequality to the product $\gamma_u \rho_u$ in $\bar{\mathrm{L}}_3$, such that $\gamma_u \rho_u \leq \frac{\bar{\Omega}_{1,u}^{(t)}}{2}\gamma_u^2 + \frac{1}{2\bar{\Omega}_{1,u}^{(t)}}\rho_u^2$. We proceed in a similar manner with constraint $\bar{\mathrm{L}}_5$ by introducing $\Omega_{2,u}^{(t)}$, which yields $\tau_u \lambda_u \leq \frac{\bar{\Omega}_{2,u}^{(t)}}{2}\tau_u^2 + \frac{1}{2\bar{\Omega}_{2,u}^{(t)}}\lambda_u^2$. Next, we replace products $\gamma_u \rho_u$ and $\tau_u \lambda_u$ with their respective convex overestimates, $\frac{\bar{\Omega}_{1,u}^{(t)}}{2}\gamma_u^2 + \frac{1}{2\bar{\Omega}_{1,u}^{(t)}}\rho_u^2$ and $\frac{\bar{\Omega}_{2,u}^{(t)}}{2}\tau_u^2 + \frac{1}{2\bar{\Omega}_{2,u}^{(t)}}\lambda_u^2$, thus yielding $\bar{\mathrm{M}}_1, \bar{\mathrm{M}}_2$. In each iteration t, we update the parameters according to $\bar{\Omega}_{1,u}^{(t)} = \frac{\rho_u^{(t-1)}}{\gamma_u^{(t-1)}}$ and $\bar{\Omega}_{2,u}^{(t)} = \frac{\lambda_u^{(t-1)}}{\tau_u^{(t-1)}}$ making it possible to sequentially adapt the convex approximation. Upon replacing $\bar{\mathrm{L}}_3, \bar{\mathrm{L}}_5$ with $\bar{\mathrm{M}}_1, \bar{\mathrm{M}}_2$ in problem $\widehat{\mathcal{Q}}_{\mathrm{CWSR}_n}$, and then solving it, an optimal solution to this modified problem will be feasible for $\widehat{\mathcal{Q}}_{\mathrm{CWSR}_n}$, $\widetilde{\mathcal{Q}}_{\mathrm{CWSR}_n}$, and $\mathcal{Q}_{\mathrm{CWSR}_n}$ since the feasible set of $\bar{\mathrm{M}}_1, \bar{\mathrm{M}}_2$ is contained in that of $\bar{\mathrm{L}}_3, \bar{\mathrm{L}}_5$. However, the solution will not necessarily be globally optimal for $\widehat{\mathcal{Q}}_{\mathrm{CWSR}_n}$, $\widetilde{\mathcal{Q}}_{\mathrm{CWSR}_n}$, and $\mathcal{Q}_{\mathrm{CWSR}_n}$ due to the possible reduction of the feasible set caused by the inner convexification in $\bar{\mathrm{M}}_1, \bar{\mathrm{M}}_2$.

• *Rank constraints:* In order to cope with rank constraints $\bar{\mathrm{L}}_9, \bar{\mathrm{L}}_{10}$, we adopt the iterative method proposed in [47], which is described as follows. We first reformulate $\bar{\mathrm{L}}_9$, $\bar{\mathrm{L}}_{10}$ as $\bar{\mathrm{M}}_3$, $\bar{\mathrm{M}}_4$, shown at the bottom of this page. Then, we penalize the objective function by adding cost function $\sum_{u \in \mathcal{U}_n' \cup \{0\}} p_u^{(t)} \zeta_u$, which promotes rank minimization and enforces $\mathbf{M}$ and $\mathbf{W}_u$ to have rank at most one, as shown in Appendix I, available online. The ζ_u, $\forall u \in \mathcal{U}_n' \cup \{0\}$, are slack variables and $p_u^{(t)} \in \mathbb{R}_+$, $\forall u \in \mathcal{U}_n' \cup \{0\}$, represent the penalty weights in iteration t. Matrices $\bar{\mathbf{M}}^{(t-1)}$ and $\bar{\mathbf{W}}_u^{(t-1)}$ are the respective solutions for $\mathbf{M}$ and $\mathbf{W}_u$, obtained in iteration $t-1$. Also, $\mathbf{T}_0^{(t)} \in \mathbb{C}^{N_{\mathrm{tx}} \times (N_{\mathrm{tx}}-1)}$ is formed by the eigenvectors of the $N_{\mathrm{tx}}-1$ smallest

$$
\begin{aligned}
\widetilde{\mathcal{Q}}_{\mathrm{CWSR}_n}: \max_{\mathbf{W},\mathbf{m},\mathbf{c},\boldsymbol{\gamma},\boldsymbol{\rho},\boldsymbol{\lambda},\boldsymbol{\tau},\beta} \quad & f_{\mathrm{CWSR}_n}(\beta) \triangleq \beta \\
\text{s.t.} \quad & \bar{\mathrm{K}}_1: \left|\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_u\right|^2 \geq (\gamma_u - 1)\rho_u, \forall u \in \mathcal{U}_n', && \text{(nonconvex)} \\
& \bar{\mathrm{K}}_2: \Delta_{\mathrm{SIC}}^2 \left|\mathbf{h}_u^{\mathrm{H}}\mathbf{m}\right|^2 + \textstyle\sum_{i \neq u, i \in \mathcal{U}_n'} \left|\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_i\right|^2 + \sigma^2 \leq \rho_u, \forall u \in \mathcal{U}_n', && \text{(convex)} \\
& \bar{\mathrm{K}}_3: \beta - \textstyle\sum_{u \in \mathcal{U}_n'} \omega_u (\log_2(\gamma_u) + C_u) \leq 0, && \text{(convex)} \\
& \bar{\mathrm{K}}_4: \left|\mathbf{h}_u^{\mathrm{H}}\mathbf{m}\right|^2 \geq (\tau_u - 1)\lambda_u, \forall u \in \mathcal{U}_n', && \text{(nonconvex)} \\
& \bar{\mathrm{K}}_5: \textstyle\sum_{i \in \mathcal{U}_n'} \left|\mathbf{h}_u^{\mathrm{H}}\mathbf{w}_i\right|^2 + \sigma^2 \leq \lambda_u, \forall u \in \mathcal{U}_n', && \text{(convex)} \\
& \bar{\mathrm{K}}_6: \textstyle\sum_{i \in \mathcal{U}_n'} C_i - \log_2(\tau_u) \leq 0, \forall u \in \mathcal{U}_n', && \text{(convex)} \\
& \bar{\mathrm{K}}_7: R_{\min} - \log_2(\gamma_u) - C_u \leq 0, \forall u \in \mathcal{U}_n', && \text{(convex)} \\
& \bar{\mathrm{K}}_8: \beta \geq 0, && \text{(linear)} \\
& \qquad \bar{\mathrm{C}}_6, \bar{\mathrm{C}}_{13}, \bar{\mathrm{C}}_{20}.
\end{aligned}
$$

$$
\bar{\mathrm{C}}_6, \bar{\mathrm{C}}_{20}, \bar{\mathrm{K}}_1, \bar{\mathrm{K}}_2, \bar{\mathrm{K}}_4, \bar{\mathrm{K}}_5 \Leftrightarrow
\begin{cases}
\bar{\mathrm{L}}_1: \sum_{u \in \mathcal{U}_n'} \mathrm{Tr}(\mathbf{W}_u) + \mathrm{Tr}(\mathbf{M}) \leq P_{\mathrm{tx}}^{\max}, \quad \bar{\mathrm{L}}_2: \mathrm{Tr}(\mathbf{M}) \leq \psi_0 P_{\mathrm{tx}}^{\max}, & \text{(linear)} \\
\bar{\mathrm{L}}_3: (\gamma_u - 1)\rho_u - \mathbf{h}_u^{\mathrm{H}}\mathbf{W}_u\mathbf{h}_u \leq 0, \forall u \in \mathcal{U}_n', & \text{(nonconvex)} \\
\bar{\mathrm{L}}_4: \Delta_{\mathrm{SIC}}^2 \mathbf{h}_u^{\mathrm{H}}\mathbf{M}\mathbf{h}_u + \sum_{i \neq u, i \in \mathcal{U}_n'} \mathbf{h}_u^{\mathrm{H}}\mathbf{W}_i\mathbf{h}_u + \sigma^2 \leq \rho_u, \forall u \in \mathcal{U}_n', & \text{(linear)} \\
\bar{\mathrm{L}}_5: (\tau_u - 1)\lambda_u - \mathbf{h}_u^{\mathrm{H}}\mathbf{M}\mathbf{h}_u \leq 0, \forall u \in \mathcal{U}_n', & \text{(nonconvex)} \\
\bar{\mathrm{L}}_6: \sum_{i \in \mathcal{U}_n'} \mathbf{h}_u^{\mathrm{H}}\mathbf{W}_i\mathbf{h}_u + \sigma^2 \leq \lambda_u, \forall u \in \mathcal{U}_n', & \text{(linear)} \\
\bar{\mathrm{L}}_7: \mathbf{W}_u \succcurlyeq 0, \forall u \in \mathcal{U}_n', & \text{(linear)} \\
\bar{\mathrm{L}}_8: \mathbf{M} \succcurlyeq 0, & \text{(linear)} \\
\bar{\mathrm{L}}_9: \mathrm{Rank}(\mathbf{W}_u) \leq 1, \forall u \in \mathcal{U}_n', \quad \bar{\mathrm{L}}_{10}: \mathrm{Rank}(\mathbf{M}) \leq \psi_0. & \text{(nonconvex)}
\end{cases}
$$

$$
\bar{\mathrm{L}}_3, \bar{\mathrm{L}}_5 \Leftrightarrow
\begin{cases}
\bar{\mathrm{M}}_1: \frac{\bar{\Omega}_{1,u}^{(t)}}{2}\gamma_u^2 + \frac{1}{2\bar{\Omega}_{1,u}^{(t)}}\rho_u^2 - \rho_u - \mathbf{h}_u^{\mathrm{H}}\mathbf{W}_u\mathbf{h}_u \leq 0, \forall u \in \mathcal{U}_n', & \text{(convex)} \\
\bar{\mathrm{M}}_2: \frac{\bar{\Omega}_{2,u}^{(t)}}{2}\tau_u^2 + \frac{1}{2\bar{\Omega}_{2,u}^{(t)}}\lambda_u^2 - \lambda_u - \mathbf{h}_u^{\mathrm{H}}\mathbf{M}\mathbf{h}_u \leq 0, \forall u \in \mathcal{U}_n', & \text{(convex)}
\end{cases}
$$

$$
\bar{\mathrm{L}}_9, \bar{\mathrm{L}}_{10} \Leftrightarrow \left\{ \bar{\mathrm{M}}_3: \zeta_0 \mathbf{I} - {\mathbf{T}_0^{(t)}}^{\mathrm{H}} \mathbf{M} \mathbf{T}_0^{(t)} \succcurlyeq \mathbf{0}, \quad \bar{\mathrm{M}}_4: \zeta_u \mathbf{I} - {\mathbf{T}_u^{(t)}}^{\mathrm{H}} \mathbf{W}_u \mathbf{T}_u^{(t)} \succcurlyeq \mathbf{0}, \forall u \in \mathcal{U}_n', \quad \text{(convex)} \right.
$$

eigenvalues of $\bar{\mathbf{M}}^{(t-1)}$, whereas $\mathbf{T}_u^{(t)} \in \mathbb{C}^{N_{\text{tx}} \times (N_{\text{tx}}-1)}$ is formed by the eigenvectors of the $N_{\text{tx}} - 1$ smallest eigenvalues of $\bar{\mathbf{W}}_u^{(t-1)}$.

5) Outlining the Algorithm and Its Extension to Solve $\mathcal{Q}'_{\text{CWEE}}$: The transformation of constraints $\bar{\text{L}}_3, \bar{\text{L}}_5, \bar{\text{L}}_9, \bar{\text{L}}_{10}$ into $\bar{\text{M}}_1 - \bar{\text{M}}_4$, leads to the following problem

$$\begin{aligned} \bar{\mathcal{Q}}^{(t)}_{\text{CWSR}_n}: \quad & \max_{\substack{\widehat{\mathbf{W}}, \mathbf{M}, \mathbf{c}, \boldsymbol{\gamma}, \boldsymbol{\rho}, \\ \boldsymbol{\lambda}, \boldsymbol{\tau}, \boldsymbol{\zeta}, \beta}} \quad f_{\text{CWSR}_n}(\beta) - \textstyle\sum_{u \in \mathcal{U}'_n \cup \{0\}} p_u^{(t)} \zeta_u \\ & \quad \text{s.t.} \qquad \bar{\text{C}}_{13}, \bar{\text{K}}_3, \bar{\text{K}}_6 - \bar{\text{K}}_8, \bar{\text{L}}_1, \bar{\text{L}}_2, \\ & \qquad\qquad\quad \bar{\text{L}}_4, \bar{\text{L}}_6 - \bar{\text{L}}_8, \bar{\text{M}}_1 - \bar{\text{M}}_4. \end{aligned}$$

On the other hand, to solve $\mathcal{Q}'_{\text{CWEE}}$, we introduce the following problem

$$\begin{aligned} \bar{\mathcal{Q}}^{(t)}_{\text{CWEE}_n}: \quad & \max_{\substack{\widehat{\mathbf{W}}, \mathbf{M}, \mathbf{c}, \boldsymbol{\gamma}, \boldsymbol{\rho}, \\ \boldsymbol{\lambda}, \boldsymbol{\tau}, \boldsymbol{\zeta}, \beta, \theta, \delta}} \quad f_{\text{CWEE}_n}(\theta) - \textstyle\sum_{u \in \mathcal{U}'_n \cup \{0\}} p_u^{(t)} \zeta_u \\ & \quad \text{s.t.} \qquad \bar{\text{C}}_{13}, \bar{\text{K}}_3, \bar{\text{K}}_6 - \bar{\text{K}}_8, \bar{\text{L}}_1, \bar{\text{L}}_2, \bar{\text{L}}_4, \\ & \qquad\qquad\quad \bar{\text{L}}_6 - \bar{\text{L}}_8, \bar{\text{M}}_1 - \bar{\text{M}}_4, \bar{\text{N}}_1 - \bar{\text{N}}_3, \end{aligned}$$

where we employed the same procedure as described in Section IV-B1 to Section IV-B4. Compared to $\bar{\mathcal{Q}}^{(t)}_{\text{CWSR}_n}$, problem $\bar{\mathcal{Q}}^{(t)}_{\text{CWEE}_n}$ features variables θ and δ, convex constraints $\bar{\text{N}}_1 : \sum_{u \in \mathcal{U}'_n} \text{Tr}(\mathbf{W}_u) + \text{Tr}(\mathbf{M}) \leq \eta_{\text{eff}} \delta$, $\bar{\text{N}}_2 : \frac{\bar{\Omega}_3^{(t)}}{2} \theta^2 + \frac{1}{2\bar{\Omega}_3^{(t)}} \delta^2 + \theta P_{\text{cir}} \leq \beta$ and $\bar{\text{N}}_3 : \theta \geq 0$, and parameter $\bar{\Omega}_3^{(t)}$. In particular, θ is used to bound the objective function $f_{\text{CWEE}_n}(\mathbf{W}, \mathbf{m}, \mathbf{c})$ from below. Variable δ is used to bound the transmit power efficiency from above, thereby yielding constraint $\bar{\text{N}}_1$. The introduction of δ and θ in the objective function leads to a multiplicative coupling $\theta\delta$, which is dealt with in the same manner as in Section IV-B4, yielding $\bar{\text{N}}_2$. Also, $\bar{\text{N}}_3$ is added to ensure the positiveness of the objective function. Parameter $\bar{\Omega}_3^{(t)}$ is updated as $\bar{\Omega}_3^{(t)} = \frac{\delta^{(t-1)}}{\theta^{(t-1)}}$ and the objective function is penalized by $\sum_{u \in \mathcal{U}'_n \cup \{0\}} p_u^{(t)} \zeta_u$.

Problems $\bar{\mathcal{Q}}^{(t)}_{\text{CWSR}_n}$ and $\bar{\mathcal{Q}}^{(t)}_{\text{CWEE}_n}$ are convex and can be solved optimally via IPMs. Both are solved iteratively, improving the objective function in each iteration until a stop criterion is met, i.e., the difference of the objective function values between successive iterations is less than a threshold ϵ or the number of iterations exceeds N_{iter}. In Appendix J, available online, we show that $\bar{\mathcal{Q}}^{(t)}_{\text{CWSR}_n}$ and $\bar{\mathcal{Q}}^{(t)}_{\text{CWEE}_n}$ converges to a KKT point. Also, by increasing the penalty weights $p_u^{(t)}$, variables ζ_u decrease in each iteration, leading to $\zeta_u \to 0$ and $\sum_{u \in \mathcal{U}'_n \cup \{0\}} p_u^{(t)} \zeta_u \to 0$. This causes $\mathbf{M}$ and $\mathbf{W}_u$ to have at most rank one, since the $N_{\text{tx}} - 1$ smallest eigenvalues of these matrices are progressively squeezed to zero. Assuming that $\bar{\mathcal{Q}}^{(t)}_{\text{CWSR}_n}$ converge in iteration $t^\star$, we have that $\mathbf{M} \approx \bar{\mathbf{M}}^{(t^\star-1)}$, where $\mathbf{M}$ is the solution in iteration $t^\star$. Via eigendecomposition of $\mathbf{M}$, we have $\mathbf{M} = \widetilde{\mathbf{R}}_0 \boldsymbol{\Sigma}_0 \widetilde{\mathbf{R}}_0^{\text{H}}$, such that $\widetilde{\mathbf{R}}_0 \widetilde{\mathbf{R}}_0^{\text{H}} = \mathbf{I}$, $\boldsymbol{\Sigma}_0 = \text{diag}(\sigma_{0,1}, \ldots, \sigma_{0,N_{\text{tx}}})$, and $\widetilde{\mathbf{R}}_0 = [\mathbf{r}_0 | \mathbf{R}_0]$. Therefore, ${\mathbf{T}_0^{(t^\star)}}^{\text{H}} \mathbf{M} \mathbf{T}_0^{(t^\star)} = {\mathbf{T}_0^{(t^\star)}}^{\text{H}} [\mathbf{r}_0 | \mathbf{R}_0] \boldsymbol{\Sigma}_0 [\mathbf{r}_0 | \mathbf{R}_0]^{\text{H}} \mathbf{T}_0^{(t^\star)}$, which can be further reduced to ${\mathbf{T}_0^{(t^\star)}}^{\text{H}} \mathbf{M} \mathbf{T}_0^{(t^\star)} = [\mathbf{0} | \mathbf{I}] \boldsymbol{\Sigma}_0 [\mathbf{0} | \mathbf{I}]^{\text{H}} = \text{diag}(\sigma_{0,2}, \ldots, \sigma_{0,N_{\text{tx}}})$, since ${\mathbf{T}_0^{(t^\star)}}^{\text{H}} \mathbf{r}_0 \approx \mathbf{0}$ and ${\mathbf{T}_0^{(t^\star)}}^{\text{H}} \mathbf{R}_0 \approx \mathbf{I}$. Considering these outcomes and $\bar{\text{M}}_3$, we obtain $\zeta_0 \mathbf{I} \succcurlyeq \text{diag}(\sigma_{0,2}, \ldots, \sigma_{0,N_{\text{tx}}})$, which leads to $\sigma_{0,2}, \ldots, \sigma_{0,N_{\text{tx}}} \to 0$ as $\zeta_0 \to 0$. Because $\sigma_{0,1}$ is not affected by this procedure, $\sigma_{0,1}$ can be different from zero or even zero, i.e., $\mathbf{M}$ can be at most rank-one. Following the same reasoning, we can obtain equivalent results for ${\mathbf{T}_u^{(t^\star)}}^{\text{H}} \mathbf{W}_u \mathbf{T}_u^{(t^\star)}, \forall u \in \mathcal{U}'_n$. The solutions that satisfy $\bar{\text{L}}_9, \bar{\text{L}}_{10}$ are recovered via eigendecomposition of $\mathbf{M}$ and $\mathbf{W}_u$, i.e., $\mathbf{m} = \sqrt{\sigma_{0,1}} \mathbf{r}_0$ and $\mathbf{w}_u = \sqrt{\sigma_{u,1}} \mathbf{r}_u$, where $\sigma_{u,1}$ and $\mathbf{r}_u$ are the largest eigenvalue and principal eigenvector of $\mathbf{W}_u$, respectively. The same analysis applies to $\bar{\mathcal{Q}}^{(t)}_{\text{CWEE}_n}$ as the constraints are the same.

6) Projecting the Continuous Rates: Due to the use of Shannon's capacity formula, the rates obtained by solving $\bar{\mathcal{Q}}^{(t)}_{\text{CWSR}_n}$ and $\bar{\mathcal{Q}}^{(t)}_{\text{CWEE}_n}$ are continuous. To meet the MCS specifications, these rates are projected, i.e., approximated to the closest feasible discrete rates. Thus, the best solution with projected rates is given by $f^{\text{proj}}_{\text{CWSR}}(\mathbf{W}, \mathbf{m}, \mathbf{c}) \triangleq \max_{n \in \mathcal{N}} \sum_{u \in \mathcal{U}'_n} \omega_u (R^{\text{proj}}_{u,n} + C^{\text{proj}}_{u,n})$ and $f^{\text{proj}}_{\text{CWEE}}(\mathbf{W}, \mathbf{m}, \mathbf{c}) \triangleq \max_{n \in \mathcal{N}} \frac{\sum_{u \in \mathcal{U}} \omega_u (\log_2(1 + \overline{\mathsf{SINR}}^{(\text{p})}_{u,n}) + \bar{C}_{u,n})}{\frac{1}{\eta_{\text{eff}}} (\sum_{u \in \mathcal{U}} \|\bar{\mathbf{w}}_{u,n}\|_2^2 + \|\bar{\mathbf{m}}_n\|_2^2) + P_{\text{cir}}}$, where $R^{\text{proj}}_{u,n}$ and $\bar{C}^{\text{proj}}_{u,n}$ are defined in (1) and (2)

$$R^{\text{proj}}_{u,n} = \left\{ R_j \mid j = \arg\min_{i \in \mathcal{J}} \overline{\mathsf{SINR}}^{(\text{p})}_{u,n} - \Gamma_i, \overline{\mathsf{SINR}}^{(\text{p})}_{u,n} \geq \Gamma_i \right\} \tag{1}$$

$$\begin{aligned} \bar{C}^{\text{proj}}_{u,n} = \Bigg\{ & \frac{\bar{C}_{u,n}}{\sum_{u \in \mathcal{U}'_n} \bar{C}_{u,n}} R_j \mid j = \arg\min_{i \in \mathcal{J}} \left\{ \min_{u \in \mathcal{U}'_n} \overline{\mathsf{SINR}}^{(\text{c})}_{u,n} \right\} \\ & - \Gamma_i, \left\{ \min_{u \in \mathcal{U}'_n} \overline{\mathsf{SINR}}^{(\text{c})}_{u,n} \right\} \geq \Gamma_i \Bigg\} \end{aligned} \tag{2}$$

whereas R_j, Γ_j were introduced in Section II. In particular, $\overline{\mathsf{SINR}}^{(\text{p})}_{u,n}$ and $\overline{\mathsf{SINR}}^{(\text{c})}_{u,n}$ are respectively the highest discrete private and common SINRs that can be achieved by UE_u in $\mathcal{U}'_n$, which are mapped to their respective discrete rates $R^{\text{proj}}_{u,n}$ and $\bar{C}^{\text{proj}}_{u,n}$. Besides, $\bar{C}_{u,n}$, $\bar{\mathbf{w}}_{u,n}$, and $\bar{\mathbf{m}}_n$ are the common rate portion of UE_u, the private precoder of UE_u, and the common precoder of the n-th combination $\mathcal{U}'_n$, respectively. After evaluating all N combinations of admitted UEs, we pick the combination achieving the highest objective function value.

7) Computational Complexity: The computational complexities of solving $\mathcal{Q}_{\text{CWSR}}$ and $\mathcal{Q}_{\text{CWEE}}$ are similar, which is given by $\mathcal{C}_{\texttt{OPT-SCA-SDR}} = \mathcal{O}(N_q N_r N_c^{0.5} N_v^2 N_d)$, where $N_q = 2\binom{U}{K}$ is the total number of combinations of admitted UEs, N_r is the number of iterations needed for convergence, $N_c = 9K + 6$ is the total number of constraints, $N_v = 2KN_{\text{tx}} + 9N_{\text{tx}} + 2K + 11$ is the number of decision variables, and $N_d = N_{\text{tx}}^2 K^3 + N_{\text{tx}}^2 K^2 + 5KN_{\text{tx}}^2 + K^3 + 3N_{\text{tx}}^2 + K^2 + 2K + 1$ is the dimension of the SDP program.

V. Simulation Results

We evaluate the WSR and WEE for several configurations, varying the number of UEs, number of admitted UEs, and transmit powers. We consider two cases, namely, two-user settings

TABLE III
SIMULATION PARAMETERS

Scenario	Objective	$P_{\mathrm{tx}}^{\mathrm{max}}$ [dBm]	σ^2 [dBm]	N_{tx}	U	K	Δ_{SIC}	η_{eff}	P_{dyn} [dBm]	P_{sta} [dBm]	Weights	Channels
I	WSR	$40, 50$	30	4	2	2	0	–	–	–	Various [3]	Deterministic [3]
II	WSR	$40, 45, 50$	30	4	2	2	0	–	–	–	Various [3]	Deterministic [3]
III	WSR	$30, 40, 50$	30	4	2	2	$[0, 1]$	–	–	–	Various [3]	Deterministic [3]
IV	WEE	$30, 40$	30	4	2	2	0	0.35	33	38	Various [3]	Deterministic [3]
V	WSR	40	3GPP [50]	16	$2, \ldots, 6$	U	0	–	–	–	Uniform	3GPP [48]
VI	WSR	$10, \ldots, 40$	3GPP [50]	16	6	3	0	–	–	–	Uniform	3GPP [48]
VII	WEE	40	3GPP [50]	16	$2, \ldots, 6$	U	0	0.35	33	38	Uniform	3GPP [48]
VIII	WEE	$10, \ldots, 40$	3GPP [50]	16	6	3	0	0.35	33	38	Uniform	3GPP [48]

(**Scenario I** to **Scenario IV**) and multiuser settings (**Scenario V** to **Scenario VIII**). For the first set of scenarios, we adopt deterministic channels and do not include user admission to gain insight regarding the impact of discrete rates, which is done by modifying constraint $\bar{\mathrm{C}}_2$ as $\sum_{u\in\mathcal{U}} \chi_u \leq K$. In particular, we consider a system consisting of a BS with $N_{\mathrm{tx}} = 4$ antennas and $U = 2$ UEs with channels $\mathbf{h}_1 = [1, 1, 1, 1]^{\mathrm{H}}$, $\mathbf{h}_2 = [1, e^{j\phi}, e^{j2\phi}, e^{j3\phi}]^{\mathrm{H}}$, where $\phi = \{\frac{\pi}{9}, \frac{2\pi}{9}, \frac{3\pi}{9}, \frac{4\pi}{9}\}$ controls the similarity of the channels, whereas the noise power is set to $\sigma^2 = 30$ dBm, as in [3]. For the second set of scenarios, we adopt UMi line-of-sight (LOS)/non-LOS (NLOS) channels [48] with carrier frequency $f_{\mathrm{c}} = 41$ GHz, $N_{\mathrm{p}} = 4$ paths, bandwidth $\mathrm{BW} = 100$ MHz, noise figure $\mathrm{NF} = 5$ dB, and noise power $\sigma^2 = -174 + \mathrm{NF} + 10\log_{10}(\mathrm{BW/Hz})$ dBm. For this case, we consider two types of channels, i.e., correlated and uncorrelated, in order to assess the performance for different channel conditions. The uncorrelated and correlate channels model the cases when the UEs are distributed across the entire sector of 120° and within a narrower sector of 10°, respectively. Also, we consider $J = 15$ MCSs with target SINRs corresponding to 10% BLER [49], shown in Table II.

For the optimization of $\bar{\mathcal{Q}}_{\mathrm{CWSR}_n}^{(t)}$ and $\bar{\mathcal{Q}}_{\mathrm{CWEE}_n}^{(t)}$, we initialize the variables $\gamma_u, \rho_u, \tau_u, \lambda_u, \delta_u, \theta_u, \forall u \in \mathcal{U}$, as $\gamma_u^{(0)} = 1$, $\rho_u^{(0)} = 1$, $\tau_u^{(0)} = 1$, $\lambda_u^{(0)} = 1$, $\delta_u^{(0)} = 1$, $\theta_u^{(0)} = 1$. In addition, we initialize the penalty factor p_u as $p_u^{(0)} = 0.01, \forall u \in \mathcal{U}$, which is updated in each iteration t as $p_u^{(t+1)} = \min\{p_{\mathrm{inc}} \cdot p_u^{(t)}, p_{\max}\}$, where $p_{\mathrm{inc}} = 4$ and $p_{\max} = 1000$. As for the stopping criterion, we consider the threshold $\epsilon = 0.0001$ and the maximum number of iterations $N_{\mathrm{iter}} = 120$. The simulation results depict the average over $N_{\mathrm{ch}} = 100$ channel realizations assuming $R_{\min} = R_1$ (see Table II), unless specified otherwise. The maximum distance between the BS and UEs is $D_{\mathrm{BS}} = 60$ m. The formulated optimization problems are solved using CVX and MOSEK. The parameter settings employed in the considered scenarios are specified in Table III. Furthermore, we compare the following algorithms.[2]

- **OPT-MISOCP:** As proposed in Section III-B for discrete rates. By setting $\psi = 0$, it reduces to SDMA.
- **OPT-SCA-SDR:** As proposed in Section IV-B for continuous rates. By setting $\psi = 0$, it reduces to SDMA.
- **RND-MISOCP:** Variant of OPT-MISOCP, which assumes random user admission.
- **RND-SCA-SDR:** Variant of OPT-SCA-SDR, which assumes random user admission.
- **PR-OPT-SCA-SDR:** Obtained from OPT-SCA-SDR upon projecting the rates, as shown in (1) and (2).
- **PR-RND-SCA-SDR:** Obtained from RND-SCA-SDR upon projecting the rates, as shown in (1) and (2).

A. Complexity, Optimality, and Convergence

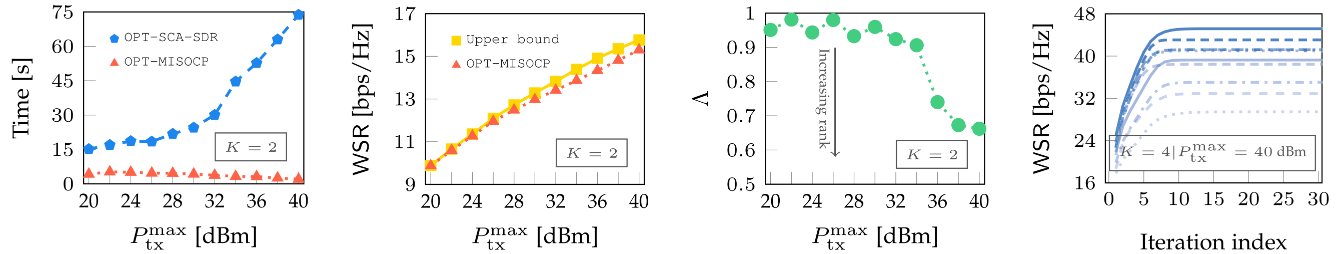
In this section, we quantify the runtime complexity of OPT-MISOCP and OPT-SCA-SDR, evaluate the optimality of OPT-MISOCP with respect to an upper bound, and analyze the convergence of OPT-SCA-SDR. For the results shown in Fig. 2, we consider the WSR problem with uncorrelated channels for $U = 4$, $K = \{2, 4\}$, weights $\omega_1 = \cdots = \omega_4 = 1$, and $N_{\mathrm{ch}} = 10$ channel realizations.

Runtime complexity: We compare the runtime complexity of OPT-MISOCP and OPT-SCA-SDR. In Fig. 2(a), we observe that for the considered parameters, OPT-MISOCP is $4 - 36$ times faster than OPT-SCA-SDR since the former exploits BnB, which circumvents the need of an exhaustive search. In contrast, OPT-SCA-SDR considers all possible combinations of admitted UEs. Furthermore, OPT-SCA-SDR is an iterative scheme, which needs to solve multiple instances of the problem until a stop criterion is met. We notice that OPT-SCA-SDR needs more time to converge as the transmit power increases. In particular, higher transit powers facilitate higher WSRs, and therefore more iterations are needed before the stopping criterion is satisfied. On the other hand, the runtime of OPT-MISOCP remains constant and even slightly decreases for higher transmit powers. This is due to constraint J_2, introduced in Section III-B7, which allows early stopping. Additional results on the worst-case complexity of OPT-MISOCP and OPT-SCA-SDR derived in Section III-B9 and Section IV-B7, respectively, are provided in Appendix L, available online.

Remark 6: We observed that for small numbers of UEs, e.g., $U = \{4, 5\}$, OPT-MISOCP has an affordable runtime. However, as U increases beyond these values, the runtime of OPT-MISOCP grows substantially, as more binary variables are involved. To keep OPT-MISOCP affordable, it can be combined with simple subcarrier allocation policy to avoid co-processing multiple UEs simultaneously and allowing for RRM parallelization.

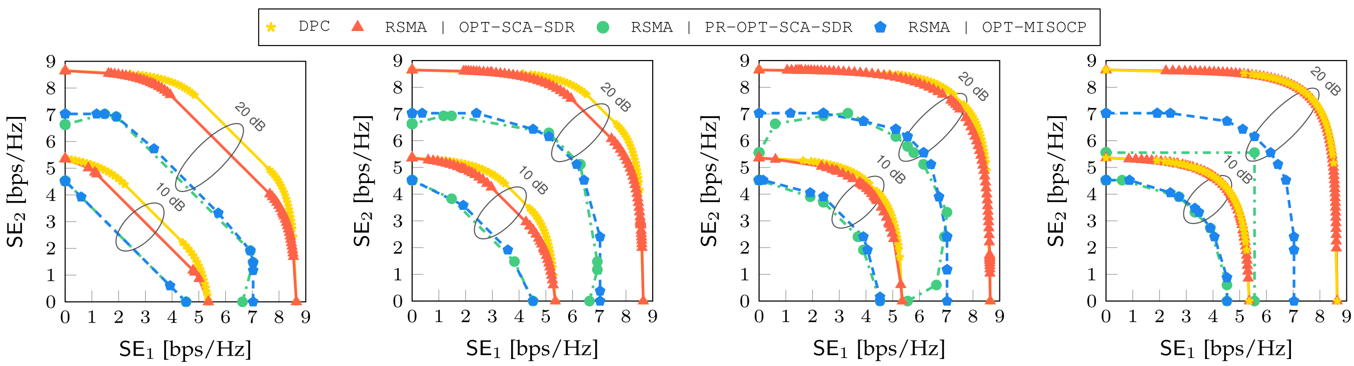
Optimality: We compare the WSR performance of OPT-MISOCP to an upper bound that we devise using SDR to demonstrate that OPT-MISOCP can yield near-optimal solutions for $\mathcal{P}'_{\mathrm{DWSR}}$. This upper bound is used to analyze the impact of the

[2] Our formulation allows us to obtain SDMA as a particular case of RSMA. However, NOMA cannot be obtained from it as NOMA requires multiple SIC stages and optimal decoding order, which our model does not feature. Still, to shed light on the performance of RSMA and NOMA, we have included results for a two-UE case in Appendix K, available online.



(a) Average runtime of OPT-SCA-SDR and OPT-MISOCP for $U = 4$, $K = 2$, considering uncorrelated channels.
(b) Average WSR of OPT-MISOCP and upper bound for $U = 4$, $K = 2$, considering uncorrelated channels.
(c) Average 'rank-oneness' of the upper bound for $U = 4$, $K = 2$, considering uncorrelated channels.
(d) Convergence of OPT-SCA-SDR for 10 channel realizations, $U = K = 4$, considering uncorrelated channels.

Fig. 2. Analysis of time complexity, optimality, and convergence.



(a) $\phi = \frac{\pi}{9}$
(b) $\phi = \frac{2\pi}{9}$
(c) $\phi = \frac{3\pi}{9}$
(d) $\phi = \frac{4\pi}{9}$

Fig. 3. (*Scenario I*) Two-user SE region of RSMA with discrete and continuous rates for $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = \{10, 20\}$ dB. Since OPT-SCA-SDR does not account for rate saturation, it continues upgrading the private rates, not necessarily leading to improved performance upon rate projection. In contrast, OPT-MISOCP considers that the rates are bounded and discrete, promoting more appropriate usage of power. Specifically, OPT-MISOCP uses the surplus of power to upgrade weaker private or common signals, preventing severe rate saturation of other signals.

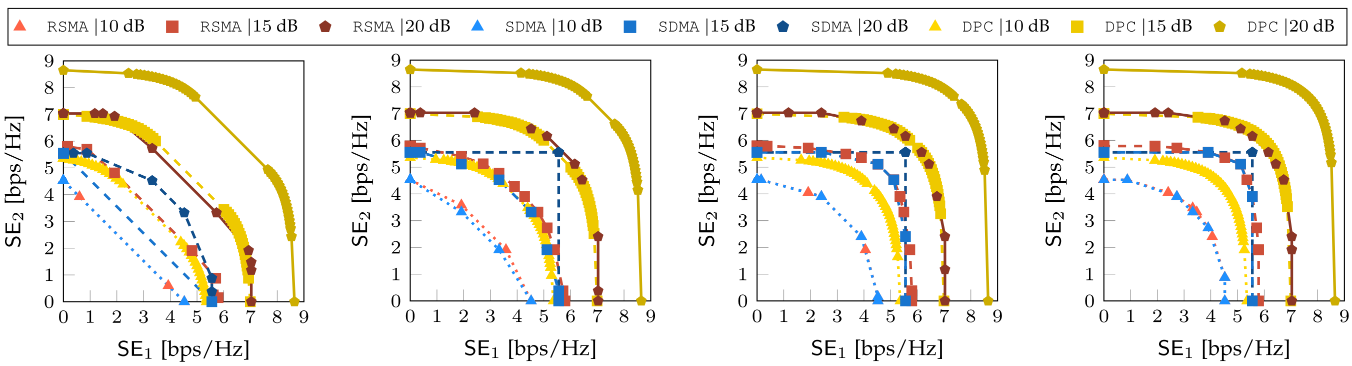
convexification procedure used in OPT-MISOCP. Since the upper bound has a larger feasible set due to the rank-one relaxation, it finds solutions that yield higher objective function values than OPT-MISOCP. However, such solutions are not necessarily feasible for problem $\mathcal{P}_{\text{DWSR}}$. In Fig. 2(b), we observe that the performance gap between OPT-MISOCP and the upper bound is generally small, although it slightly increases to 3% for higher transmit powers. To explain this result, we show in Fig. 2(c), the ratio of the principal eigenvalue to the sum of all eigenvalues, which we denote by Λ, i.e., Λ portrays the 'rank-oneness' of the upper bound solutions. Specifically, it is defined as $\Lambda = \frac{1}{\min\{1,\text{Rank}(\mathbf{X}_0)\}+\sum_{u\in\mathcal{U}}\min\{1,\text{Rank}(\mathbf{X}_u)\}} \frac{\sum_{u\in\mathcal{U}\cup\{0\}}\hat{\lambda}_{\max,u}}{\sum_m \hat{\lambda}_{m,u}}$, where $\hat{\lambda}_{m,u}$ is the m-th eigenvalue of $\mathbf{X}_u \succcurlyeq \mathbf{0}$, $u \in \mathcal{U} \cup \{0\}$. Here, $\mathbf{X}_u$ is the private precoder for UE_u and $\mathbf{X}_0$ is the precoder for the common signal, obtained by the upper bound. Λ reveals that the upper bound solutions have ranks higher than one, and therefore are not feasible for problem $\mathcal{P}_{\text{DWSR}}$, thus explaining the performance gap.

Convergence: In Fig. 2(d), we show the convergence of OPT-SCA-SDR for $N_{\text{ch}} = 10$ channel realizations.

Scenario I: Two-User SE Region for Continuous/Discrete RSMA Rates

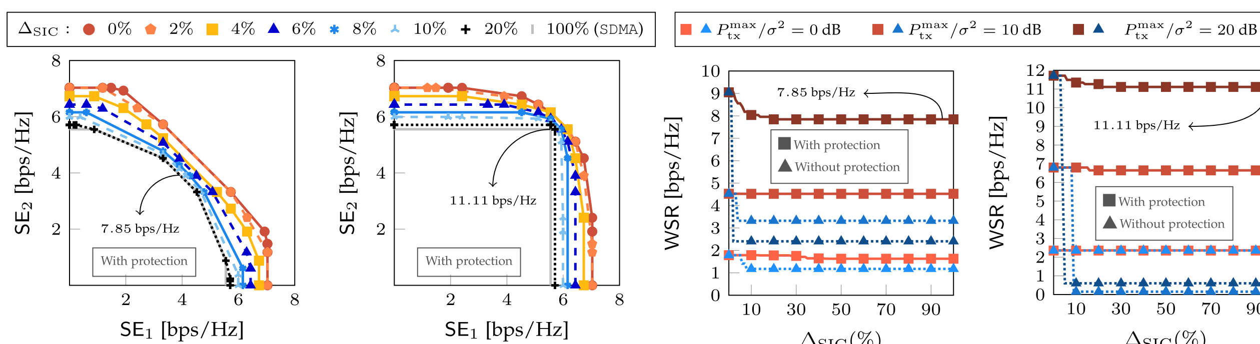
In Fig. 3, we compare the SE of RSMA with discrete and continuous rates to investigate the impact of rate discretization. For all considered cases, OPT-MISOCP and PR-OPT-SCA-SDR exhibit similar performance when $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = 10$ dB. This occurs because the rates obtained by OPT-SCA-SDR are small due to the low transmit power, and therefore projection does not have a significant impact. However, the performance gap between them can become large when $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = 20$ dB due to the higher rates achieved, which can lead to more noticeable projection losses. For instance, the difference is negligible in Fig. 3(a), whereas it is more evident in Fig. 3(c). The reason is that the channels become less correlated as ϕ increases, making the common rate less relevant for OPT-SCA-SDR. This causes OPT-SCA-SDR to be noticeably impacted by rate projection, as the private rates may experience heavy saturation while the common rate remains small. Fig. 3(d) shows an extreme case with low channel correlation, which causes OPT-SCA-SDR to opt for SDMA. In this case, the loss due to projection is higher than in Fig. 3(b) and (c) since the common rate is zero, and the private rates saturate at R_J (see Table II). On the other hand, OPT-MISOCP can prevent rate saturation losses as it takes the rate discretization into account, and expends any surplus of power to improve the common rate. For reference, we have included dirty paper coding (DPC), which is capacity-achieving for continuous rates [51]. We observe that the proposed OPT-SCA-SDR can approach the performance of DPC, especially in Fig. 3(d), thus demonstrating that OPT-SCA-SDR produces high-quality solutions.

In summary, the optimal partitioning of information, transmitted via unicast and multicast signals, can differ significantly



(a) $\phi = \frac{\pi}{9}$ (b) $\phi = \frac{2\pi}{9}$ (c) $\phi = \frac{3\pi}{9}$ (d) $\phi = \frac{4\pi}{9}$

Fig. 4. (*Scenario II*) Two-user SE region of RSMA and SDMA with discrete rates using OPT-MISOCP for $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = \{10, 15, 20\}$ dB. The advantage of RSMA stems from its capability of using the surplus of power to transmit the common signal, even in scenarios with highly uncorrelated channels, which SDMA is unable to do.



(a) $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = 20$ dB and $\phi = \frac{\pi}{9}$ (b) $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = 20$ dB and $\phi = \frac{4\pi}{9}$ (c) $\omega_1 = \omega_2 = 1$ and $\phi = \frac{\pi}{9}$ (d) $\omega_1 = \omega_2 = 1$ and $\phi = \frac{4\pi}{9}$

Fig. 5. (*Scenario III*) Two-user SE region of RSMA with discrete rates and imperfect SIC using OPT-MISOCP for $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = \{0, 10, 20\}$ dB and various Δ_{SIC} values. Accounting for potentially imperfect SIC has an enormous performance benefit. In the worst case, RSMA collapses to SDMA, still providing outstanding performance compared to the case without protection. In the presence of large unmanaged residuals of the common signal, due to an imperfect SIC, the private rates cannot be guaranteed, thus collapsing to zero due to the inability to fulfill the target SINRs required for successful decoding.

across different scenarios since the partitioning depends largely on the channel characteristics. The optimal partitioning also depends on the specific constraints of the RRM problem. Specifically, OPT-MISOCP is subject to discrete rates, while OPT-SCA-SDR assumes continuous rates. Although the resulting partitionings differ in these two cases, the respective values are optimal in each case, given the constraints.

Scenario II: Two-User SE Region with Discrete Rates for RSMA and SDMA

In Fig. 4, we compare the SE of RSMA and SDMA using discrete rates to elucidate the performance gap between them for different transmit powers. In Fig. 4(a), RSMA and SDMA have nearly the same performance when $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = 10$ dB, however, RSMA outperforms SDMA when $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = \{15, 20\}$ dB. SDMA is unable to cope well with high channel correlation, showing little improvement even as the transmit power increases. In contrast, RSMA can take advantage of high channel correlation to achieve considerable improvement. In Fig. 4(b), (c), and (d), RSMA outperforms SDMA by a small margin when $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = \{10, 15\}$ dB as the channels are less correlated, thus making the transmission of the common signal more expensive. However, RSMA clearly outperforms SDMA when $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = 20$ dB since SDMA saturates (i.e., UEs are served at rate $R_J = 5.5547$ bps/Hz), whereas RSMA can still improve as it can use the surplus of power to support a common signal.

Scenario III: Two-User SE Region with Imperfect SIC for RSMA

In Fig. 5, we evaluate the SE of RSMA for various levels of protection against imperfect SIC as well as without protection. When we consider protection, we assume a given $\Delta_{\text{SIC}} \neq 0\%$, which is taken into account for the optimization. Therefore, the BS guarantees the allocated rates for the UEs up to the selected value of Δ_{SIC}. When we neglect protection, we assume $\Delta_{\text{SIC}} = 0\%$ for the optimization even though the UEs may suffer from imperfect SIC. Therefore, the allocated rates may not be guaranteed. In Fig. 5(a) and (b), protection against imperfect SIC is considered. We observe that endowing RSMA with a higher robustness against imperfect SIC, i.e., larger Δ_{SIC}, produces a more noticeable decrease in the SE because the private SINRs are optimized to deal with additional

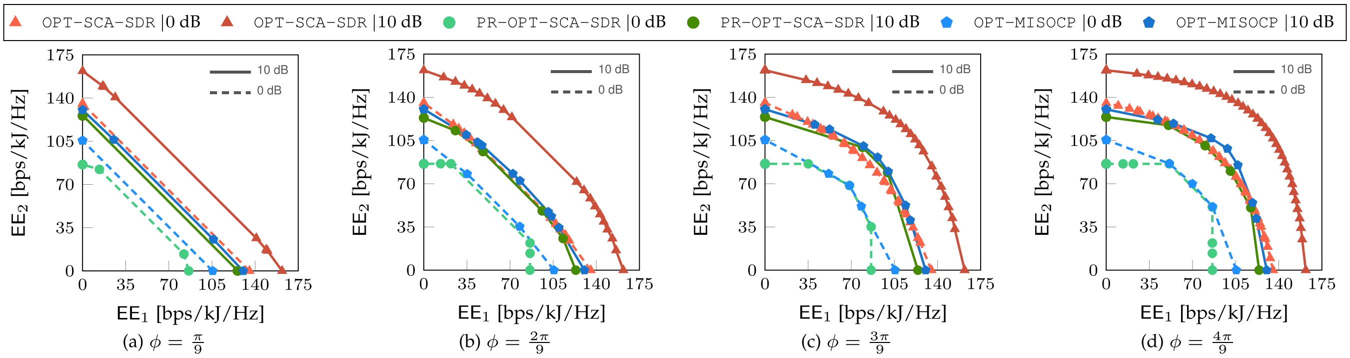


Fig. 6. (*Scenario IV*) Two-user EE region of RSMA with discrete and continuous rates for $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = \{0, 10\}$ dB, $\eta_{\text{eff}} = 0.35$, $P_{\text{dyn}} = 33$ dBm, and $P_{\text{sta}} = 38$ dBm. As it uses the transmit power more judiciously, OPT-MISOCP has a notable advantage over PR-OPT-SCA-SDR, ensuring high discrete rates with minimal power consumption, leading to improved EE. In contrast, PR-OPT-SCA-SDR is not aware of rate discretization, and therefore the precoders have larger powers than necessary, which impacts the EE upon rate projection.

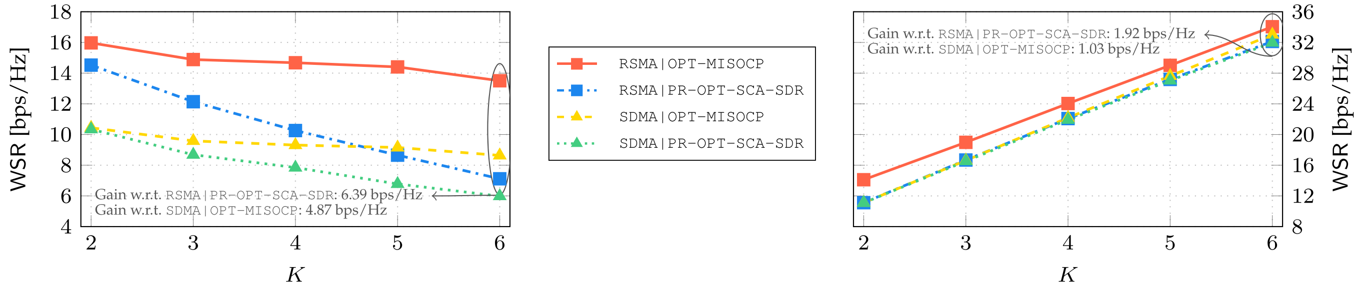
interference due to $\Delta_{\text{SIC}} \neq 0\%$ (see Section III-A3). Also, we observe that values up to $\Delta_{\text{SIC}} = 4\%$ do not affect the SE performance substantially while providing adequate protection. However, RSMA almost collapses to SDMA when $\Delta_{\text{SIC}} = 20\%$, as the common rates become very small. In fact, RSMA smartly switches to SDMA for values larger than $\Delta_{\text{SIC}} = 20\%$ since the high protection against imperfect SIC prevents enhancement of the private SINRs. The results for SDMA are identical to those for RSMA with $\Delta_{\text{SIC}} = 100\%$. In Fig. 5(c) and (d), we evaluate the impact of not accounting for imperfect SIC on the WSR performance, where we consider the same scenarios in Fig. 5(a) and (b), and equal weights, i.e., $\omega_1 = \omega_2 = 1$. In Fig. 5(c), the impact of imperfect SIC is small when $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = 0$ dB because information is predominantly transmitted via the common signal which is not affected by imperfect SIC. When $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = \{10, 20\}$ dB, the common and private rates increase since higher MCSs can be selected. This also implies that potential unmanaged residuals of the common signal may cause the private rates to collapse more noticeably, e.g., the SE drops from 7.85 bps/Hz to 2.41 bps/Hz (when $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = 20$ dB) and from 4.52 bps/Hz to 3.32 bps/Hz (when $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = 10$ dB). However, the system performs well when protection against imperfect SIC is considered. In particular, for high Δ_{SIC}, RSMA transitions to SDMA thereby avoiding further private SINRs degradation. In Fig. 5(d), we observe the same trend as in Fig. 5(c), although the degradation due to imperfect SIC is more conspicuous when protection against imperfect SIC is neglected. This occurs because the channels are highly uncorrelated, making the private rates even more prominent than in Fig. 5(c), with the consequent potential risk of much larger degradation in case of SIC failure.

Scenario IV: Two-User EE Region with Continuous/Discrete Rates for RSMA

In Fig. 6, we compare the EE of RSMA with continuous and discrete rates to investigate the impact of rate discretization. In Fig. 6(a), (b), (c), and (d), the EE of both OPT-MISOCP and PR-OPT-SCA-SDR improve when $P_{\text{tx}}^{\max}$ increases from 0 dB to 10 dB, as a higher transmit power allows to find an improved EE operating point with a better tradeoff between the achieved rates and the expended power. When $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = 0$ dB (dashed lines), OPT-MISOCP surpasses PR-OPT-SCA-SDR showing gains as large as 18 bps/kJ/Hz, particularly when the UE weights are not equal. When $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = 10$ dB (solid lines), OPT-MISOCP also outperforms PR-OPT-SCA-SDR although the gap is smaller. The reason for this effect is that OPT-MISOCP can better exploit the limited transmit power when $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = 0$ dB as it is able to handle discrete rates, whereas PR-OPT-SCA-SDR wastes power yielding rates higher than necessary, thus incurring a loss after projection. However, when $\frac{P_{\text{tx}}^{\max}}{\sigma^2} = 10$ dB, the power limitation is alleviated, and therefore PR-OPT-SCA-SDR can reduce the performance gap with respect to OPT-MISOCP. We oserve that as channels become less correlated, the EE of OPT-MISOCP and PR-OPT-SCA-SDR improve because interference can be handled more effectively and with less transmit power.

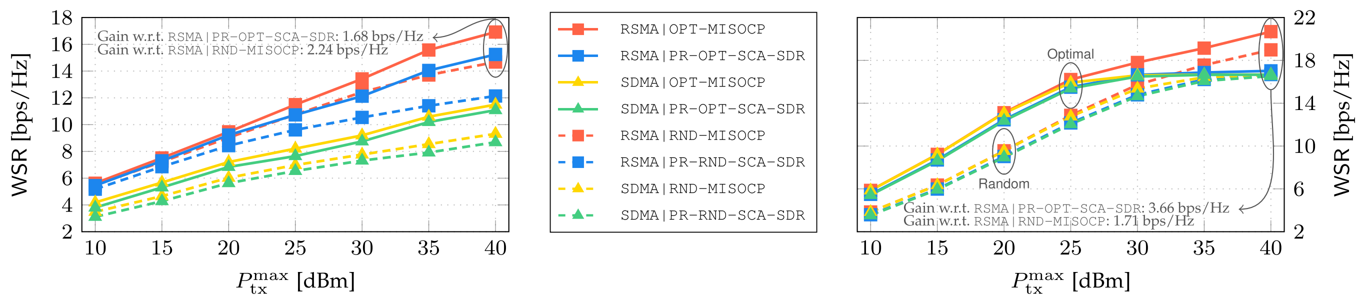
Scenario V: Impact of the Number of Admitted UEs on WSR Performance

In Fig. 7, we compare the WSR of RSMA and SDMA when the number admitted UEs varies. In Fig. 7(a), we consider correlated channels, for which we observe that an increasing number of UEs leads to WSR degradation. This occurs because the UEs are located in close proximity of each other, exacerbating interference for every additional UE admitted. We observe that RSMA | OPT-MISOCP has a noticeable advantage over SDMA | OPT-MISOCP since it can exploit the channel similarity via the common signal. We observe a similar behavior for RSMA | PR-OPT-SCA-SDR and SDMA | PR-OPT-SCA-SDR although the difference between them decreases as U increases. Also, not considering rate discretization can severely affect RSMA | PR-OPT-SCA-SDR, reducing its performance to the extent of being outperformed by SDMA | OPT-MISOCP when $U = \{5, 6\}$. In Fig. 7(b), we consider uncorrelated channels, for which we observe that increasing the number of UEs leads to an improved WSR. This is expected as interference is more easily dealt with in this case. Also, RSMA | PR-OPT-SCA-SDR and SDMA | PR-OPT-SCA-SDR achieve the same performance because RSMA does not devise a common signal. However, RSMA | OPT-MISOCP surpasses SDMA | OPT-MISOCP as it is able to exploit the surplus of power to devise the common signal.



(a) Correlated channels. (b) Uncorrelated channels.

Fig. 7. (*Scenario V*) WSR of RSMA and SDMA as a function of the number of admitted UEs. In Fig. 7(a), RSMA | OPT-MISOCP has an advantage of 6.39 bps/Hz (↑ 89.7% gain) and 4.87 bps/Hz (↑ 56.3% gain) with respect to RSMA | PR-OPT-SCA-SDR and SDMA | OPT-MISOCP, respectively, when $U = 6$. In Fig. 7(b), RSMA | OPT-MISOCP has an advantage of 1.92 bps/Hz (↑ 5.9% gain) and 1.03 bps/Hz (↑ 3.1% gain) compared to RSMA | PR-OPT-SCA-SDR and SDMA | OPT-MISOCP, respectively, when $K = 6$.



(a) Correlated channels. (b) Uncorrelated channels.

Fig. 8. (*Scenario VI*) WSR of RSMA and SDMA with optimal and random UE admission as a function of the transmit power. In Fig. 8(a), RSMA | OPT-MISOCP has an advantage of 1.68 bps/Hz (↑ 10.9% gain) and 2.24 bps/Hz (↑ 15.3% gain) with respect to RSMA | PR-OPT-SCA-SDR and RSMA | RND-MISOCP, respectively, when $P_{\rm tx}^{\max} = 40$ dBm. In Fig. 8(b), RSMA | OPT-MISOCP has an advantage of 3.66 bps/Hz (↑ 21.6% gain) with respect to RSMA | PR-OPT-SCA-SDR and 1.71 bps/Hz (↑ 8.9% gain) with respect to RSMA | RND-MISOCP, when $P_{\rm tx}^{\max} = 40$ dBm.

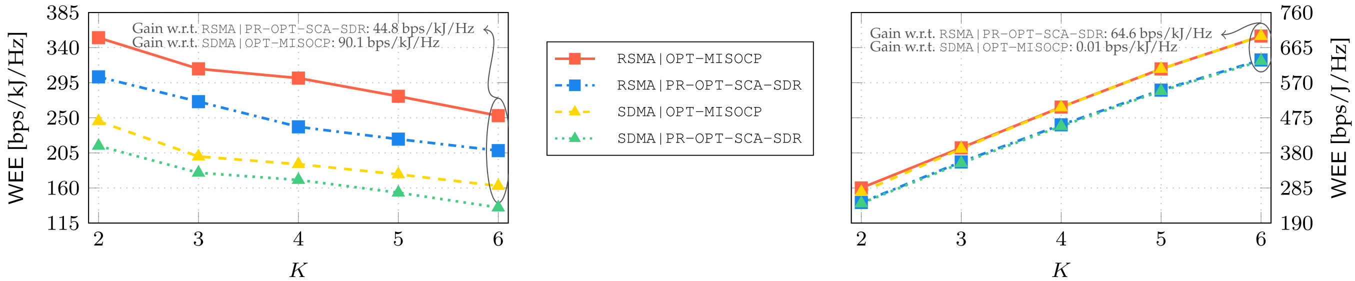
Besides, SDMA | OPT-MISOCP performs slightly better than RSMA | PR-OPT-SCA-SDR as it avoids projection losses.

Scenario VI: Impact of the Transmit Power on the WSR Performance

In Fig. 8, we evaluate the WSR as a function of the transmit power. In Fig. 8(a), we consider correlated channels, for which optimal admission leads to a consistently higher WSR compared to random admission of UEs. Besides, RSMA outperforms SDMA in all cases due to the high channel similarity, which allows capitalizing on the multicast signal. Specifically, the performance gap widens as the transmit power increases since higher rates can be allocated to the common signal, whereas SDMA is hampered by high interference. We also observe that RSMA | OPT-MISOCP outperforms RSMA | PR-OPT-SCA-SDR for all considered cases, whereas RSMA | RND-MISOCP performs similarly to RSMA | PR-OPT-SCA-SDR even though RSMA | RND-MISOCP does not control which UEs are admitted. In Fig. 8(b), we consider uncorrelated channels, where optimal admission also facilitates additional gains for both RSMA and SDMA compared to random UE admission, particularly when the transmit power is more constrained. Besides, SDMA | OPT-MISOCP performs marginally better than RSMA | PR-OPT-SCA-SDR because the latter collapses to SDMA due to the low channel correlation, thereby experiencing severe saturation upon rate projection. On the other hand, the gains due to optimal UE admission tend to diminish for higher transmit powers. For high transmit powers, RSMA | RND-MISOCP surpasses RSMA | PR-OPT-SCA-SDR as the former accounts for rate discretization, thus avoiding losses due to rate projection.

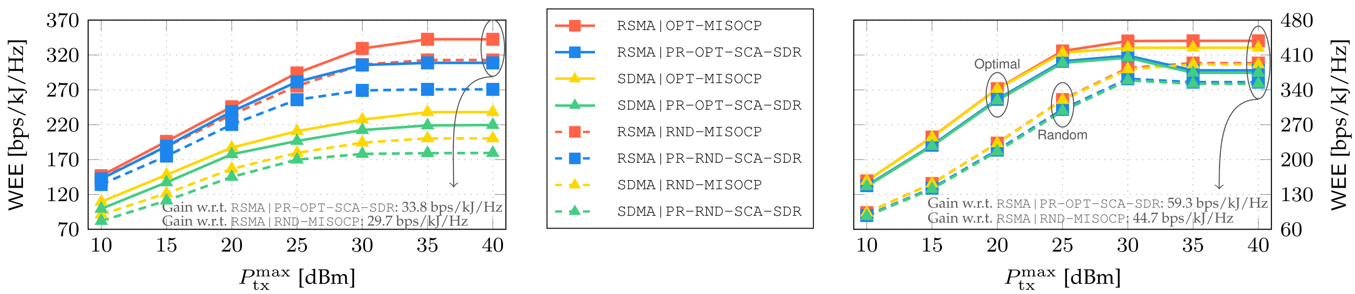
Scenario VII: Impact of the Number of Admitted UEs on WEE Performance

In Fig. 9, we compare the WEE of RSMA and SDMA as a function of the number of admitted UEs. In Fig. 9(a), we consider correlated channels, for which RSMA and SDMA experience a WEE degradation as the number of admitted UEs increases. This occurs because the transmit power needs to be distributed among more UEs, thus affecting the SINRs and the allocated rates. However, RSMA attains a higher performance than SDMA since RSMA is capable of harnessing the high channel similarity. Further, we observe that RSMA | OPT-MISOCP and SDMA | OPT-MISOCP respectively outperform RSMA | PR-OPT-SCA-SDR and SDMA | PR-OPT-SCA-SDR by at least 20%. In Fig. 9(b), we consider uncorrelated channels, for which RSMA collapses to SDMA in most cases, since the common rate is very small or zero due to a low channel correlation. Furthermore, the common rate improves marginally when the number of UEs increases, as it requires a substantially larger transmit power. Specifically,



(a) Correlated channels. (b) Uncorrelated channels.

Fig. 9. (*Scenario VII*) WEE of RSMA and SDMA as a function of the number of admitted UEs. In Fig. 9(a), RSMA | OPT-MISOCP outperforms RSMA | OPT-PR-SCA-SDR and SDMA | OPT-MISOCP by 44.8 bps/kJ/Hz ($\uparrow$ 21.5% gain) and 90.1 bps/kJ/Hz ($\uparrow$ 55.4% gain), respectively, when $K = 6$. In Fig. 9(b), RSMA | OPT-MISOCP outperforms RSMA | OPT-PR-SCA-SDR by 64.6 bps/kJ/Hz ($\uparrow$ 10.2% gain), when $K = 6$.



(a) Correlated channels. (b) Uncorrelated channels.

Fig. 10. (*Scenario VIII*) WEE of RSMA and SDMA with optimal and random UE admission as a function of the transmit power. In Fig. 10(a), RSMA | OPT-MISOCP outperforms RSMA | OPT-PR-SCA-SDR and RSMA | RND-MISOCP by 33.8 bps/kJ/Hz ($\uparrow$ 10.9% gain) and 29.7 bps/kJ/Hz ($\uparrow$ 9.5% gain), respectively, when $P_{tx}^{max} = 40$ dBm. In Fig. 10(b), RSMA | OPT-MISOCP outperforms RSMA | OPT-PR-SCA-SDR and RSMA | RND-MISOCP by 59.3 bps/kJ/Hz ($\uparrow$ 15.6% gain) and 44.7 bps/kJ/Hz ($\uparrow$ 11.4% gain), respectively, when $P_{tx}^{max} = 40$ dBm.

as the number of UEs increases, utilizing the common signal becomes less energy-efficient. We observe that RSMA | OPT-MISOCP outperforms RSMA | PR-OPT-SCA-SDR for all considered values of K.

Scenario VIII: Impact of the Transmit Power on the WEE Performance

In Fig. 10, we evaluate the WEE as a function of the transmit power. In Fig. 10(a), we consider correlated channels, for which RSMA outperforms SDMA as it can exploit the high channel correlation. We also observe that optimal admission performs significantly better than random UE admission, as it allows to select UEs with mutually beneficial channel characteristics that promote EE gains. In addition, increasing the transmit power boosts the WEE as improved operating points can be found. However, this increment saturates after a certain point, as the power required to reach higher rates becomes too costly for a marginal gain in WEE. Besides, RSMA | RND-MISOCP performs similarly to RSMA | PR-OPT-SCA-SDR since its ability to handle discrete rates compensates indirectly for the random selection of UEs. In Fig. 10(b), we consider uncorrelated channels, for which we observe that optimal UE admission can lead to substantial gains. Also, RSMA | OPT-MISOCP and SDMA | OPT-MISOCP outperform RSMA | OPT-PR-SCA-SDR and RSMA | RND-PR-SCA-SDR, respectively. Moreover, RSMA | OPT-PR-SCA-SDR and SDMA | OPT-PR-SCA-SDR experience a WEE degradation for larger values of the transmit power because of rate saturation.

VI. Future Research Directions

In the following, we discuss future avenues of research that can be pursued to extend our current work.

Practical RRM algorithms: With an increasing pool of radio resources, the execution times of RRM algorithms are expected to rise. Thus, having fast, efficient, real-time algorithms capable of operating within short time frames is crucial for practical deployments. This is especially important for reducing end-to-end communication latency and optimizing radio resource utilization efficiency. Achieving fast RRM is plausible through the use of heuristics, which can simplify RRM's complexity by, for example, pre-selecting a subset of admitted UEs to minimize the number of binary variables. However, choosing such a subset of UEs is not trivial in RSMA. Greedy strategies based, for instance, on channel correlation, widely used in SDMA and NOMA, may not be effective in RSMA. Unlike SDMA, which favors low channel correlation, and NOMA, which favors high channel correlation, RSMA presents unique challenges. In particular, the multicast signal of RSMA benefits from correlated channels, whereas the unicast signals benefit from uncorrelated channels. Hence, heuristics for determining a subset of admitted UEs that ensures high performance of RSMA are yet to be developed and require further research. Considering that a significant portion of the computational complexity is linked to the binary variables needed for UE admission and discrete rate selection, one strategy to cope with it is to relax the binary variables and penalize their integrality violation in the objective function, as in [52]. However, fine-tuning the penalty factors associated with these binary variables may become tedious, with no guarantee of obtaining integer solutions in all cases. Overall, further research is needed to develop real-time and efficient algorithms for the RRM of RSMA, particularly when an extended pool of radio resources is considered, which is expected to increase computational complexity. Additionally, integrating mobile data traffic patterns into RRM strategies, as discussed in [53], may offer further opportunities for optimization.

Profiting from caching: Leveraging the caching capabilities of UEs holds great promise for further improving RSMA performance, as demonstrated by its efficacy in improving the SE of NOMA [54]. Specifically, RSMA can benefit from exploiting content cached at UEs in several ways. Cached content can significantly reduce latency by allowing the BS to transmit only the missing fragments of the requested content. Moreover, cached content can bolster signal quality by aiding in interference cancellation, thus increasing SINR and data rates. However, integrating caching into the RRM design of RSMA poses challenges, including the cache capacity, the size of the requested content, and the specific content cached. In addition, determining which content fragments to cache is non-trivial due to limited cache capacities. Also, exchanging information between the BS and UEs regarding the cached content incurs additional overhead. Furthermore, accounting for cache power consumption is critical, as maintaining cache freshness demands energy. Despite these challenges, optimizing caching jointly with other radio resources holds significant potential for enhancing RSMA's performance.

Integration with OFDM: Given the widespread adoption of orthogonal frequency-division multiplexing (OFDM) in modern wireless communications systems, recent research has focused on integrating OFDM with next-generation multiple-access candidates, such as RSMA. Notably, OFDM-RSMA integration has demonstrated significant advantages in mitigating inter-carrier interference (ICI) [55] caused by Doppler spread and inter-numerology interference (INI) [56] caused by different subcarrier spacing (SS) numerologies. These findings have opened new avenues for future research. In particular, the integration of beamforming design in OFDM-RSMA systems remains to be thoroughly investigated, as most previous works focused on SISO scenarios. It is also interesting to explore flexible SS numerology selection, an aspect overlooked in prior works, which only assessed the impact of different numerologies on performance. Furthermore, with discussions on more advanced OFDM variants like orthogonal time-frequency space (OTFS) modulation underway, it is timely to investigate the combination of RSMA with these emerging waveforms, which promise to support ultra-high mobility.

CSI estimation: As RSMA utilizes the same CSI as SDMA for RRM, acquiring and estimating CSI in RSMA systems is expected to remain consistent with the established standardized methods. Specifically, standardized CSI estimation procedures are already in place for unicast services [57]. Additionally, 3GPP has implemented CSI estimation procedures tailored for multicast-broadcast services, designed primarily to transmit multimedia content to multiple UEs efficiently [58]. As RSMA is based on unicast and multicast transmissions, it can seamlessly leverage existing CSI estimation procedures without any changes. However, future RSMA systems may benefit from the joint design of, e.g., reference signals for CSI estimation. This joint design strategy could reduce the overhead compared to employing individual and independent reference signals for unicast and multicast transmissions, opening up novel research directions on CSI estimation for RSMA systems.

VII. Conclusion

In this paper, two new RRM problems were proposed to investigate the SE and EE of RSMA, taking into account characteristics of practical wireless systems, namely the use of discrete rates in contrast to the widely embraced continuous rates, the need for selective UE admission instead of ubiquitously serving all UEs, and imperfect SIC in lieu of ideal SIC. In particular, we investigated the maximization of the WSR and WEE of RSMA as optimization problems and jointly optimized the beamforming, the UE admission, and the allocation of discrete rates, while accounting for an imperfect SIC. Furthermore, given the widespread adoption of Shannon's capacity formula for SINR-rate modeling in RRM designs, we also considered the case of continuous rates. The considered RRM problems resulted in nonconvex MINLPs, which are generally difficult to solve. Nevertheless, we developed two algorithms capable of finding high-quality solutions. The first algorithm addresses the RRM with discrete rates and transforms the nonconvex MINLP into a MISOCP, which can be solved globally optimally via BnB and IPMs. This algorithm features custom cutting planes that reduce the runtime. The second algorithm addresses the RRM with continuous rates, and solves the nonconvex MINLP using binary enumeration, SDR, and SCA, converging to a KKT point. We revealed that ignoring the practical characteristics of wireless systems in RRM design can have serious repercussions on performance. Specifically, we demonstrated the importance of accounting for discrete rates in the RRM model to avoid potentially severe rate projection losses. In addition, we recognized the importance of selectivity for UE admission, which yields greater gains, as it allows to serve UEs with mutually beneficial channel characteristics that can improve the WSR or WEE. Finally, our results confirmed the benefits of accounting for imperfect SIC to guarantee the allocated rate. Our simulations show that RSMA designed for discrete rates achieves gains of up to 89.7% (WSR) and 21.5% (WEE) compared to projecting continuous rates onto the admissible set of discrete rates since projection losses are avoided. Furthermore, user admission proves crucial for RSMA as it yields additional gains of up to 15.3% (WSR) and 11.4% (WEE) compared to random user admission when discrete rates are considered.

References

[1] B. Clerckx et al., "A primer on rate-splitting multiple access: Tutorial, myths, and frequently asked questions," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 5, pp. 1265–1308, May 2023.

[2] B. Clerckx et al., “Is NOMA efficient in multi-antenna networks? A critical look at next generation multiple access techniques,” *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1310–1343, 2021.

[3] Y. Mao, B. Clerckx, and V. O. K. Li, “Rate-splitting multiple access for downlink communication systems: Bridging, generalizing, and outperforming SDMA and NOMA,” *EURASIP J. Wirel. Commun. Netw.*, vol. 1, May 2018, Art. no. 133.

[4] 3GPP, “Technical specification group radio access network; study on non-orthogonal multiple access (NOMA) for NR,” 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38. 812, 12 2017, version 16.0.0. [Online]. Available: https://www.3gpp.org/ftp//Specs/archive/38_series/38.812/38812-g00.zip

[5] S. Shi, M. Schubert, and H. Boche, “Rate optimization for multiuser MIMO systems with linear processing,” *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 4020–4030, Aug. 2008.

[6] Y. Mao, O. Dizdar, B. Clerckx, R. Schober, P. Popovski, and H. V. Poor, “Rate-splitting multiple access: Fundamentals, survey, and future research trends,” *IEEE Commun. Surv. Tut.*, vol. 24, no. 4, pp. 2073–2126, Fourth Quarter 2022.

[7] Y. Liu et al., “Evolution of NOMA toward next generation multiple access (NGMA) for 6G,” *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1037–1071, Apr. 2022.

[8] R.-J. Reifert, S. Roth, A. A. Ahmad, and A. Sezgin, “Energy efficiency in rate-splitting multiple access with mixed criticality,” in *Proc. IEEE ICC Workshops*, 2022, pp. 681–686.

[9] C. Xu, B. Clerckx, S. Chen, Y. Mao, and J. Zhang, “Rate-splitting multiple access for multi-antenna joint radar and communications,” *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1332–1347, Nov. 2021.

[10] S. Naser, L. Bariah, S. Muhaidat, M. Al-Qutayri, M. Uysal, and P. C. Sofotasios, “Interference management strategies for multiuser multicell MIMO VLC systems,” *IEEE Trans. Commun.*, vol. 70, no. 9, pp. 6002–6019, Sep. 2022.

[11] L. F. Abanto-Leon, M. Hollick, B. Clerckx, and G. H. Sim, “Sequential parametric optimization for rate-splitting precoding in non-orthogonal unicast and multicast transmissions,” in *Proc. IEEE Int. Conf. Commun.*, 2022, pp. 1–7.

[12] Y. Mao, B. Clerckx, and V. O. K. Li, “Rate-splitting for multi-antenna non-orthogonal unicast and multicast transmission,” in *Proc. IEEE 19th Int. Workshop Signal Process. Adva. Wireless Commun.*, 2018, pp. 1–5.

[13] H. Fu, S. Feng, W. Tang, Z. Wei, and D. W. K. Ng, “Secrecy outage-constrained robust resource allocation design for MU-MISO RSMA systems,” in *Proc. IEEE GLOBECOM Workshops*, 2021, pp. 1–7.

[14] H. Pang, F. Ji, L. Xu, Y. Liu, and M. Wen, “Resource allocation for RSMA-based coordinated direct and relay transmission,” *IEEE Wireless Commun. Lett.*, vol. 12, no. 3, pp. 505–509, Mar. 2023.

[15] H. Fu, S. Feng, and D. W. K. Ng, “Resource allocation design for IRS-aided downlink MU-MISO RSMA systems,” in *Proc. IEEE Int. Conf. Commun. Workshops*, 2021, pp. 1–6.

[16] A. Rahmati, Y. Yapici, N. Rupasinghe, I. Guvenc, H. Dai, and A. Bhuyan, “Energy efficiency of RSMA and NOMA in cellular-connected mmwave UAV networks,” in *Proc. IEEE Int. Conf. Commun. Workshops*, 2019, pp. 1–6.

[17] F. Xing, S. He, V. C. M. Leung, and H. Yin, “Energy efficiency optimization for rate-splitting multiple access-based indoor visible light communication networks,” *IEEE J. Sel. Areas Commun.*, vol. 40, no. 5, pp. 1706–1720, May 2022.

[18] Z. Yang, M. Chen, Z. Zhang, and C. Huang, “Energy efficient semantic communication over wireless networks with rate splitting,” *IEEE J. Sel. Areas Commun.*, vol. 41, no. 5, pp. 1484–1495, May 2023.

[19] O. Dizdar, A. Kaushik, B. Clerckx, and C. Masouros, “Energy efficient dual-functional radar-communication: Rate-splitting multiple access, low-resolution DACs, and RF chain selection,” *IEEE Open J. Commun. Soc.*, vol. 3, pp. 986–1006, 2022.

[20] Y. Mao, B. Clerckx, and V. O. K. Li, “Rate-splitting for multi-antenna non-orthogonal unicast and multicast transmission: Spectral and energy efficiency analysis,” *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8754–8770, Dec. 2019.

[21] B. Matthiesen, Y. Mao, A. Dekorsy, P. Popovski, and B. Clerckx, “Globally optimal spectrum- and energy-efficient beamforming for rate splitting multiple access,” *IEEE Trans. Signal Process.*, vol. 70, pp. 5025–5040, 2022.

[22] G. Zhou, Y. Mao, and B. Clerckx, “Rate-splitting multiple access for multi-antenna downlink communication systems: Spectral and energy efficiency tradeoff,” *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 4816–4828, Jul. 2022.

[23] N. Q. Hieu, D. T. Hoang, D. Niyato, and D. I. Kim, “Optimal power allocation for rate splitting communications with deep reinforcement learning,” *IEEE Wireless Commun. Lett.*, vol. 10, no. 12, pp. 2820–2823, Dec. 2021.

[24] Z. Yang, M. Chen, W. Saad, and M. Shikh-Bahaei, “Optimization of rate allocation and power control for rate splitting multiple access (RSMA),” *IEEE Trans. Commun.*, vol. 69, no. 9, pp. 5988–6002, Sep. 2021.

[25] A. R. Flores and R. C. de Lamare, “Robust and adaptive power allocation techniques for rate splitting based MU-MIMO systems,” *IEEE Trans. Commun.*, vol. 70, no. 7, pp. 4656–4670, Jul. 2022.

[26] T. Cai, J. Zhang, S. Yan, L. Meng, J. Sun, and N. Al-Dhahir, “Resource allocation for secure rate-splitting multiple access with adaptive beamforming,” in *Proc. IEEE Int. Conf. Commun. Workshops*, 2021, pp. 1–6.

[27] N. Q. Hieu, D. T. Hoang, D. Niyato, D. N. Nguyen, D. I. Kim, and A. Jamalipour, “Joint power allocation and rate control for rate splitting multiple access networks with covert communications,” *IEEE Trans. Commun.*, vol. 71, no. 4, pp. 2274–2287, Apr. 2023.

[28] 3GPP, “5G; NR; Physical layer procedures for data,” 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.214, 2020, version 16.2.0. [Online]. Available: https://www.3gpp.org/ftp//Specs/archive/38_series/38.214/38214-g20.zip

[29] G. H. Sim and J. Widmer, “Finite horizon opportunistic multicast beamforming,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1452–1465, Mar. 2017.

[30] O. Dizdar, Y. Mao, W. Han, and B. Clerckx, “Rate-splitting multiple access for downlink multi-antenna communications: Physical layer design and link-level simulations,” in *Proc. IEEE 31st Annu. Int. Symp. Pers. Indoor Mobile Radio Commun.*, 2020, pp. 1–6.

[31] Y. Cheng, A. Philipp, and M. Pesavento, “Dynamic rate adaptation and multiuser downlink beamforming using mixed integer conic programming,” in *Proc. 20th Eur. Signal Process. Conf.*, 2012, pp. 824–828.

[32] Y. Cheng and M. Pesavento, “Joint discrete rate adaptation and downlink beamforming using mixed integer conic programming,” *IEEE Trans. Signal Process.*, vol. 63, no. 7, pp. 1750–1764, Apr. 2015.

[33] H.-T. Wai, Q. Li, and W.-K. Ma, “Discrete sum rate maximization for MISO interference broadcast channels: Convex approximations and efficient algorithms,” *IEEE Trans. Signal Process.*, vol. 64, no. 16, pp. 4323–4336, Aug. 2016.

[34] S. X.-Y. Ni and A. M.-C. So, “Mixed-integer semidefinite relaxation of joint admission control and beamforming: An SOC-based outer approximation approach with provable guarantees,” in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun.*, 2018, pp. 1–5.

[35] A. Bandi, M. R. B. Shankar, S. Chatzinotas, and B. Ottersten, “A joint solution for scheduling and precoding in multiuser MISO downlink channels,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 475–490, Jan. 2020.

[36] E. Matskani, N. D. Sidiropoulos, Z.-Q. Luo, and L. Tassiulas, “Convex approximation techniques for joint multiuser downlink beamforming and admission control,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2682–2693, Jul. 2008.

[37] R. Wang, W. Kang, G. Liu, R. Ma, and B. Li, “Admission control and power allocation for NOMA-based satellite multi-beam network,” *IEEE Access*, vol. 8, pp. 33631–33643, 2020.

[38] L. F. Abanto-Leon, A. Asadi, A. Garcia-Saavedra, G. H. Sim, and M. Hollick, “Radiorchestra: Proactive management of millimeter-wave self-backhauled small cells via joint optimization of beamforming, user association, rate selection, and admission control,” *IEEE Trans. Wireless Commun.*, vol. 22, no. 1, pp. 153–173, Jan. 2023.

[39] X. Ou, X. Xie, H. Lu, and H. Yang, “Resource allocation in MU-MISO rate-splitting multiple access with SIC errors for URLLC services,” *IEEE Trans. Commun.*, vol. 71, no. 1, pp. 229–243, Jan. 2023.

[40] G. Chopra, A. Chowdary, and A. Kumar, “Bounds on power and common message fractions for RSMA with imperfect SIC,” 2022. [Online]. Available: https://arxiv.org/abs/2203.02748

[41] N. Mouni, A. Kumar, and P. K. Upadhyay, “Adaptive user pairing for NOMA systems with imperfect SIC,” *IEEE Wireless Commun. Lett.*, vol. 10, no. 7, pp. 1547–1551, Jul. 2021.

[42] I. A. Mahady, E. Bedeer, S. Ikki, and H. Yanikomeroglu, “Sum-rate maximization of NOMA systems under imperfect successive interference cancellation,” *IEEE Commun. Lett.*, vol. 23, no. 3, pp. 474–477, Mar. 2019.

[43] X. Wang, R. Chen, Y. Xu, and Q. Meng, “Low-complexity power allocation in NOMA systems with imperfect SIC for maximizing weighted sum-rate,” *IEEE Access*, vol. 7, pp. 94238–94253, 2019.

[44] H. Wang, Z. Zhang, and X. Chen, “Energy-efficient power allocation for non-orthogonal multiple access with imperfect successive interference cancellation,” in *Proc. 9th Int. Conf. Wireless Commun. Signal Process.*, 2017, pp. 1–6.

[45] K. K. Leung and W. Li-Chun, “Integrated link adaptation and power control to improve error and throughput performance in broadband wireless packet networks,” *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 619–629, Oct. 2002.

[46] A. Beck, A. Ben-Tal, and L. Tetruashvili, “A sequential parametric convex approximation method with applications to nonconvex truss topology design problems,” *J. Glob. Optim.*, vol. 47, no. 1, pp. 29–51, 2010.

[47] C. Sun and R. Dai, “An iterative rank penalty method for nonconvex quadratically constrained quadratic programs,” *SIAM J. Control Optim.*, vol. 57, no. 6, pp. 3749–3766, 2019.

[48] 3GPP, “Study on scenarios and requirements for next generation access technologies,” 3rd Generation Partnership Project (3GPP), Technical Report (TR) 38.913, 05 2017, version 14.2.0. [Online]. Available: https://www.3gpp.org/ftp//Specs/archive/38_series/38.913/38913-e20.zip

[49] R. Kovalchukov, D. Moltchanov, Y. Gaidamaka, and E. Bobrikova, “An accurate approximation of resource request distributions in millimeter wave 3GPP new radio systems,,” in *Internet of Things, Smart Spaces, and Next Generation Networks and Systems*, S. Galinina, S. A. Balandin, and Y. Koucheryavy Eds., Cham, Switzerland: Springer, 2019, pp. 572–585.

[50] 3GPP, “Technical specification group radio access network; V2X services based on NR; user equipment (UE) radio transmission and reception,” 3rd Generation Partnership Project (3GPP), Technical Report (TR) 38.886, 2022, version 16.3.0. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/38_series/38.886/38886-g30.zip

[51] H. Viswanathan, S. Venkatesan, and H. Huang, “Downlink capacity evaluation of cellular networks with known-interference cancellation,” *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 802–811, Jun. 2003.

[52] A. Khalili, M. R. Mili, M. Rasti, S. Parsaeefard, and D. W. K. Ng, “Antenna selection strategy for energy efficiency maximization in uplink OFDMA networks: A multi-objective approach,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 595–609, Jan. 2020.

[53] P. Fernandez-Perez, C. Fiandrino, and J. Widmer, “Characterizing and modeling mobile networks user traffic at millisecond level,” in *Proc. 17th ACM Workshop Wireless Netw. Testbeds Exp. Eval. Characterization*, 2023, pp. 64–71.

[54] L. Xiang, D. W. K. Ng, X. Ge, Z. Ding, V. W. S. Wong, and R. Schober, “Cache-aided non-orthogonal multiple access: The two-user case,” *IEEE J. Sel. Areas Commun.*, vol. 13, no. 3, pp. 436–451, Jun. 2019.

[55] M. M. Sahin, O. Dizdar, B. Clerckx, and H. Arslan, “Multicarrier rate-splitting multiple access: Superiority of OFDM-RSMA over OFDMA and OFDM-NOMA,” *IEEE Commun. Lett.*, vol. 27, no. 11, pp. 3088–3092, Nov. 2023.

[56] M. M. Sahin, O. Dizdar, B. Clerckx, and H. Arslan, “OFDM-RSMA: Robust transmission under inter-carrier interference,” 2023. [Online]. Available: https://arxiv.org/pdf/2310.01686.pdf

[57] 3GPP, “5G; NR; physical channels and modulation,” *3rd Gener. Partnership Project (3GPP)*, Technical Specification (TS) 38.211, 2022, version 17.2.0. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/38_series/38.211/38211-h20.zip

[58] 3GPP, “Universal mobile telecommunications system (UMTS); LTE; multimedia broadcast/multicast service (MBMS); architecture and functional description,” 3rd Generation Partnership Project (3GPP), Technical specification (TS) 23. 246, Apr. 2022, version 17.0.0. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/23_series/23.246/23246-h00.zip

[59] G. P. McCormick, “Computability of global solutions to factorable nonconvex programs: Part I–Convex underestimating problems,” *Math. Prog.*, vol. 10, no. 1, pp. 147–175, 1976.

[60] L.-N. Tran, M. F. Hanif, A. Tolli, and M. Juntti, “Fast converging algorithm for weighted sum rate maximization in multicell MISO downlink,” *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 872–875, Dec. 2012.

[61] F. Alavi, K. Cumanan, Z. Ding, and A. G. Burr, “Robust beamforming techniques for non-orthogonal multiple access systems with bounded channel uncertainties,” *IEEE Commun. Lett.*, vol. 21, no. 9, pp. 2033–2036, Sep. 2017.

[62] H. Fu, S. Feng, W. Tang, and D. W. K. Ng, “Robust secure beamforming design for two-user downlink MISO rate-splitting systems,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 8351–8365, Dec. 2020.

[63] D. Xu, X. Yu, Y. Sun, D. W. K. Ng, and R. Schober, “Resource allocation for IRS-assisted full-duplex cognitive radio systems,” *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7376–7394, Dec. 2020.

Luis F. Abanto-Leon received the master’s degree in communications engineering from Tohoku University, Japan, in 2015. He is currently working toward the PhD degree with Technische Universität (TU) Darmstadt, Department of Computer Science. His research focuses on optimization theory, signal processing, and algorithm design for radio resource management for 5G wireless networks.

Aravindh Krishnamoorthy (Graduate Student Member, IEEE) received the master’s degree in communication and multimedia engineering from the Friedrich-Alexander University of Erlangen-Nuremberg (FAU), Germany, in 2015. He is currently working toward the PhD degree with the Institute for Digital Communications, FAU, under the sponsorship of Fraunhofer Institute for Integrated Circuits (IIS), Erlangen, Germany. He returned to academics in 2012, after a successful industrial career spanning several years in Philips Semiconductors and Ericsson. His industrial experience includes algorithm design for audio signal processing and system and algorithm design for 4th and 5th generation (4 G and 5 G) wireless communication systems. His current research interests include information and communication theory, probability theory, and (numerical) linear algebra. Hi is also an active reviewer for several IEEE journals.

Andres Garcia-Saavedra received the PhD degree from the University Carlos III of Madrid (UC3M), in 2013. He then joined Trinity College Dublin (TCD), Ireland, as a research fellow until 2015. Currently, he is a principal researcher with NEC Laboratories Europe. His research interests include the application of fundamental mathematics to real-life wireless communication systems.

Gek Hong Sim received the bachelor’s degree in telecommunication and the first master’s degree in engineering science from Multimedia University, Malaysia, in 2007 and 2011, respectively, and the master’s and PhD degrees in telematics engineering from the University Carlos III Madrid, in 2012 and 2015, respectively. She is a postdoctoral researcher with the Department of Computer Science, Technische Universität Darmstadt, Germany. Her research interests include multicast scheduling, precoding, and MAC layer optimization for millimeter-wave networks.

Robert Schober (Fellow, IEEE) received the diplom (Univ.) and PhD degrees in electrical engineering from Friedrich-Alexander University of Erlangen-Nuremberg (FAU), Germany, in 1997 and 2000, respectively. From 2002 to 2011, he was a professor and Canada research chair with the University of British Columbia (UBC), Vancouver, Canada. Since January 2012 he is an Alexander von Humboldt Professor and the chair for Digital Communication with FAU. His research interests include fall into the broad areas of Communication Theory, Wireless and Molecular Communications, and Statistical Signal Processing. Robert received several awards for his work including the 2002 Heinz Maier Leibnitz Award of the German Science Foundation (DFG), the 2004 Innovations Award of the Vodafone Foundation for Research in Mobile Communications, a 2006 UBC Killam Research Prize, a 2007 Wilhelm Friedrich Bessel Research Award of the Alexander von Humboldt Foundation, the 2008 Charles McDowell Award for Excellence in Research from UBC, a 2011 Alexander von Humboldt Professorship, a 2012 NSERC E.W.R. Stacie Fellowship, a 2017 Wireless Communications Recognition Award by the IEEE Wireless Communications Technical Committee, and the 2022 IEEE Vehicular Technology Society Stuart F. Meyer Memorial Award. Furthermore, he received numerous Best Paper Awards for his work including the 2022 ComSoc Stephen O. Rice Prize. Since 2017, he has been listed as a Highly Cited Researcher by the Web of Science. Robert is a Fellow of the Canadian Academy of Engineering, a Fellow of the Engineering Institute of Canada, and a Member of the German National Academy of Science and Engineering. He was also the editor-in-chief of the *IEEE Transactions on Communications*, VP Publications of the IEEE Communication Society (Com-Soc), ComSoc Member at Large, and ComSoc Treasurer. He is the senior editor of the Proceedings of the IEEE and ComSoc president-elect.

Matthias Hollick received the PhD degree from Technische Universität (TU) Darmstadt, in 2004. He is currently the head of the Secure Mobile Networking Lab, Department of Computer Science, TU Darmstadt, Germany. He has been researching and teaching with TU Darmstadt, Universidad Carlos III de Madrid, and the University of Illinois at Urbana–Champaign. His research focus is on resilient, secure, privacy-preserving, and quality-of-service-aware communication for mobile and wireless systems and networks.