

Received October 15, 2017, accepted November 7, 2017, date of publication November 24, 2017, date of current version December 22, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2777601

Outage-Constrained Resource Allocation in Uplink NOMA for Critical Applications

DANIEL TWEED¹, (Student Member, IEEE), MAHSA DERAKHSHANI², (Member, IEEE),
SAEED PARSAEFARD³, (Member, IEEE), AND THO LE-NGOC¹, (Fellow, IEEE)

¹Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada

²Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, Loughborough LE11 3TU, U.K.

³Iran Telecommunications Research Center, Tehran 1439955471, Iran

Corresponding author: Daniel Tweed (daniel.tweed@mail.mcgill.ca)

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada CRD Grant with Huawei Technologies Canada Company Ltd.

ABSTRACT In this paper, we consider the resource allocation problem for uplink non-orthogonal multiple access (NOMA) networks whose users represent power-restricted but high priority devices, such as those used in sensor networks supporting health and public safety applications. Such systems require high reliability and robust resource allocation techniques are needed to ensure performance. We examine the impact on system and user performance due to residual cancellation errors resulting from imperfect successive interference cancellation (SIC) and apply the chance-constrained robust optimization approach to tackle this type of error. In particular, we derive an expression for the user outage probability as a function of SIC error variance. This result is used to formulate a robust joint resource allocation problem that minimizes user transmit power subject to rate and outage constraints of critical applications. As the proposed optimization problem is inherently non-convex and NP-hard, we apply the techniques of variable relaxation and complementary geometric programming to develop a computationally tractable two-step iterative algorithm based on successive convex approximation. Simulation results demonstrate that, even for high levels of SIC error, the proposed robust algorithm for NOMA outperforms the traditional orthogonal multiple access case in terms of user transmit power and overall system density, i.e., serving more users over fewer sub-carriers. The chance-constrained approach necessitates a power-robustness tradeoff compared with non-robust NOMA but effectively enforces maximum user outage and can result in transmit power savings when users can accept a higher probability of outage.

INDEX TERMS Non-orthogonal multiple access, dynamic resource allocation, robust optimization theory, complementary geometric programming.

I. INTRODUCTION

A. MOTIVATION

The targets set for fifth-generation (5G) networks require significant increases in spectral and power efficiency, data rates, and traffic density [1]. Non-orthogonal multiple access (NOMA) is a promising technique being positioned to meet these targets by allowing several users to concurrently share radio resources through so-called power-domain multiplexing [2], [3]. The NOMA resource management problem is inherently complex and to leverage these multiplexing gains efficient dynamic resource allocation is essential. NOMA has drawn a lot of attention recently with many approaches being proposed, including sub-optimal solutions of low computational complexity which produce near-optimal results.

These approaches mainly rely on accurate channel statistics and perfect processing of received signals yet, due to the nature of wireless channels and user mobility, such accuracy may not be feasible. To support critical applications, where user outage can be equivalent to system failure, robust optimization approaches are needed to provide the necessary reliability, which is the main focus of this paper.

B. BACKGROUND AND RELATED WORKS

Due to importance of NOMA and its related dynamic resource management in 5G, there exists a vast number of related works which cannot be summarized in a straightforward manner. Downlink (DL) NOMA has been extensively studied and uplink (UL) NOMA is of increasing interest due

to the greater potential multiplexing gains and the expected increase in the number of devices for which UL traffic will outweigh DL traffic, such as distributed sensor networks. In this work, we are trying to reply to the following aspects via this literature review: 1) The importance of UL NOMA and its resource management compared to that of the DL NOMA; and 2) Related works considering the uncertainty in resource management for NOMA and the main novelty of this paper compared to the past works.

When first proposed in [4] for DL, NOMA demonstrated gains over orthogonal multiple access (OMA) techniques in terms of traffic density, and power and spectral efficiency. DL NOMA can achieve better outage performance and ergodic sum rate than OMA, subject to careful sub-carrier and power allocation; however, the successive interference cancellation (SIC) required in NOMA to resolve individual signals introduces significant complexity [5]. As a result, NOMA has so far been limited to two-user DL systems such as multi-user superposition transmission incorporated into LTE-A, but has generated significant interest with many resource allocation techniques being proposed [6]. For example, in [7] a jointly optimal power and sub-carrier allocation is presented based on monotonic optimization. Sub-optimal solutions which achieve close to optimal performance based on successive convex approximation (SCA) and variable relaxation are also presented. In [8], integer linear programming for sub-carrier allocation is used and algorithms for power allocation based on SCA for low-complexity (SCALE), the arithmetic-geometric mean approximation (AGMA), and difference of convex functions are examined. In [9], the NOMA power minimization problem is proved to be NP-hard and an algorithm to solve a relaxed convex form is presented. The work in [10] investigates the application of DL NOMA in virtualized wireless networks (VWN) with quality of service (QoS) guaranteed by slice-level group rate reservations.

In UL NOMA, SIC is performed by the base station (BS) which can manage the complexity to potentially support more users in each sub-carrier than the DL case. However, in UL NOMA interference for each user is a function of the channel state information (CSI) of all users in the cell [2]. Further, issues of synchronization are presented for UL NOMA due to the multi-path aspect and different propagation delays of distributed mobile users, though the recent work in [11] presents a low-complexity asynchronous SIC. UL NOMA has been less extensively studied than DL but has recently been gaining more attention. A dynamic power allocation scheme for two-user UL and DL NOMA systems is presented in [12] which is structured to guarantee user rates improve compared to OMA. Considering a stepped level back-off based power allocation, [13] derives an expression for sum rate and outage probability for two-user UL NOMA systems. A performance analysis of power minimization in UL NOMA systems supporting large deployments of low power machine type communication devices is presented in [14]. In [15], users in multi-cell UL NOMA systems are modelled

as independent Poisson point processes to derive the user and mean rate coverage probabilities.

As noted, DL NOMA has been extensively studied to increase spectral efficiency, while UL NOMA is of increasing interest due to the greater potential multiplexing gains and the expected increase in the number of devices for which UL traffic will outweigh DL traffic, such as distributed sensor networks. Consequently, this paper focuses on the UL NOMA for critical applications in 5G.

Uncertainty in system and channel statistics needed for NOMA resource allocation has also been studied. For example, statistical CSI for user grouping, power allocation, and decoding order is studied in [16] and the user outage balancing problem in DL NOMA is solved using minimum weighted success probability maximization. [17] examines DL NOMA for the case of perfect CSI being available to all users but with only 1-bit feedback to the BS used to determine the power allocation policy and presents a closed-form expression for the user outage probability. Reference [18] studies power minimization for rate maximization under outage constraints in DL NOMA systems where only average CSI is available at the BS. In [19], worst-case CSI uncertainty in DL NOMA is considered to propose a robust joint resource allocation algorithm based on SCALE for power allocation and integer non-linear programming for sub-carrier allocation.

C. CONTRIBUTIONS

In this work, we consider UL NOMA in the framework of VWN, which is positioned as the preferred architecture for 5G due to the available increase in network utility as well as improved spectrum and power efficiencies [20]. In VWN, sets of specific groups of users, called slices, access a share of the network resources while the required QoS for slice users should be preserved and isolation between slices should be provided based on negotiated service level agreements (SLA) [21]. In this context, we consider slices as groups of users serving a particular critical application, with QoS requirements driven by the specific needs of that application. To mathematically represent these issues of slicing and isolation in resource management problems, the minimum rate of each user per slice is preserved. However, preserving minimum rate in dynamic wireless networks can be challenging and in the context of UL NOMA, where SIC is used to resolve individual signals, errors in performing SIC can degrade system and user performance [2]. Such errors can result from any of several sources including synchronization, inaccurate CSI, and the inherent level of accuracy in performing SIC.

To consider the above sources of error and protect the QoS and isolation of slices, we apply the techniques of chance-constrained optimization theory, where the maximum outage probability of each critical application is kept below a predefined value [22], [23]. Since the resource allocation problem with this type of constraint suffers from significant computational complexity, we first apply the Chebyshev approximation via the Chebyshev-Cantelli inequality to reach a more

tractable formulation. We then use this result to formulate a robust outage-constrained resource allocation problem which minimizes transmit power of users, which is highly desirable for critical applications which rely on low-power sensors, while ensuring slice isolation and outage performance.

The robust problem constrains on a lower-bound on achieved rate to ensure outage performance and remains both non-convex and NP-hard [9]. To tackle this complexity, variable relaxation and complementary geometric programming (CGP) via AGMA to approximate non-convex constraints are used to develop an efficient iterative algorithm. The proposed algorithm is then evaluated in terms of user transmit power, outage, and system density performance. Notwithstanding a power-robustness trade-off, even for high levels of SIC error, the proposed algorithm for UL NOMA can support more users at lower average transmit power on fewer sub-carriers than the corresponding OMA system. Since the proposed algorithm jointly optimizes over power and sub-carrier allocations and there is no minimum user grouping, in cases where channel conditions are unfavourable to power-domain multiplexing an orthogonal solution can be produced. Further, the proposed algorithm can allow slices to set QoS targets to manage this power-robustness trade-off via adjusting the predefined threshold of outage probability to benefit from power savings when a higher probability of outage is acceptable.

In contrast to [16]–[19], which focus on the availability or accuracy of CSI in DL systems, we examine the performance degradation in UL NOMA due to residual cancellation errors via imperfect SIC, independent of the source of error. Further, we specifically focus on increasing the gains possible under UL NOMA by considering higher order multiplexing than the typically considered two-user systems in other works, and optimize jointly over sub-carrier and power allocation rather than considering fixed groupings, power allocations, or fixed power levels or policies.

D. ORGANIZATION OF PAPER

The remainder of this paper is organized as follows. In Section II, a system model for a virtualized UL NOMA system is presented and the problem formulation based on an outage probability analysis is presented in Section III. The derivation of the proposed joint resource allocation algorithm is provided in Section IV. Section V presents simulation results and their related analysis, followed by the conclusion in Section VI.

II. SYSTEM MODEL

Consider a single BS employing UL NOMA to support a VWN in which each slice $s \in \mathcal{S}$ serves battery-dependent users supporting a particular critical application. For each slice, its group of users is $\mathcal{K}_s = \{1, \dots, K_s\}$ and the total number of users in the system is $K = \sum_{s \in \mathcal{S}} K_s$. Each slice has negotiated QoS based on the application priority and requirements of its users as a minimum rate, R_s^{rsv} , which must be ensured for each user in supporting critical applications.

The set of available sub-carriers, $\mathcal{N} = \{1, \dots, N\}$, is shared by all K users and SIC is applied at the BS to resolve the individual signals when a given sub-carrier is used by more than one user concurrently.

To perform SIC, CSI is used to rank users in each sub-carrier they are assigned to and allocate user transmission power levels so that individual signals remain sufficiently distinct and can be resolved from the superposed received signal. Letting $h_{k_s,n}$ denote the channel gain on the n^{th} channel for user k_s , in each sub-carrier users are indexed such that $h_{1,n} > h_{2,n} > \dots > h_{K,n}$. The BS then decodes the user ranked i on sub-carrier n by applying SIC to remove the signals of all users whose rank is lower than i . The remaining users, whose rank is greater than i , are treated as unresolvable interference. Thus, assuming the user transmits with power $\beta_{i,n}$, the signal-to-interference-plus-noise ratio (SINR) experienced when decoding user i on sub-carrier n is

$$\text{SINR}_{i,n} = \frac{\beta_{i,n} h_{i,n}}{\sigma_{i,n}^2 + \sum_{j=i+1}^K \beta_{j,n} h_{j,n} + I_{i,n}^e}, \quad \forall i, n \quad (1)$$

where $\sigma_{i,n}^2$ is additive white Gaussian noise (AWGN) and $I_{i,n}^e$ is the residual interference which may result from imperfect cancellation of the transmissions of users $1 \leq j < i$.

If we let $\gamma_{i,n} = 1 + \text{SINR}_{i,n}$ and define $\alpha_{k_s,n} \in \{0, 1\}$ to be the sub-carrier allocation indicator, where $\alpha_{k_s,n} = 1$ means that sub-carrier n is allocated to user k_s , the achieved rate for user k_s , whose rank i is determined independently for each sub-carrier n , is

$$R_{k_s} = \sum_{n \in \mathcal{N}} \alpha_{k_s,n} \log(\gamma_{i,n}). \quad (2)$$

The constraint on achieved rate to meet reservations can then be expressed as

$$\text{C1} : R_{k_s} \geq R_s^{\text{rsv}}, \quad \forall k_s \in \mathcal{K}_s, \forall s \in \mathcal{S}. \quad (3)$$

For practicality, we limit the user transmit power to P_{max} and the number of users per sub-carrier is constrained to N_{max} by the following two constraints

$$\text{C2} : \sum_{n \in \mathcal{N}} \beta_{k_s,n} \leq P_{\text{max}}, \quad \forall s \in \mathcal{S}, \forall k_s \in \mathcal{K}_s, \quad (4)$$

$$\text{C3} : \sum_{s \in \mathcal{S}} \sum_{k_s \in \mathcal{K}_s} \alpha_{k_s,n} \leq N_{\text{max}}, \quad \forall n \in \mathcal{N}. \quad (5)$$

Further, we restrict power allocation to only those sub-carriers which are allocated to users with

$$\text{C4} : \beta_{k_s,n} - \alpha_{k_s,n} \times P_{\text{max}} \leq 0, \quad \forall n \in \mathcal{N}, \forall s \in \mathcal{S}, \forall k_s \in \mathcal{K}_s. \quad (6)$$

Then, the minimum power needed to meet slice reservations within practical system limitations is obtained by solving the following optimization problem

$$\begin{aligned} \min_{\alpha, \beta} \quad & \max_{\substack{\forall s \in \mathcal{S} \\ \forall k_s \in \mathcal{K}_s, n \in \mathcal{N}}} \sum \alpha_{k_s,n} \beta_{k_s,n}, \\ \text{Subject to:} \quad & \text{C1-C4} \end{aligned} \quad (7)$$

where α and β are the $K \times N$ matrices of $\alpha_{k_s,n}$ sub-carrier allocation indicators and $\beta_{k_s,n}$ user transmit power factors.

III. ROBUST FORMULATION WITH OUTAGE PROBABILITY ANALYSIS

The optimization in (7) assumes that SINR can be calculated from accurate CSI and a known level of SIC residual cancellation error; however, CSI may not be accurately known and, assuming that SIC errors occur, the occurrence and magnitude of residual cancellation errors are non-deterministic and depend on one or more factors including, but not limited to, the type of SIC employed, thermal and environmental noise, system parameters such as number of users, user mobility and synchronization of received signals, and accuracy of CSI used in the resource allocation and decoding. From (1), we see that after SIC is performed at the BS residual cancellation errors will degrade the achieved SINR and may reduce users achieved rate below their reserved R_s^{rsv} , i.e., the user will be in outage.

To address this, we apply the techniques of robust optimization theory to reformulate the problem based on chance-constrained resource allocation which limits user outage probability to maintain isolation considering the uncertainty in achieved SINR [22]. In this context, the uncertain value is modelled as its own estimated value and additive error, and the nominal optimization problem, i.e., (7), is mapped to its own robust counterpart considering the uncertain parameter [23]. This mapping is usually done via two basic approaches: 1) If the stochastic information of error is available, the nominal optimization problem is mapped to its Bayesian (stochastic) counterpart where the stochastic function of objective function and constraints (e.g., mean value) are optimized; or 2) When the error is bounded in some specific region, called the uncertainty region, worst-case optimization theory is applied where the optimal value is derived for the worst-case condition of error. Both of these two approaches have their own drawbacks. For instance, via the Bayesian approach, only average performance is achieved which is not acceptable for most critical applications; while via the worst-case approach, the network performance is considerably reduced via considering the error in the maximum extent [23].

In this section we resort to robust optimization theory to protect the minimum reserved rate of each user within each slice against residual cancellation error. However, since the uncertain value belongs to the constraint of our optimization problem, we apply the chance-constrained approach where the probability of violation of the constraint is limited to a certain value [22]. Such techniques consider the expected values of the data and accept sub-optimal solutions which remain feasible if the data changes; however, this introduces a trade-off between robustness and optimality via changing the limit of violation of constraints, as we will explain [24].

Here in this work, SINR degradation from SIC residual cancellation error results in a non-zero residual interference

from cancelled signals and can be modelled by

$$I_{i,n}^e = \sum_{j=1}^{i-1} \beta_{j,n} h_{j,n} \|e_{j,n}\|^2, \quad (8)$$

where we assume that $e_{j,n} \sim \mathcal{CN}(0, \sigma_e^2)$.¹ As a result, $\frac{1}{\sigma_e^2} \|e_{j,n}\|^2$ is a random variable which has a chi-squared distribution with 2 degrees of freedom.

The uncertain parameter, $e_{j,n}$ in (8), affects the constraint of our optimization problem, i.e., C1 in (3), and not the objective function, which allows us to apply chance-constrained optimization theory [22]. To apply this approach, we reformulate C1 and consider the maximum outage probability, $0 \leq \epsilon_s \leq 1$, which has been negotiated by each slice based on service levels required by their users. Therefore, C1 can be re-expressed for all users $k_s \in \mathcal{K}_s$ as its own robust counterpart

$$\Pr [R_{k_s} \leq R_s^{\text{rsv}}] \leq \epsilon_s, \quad (9)$$

which is equal to

$$\Pr [R_{k_s} \geq R_s^{\text{rsv}}] \geq 1 - \epsilon_s. \quad (10)$$

Note that (10) can be considered as a maximum outage probability of reserved rate of each user. Depending on the SLA and the request of critical applications, this threshold can be adjusted, e.g., for highly critical mission applications, one can use very small value of ϵ_s , resulting in significant computational complexity and less optimality [22], [23]. However, one can also allow for increased efficiencies in cases when looser constraints are acceptable.

In this context, to tackle the computational complexity and reach a more tractable formulation, the constraint with uncertain parameters is relaxed and a more tractable formulation is used in place of the original one [22], [23]. Here, we apply the Chebyshev approximation using the Chebyshev-Cantelli inequality, defined for $\eta > 0$ as

$$\Pr (X - \mathbf{E}[X] \geq \eta) \leq \frac{\mathbf{Var}[X]}{\mathbf{Var}[X] + \eta^2}, \quad (11)$$

to replace (10). Substituting $X = R_{k_s}$ and $\eta = R_s^{\text{rsv}} - \mathbf{E}[R_{k_s}]$, we have

$$\begin{aligned} \Pr [R_{k_s} - \mathbf{E}[R_{k_s}] \geq R_s^{\text{rsv}} - \mathbf{E}[R_{k_s}]] \\ \leq \frac{\mathbf{Var}[R_{k_s}]}{\mathbf{Var}[R_{k_s}] + (R_s^{\text{rsv}} - \mathbf{E}[R_{k_s}])^2}. \end{aligned} \quad (12)$$

From the inequality in (12), the constraint in (10) can be relaxed to a deterministic form as

$$1 - \epsilon_s \leq \frac{\mathbf{Var}[R_{k_s}]}{\mathbf{Var}[R_{k_s}] + (R_s^{\text{rsv}} - \mathbf{E}[R_{k_s}])^2}, \quad (13)$$

¹This assumption is made by considering the potential sources of error, i.e., thermal noise, CSI inaccuracy, and asynchronization. For each source, and others not explicitly accounted for, the magnitude of the resulting cancellation error either follows a Gaussian distribution, e.g., thermal noise and asynchronicity received signals, as shown in [11], or at worst can be assumed to be independent and identically distributed (i.i.d) for any sources of error resulting from co-channel interference of other users in the system. For each source of error, the sum of i.i.d. random variables quickly converges to a Gaussian variable.

which can be re-arranged as

$$\mathbf{E}[R_{k_s}] + \sqrt{\mathbf{Var}[R_{k_s}] \frac{\epsilon_s}{1 - \epsilon_s}} \geq R_s^{\text{rsv}}. \quad (14)$$

While (14) is more tractable than (10), we have to calculate the mean and variance of R_{k_s} . In order to obtain the required statistics, recall that the achieved rate on each sub-carrier is independent. Then, from (2), we have

$$\mathbf{E}[R_{k_s}] = \sum_{n \in \mathcal{N}} \alpha_{k_s, n} \mathbf{E}[\log(\gamma_{i, n})], \quad (15)$$

$$\mathbf{Var}[R_{k_s}] = \sum_{n \in \mathcal{N}} \alpha_{k_s, n} \mathbf{Var}[\log(\gamma_{i, n})]. \quad (16)$$

To find $\mathbf{E}[\log(\gamma_{i, n})]$ and $\mathbf{Var}[\log(\gamma_{i, n})]$, we will approximate the rate function with the help of the Taylor series of a logarithmic function. Approximating with two terms, we have

$$\log(\gamma_{i, n}) \approx \log(\mathbf{E}[\gamma_{i, n}]) + \frac{1}{\mathbf{E}[\gamma_{i, n}]} (\gamma_{i, n} - \mathbf{E}[\gamma_{i, n}]). \quad (17)$$

We can take the expected value and variance of the both sides to obtain

$$\mathbf{E}[\log(\gamma_{i, n})] \approx \log(\mathbf{E}[\gamma_{i, n}]), \quad (18)$$

$$\mathbf{Var}[\log(\gamma_{i, n})] \approx \frac{\mathbf{Var}[\gamma_{i, n}]}{\mathbf{E}^2[\gamma_{i, n}]}. \quad (19)$$

Now, we need to calculate $\mathbf{E}[\gamma_{i, n}]$ and $\mathbf{Var}[\gamma_{i, n}]$. Similarly, we use the Taylor series to approximate the one plus SINR expression as a function of $I_{i, n}^e$. With two terms, we have

$$\begin{aligned} & \gamma_{i, n}(I_{i, n}^e) \\ & \approx 1 + \frac{a_{i, n}}{b_{i, n} + \mathbf{E}[I_{i, n}^e]} - \frac{a_{i, n}}{(b_{i, n} + \mathbf{E}[I_{i, n}^e])^2} (I_{i, n}^e - \mathbf{E}[I_{i, n}^e]) \end{aligned} \quad (20)$$

and consequently

$$\mathbf{E}[\gamma_{i, n}] \approx 1 + \frac{a_{i, n}}{b_{i, n} + \mathbf{E}[I_{i, n}^e]}, \quad (21)$$

$$\mathbf{Var}[\gamma_{i, n}] \approx \frac{a_{i, n}^2}{(b_{i, n} + \mathbf{E}[I_{i, n}^e])^4} \mathbf{Var}[I_{i, n}^e], \quad (22)$$

where, for the sake of notational simplicity, we have defined

$$a_{i, n} = \beta_{i, n} h_{i, n}, \quad (23)$$

$$b_{i, n} = \sigma_{i, n}^2 + \sum_{j=i+1}^K \beta_{j, n} h_{j, n}. \quad (24)$$

As a result, considering (15), (16), (18), (19), (21) and (22), we can express the mean and variance of $R_{k_s, n}$ as

$$\mathbf{E}[R_{k_s, n}] \approx \alpha_{k_s, n} \log \left(1 + \frac{a_{i, n}}{b_{i, n} + \mathbf{E}[I_{i, n}^e]} \right), \quad (25)$$

$$\begin{aligned} \mathbf{Var}[R_{k_s, n}] & \approx \alpha_{k_s, n} \left(\frac{a_{i, n}}{b_{i, n} + \mathbf{E}[I_{i, n}^e]} \right)^2 \\ & \times \frac{\mathbf{Var}[I_{i, n}^e]}{(a_{i, n} + b_{i, n} + \mathbf{E}[I_{i, n}^e])^2}, \end{aligned} \quad (26)$$

From the definition of $I_{i, n}^e$ in (8), we have

$$\mathbf{E}[I_{i, n}^e] = \sum_{j=1}^{i-1} 2a_{j, n} \sigma_e^2, \quad (27)$$

$$\mathbf{Var}[I_{i, n}^e] = \sum_{j=1}^{i-1} 4a_{j, n}^2 \sigma_e^4, \quad (28)$$

Finally, substituting (25) and (26) in (14), the approximation of the outage probability constraint in (10), $\forall k_s \in \mathcal{K}_s, \forall s \in \mathcal{S}$ can be written as

$$\begin{aligned} \tilde{\text{C1}} : & \sum_{n \in \mathcal{N}} \alpha_{k_s, n} \log \left(1 + \frac{a_{i, n}}{b_{i, n} + \mathbf{E}[I_{i, n}^e]} \right) + \sqrt{\frac{\epsilon_s}{1 - \epsilon_s}} \\ & \times \sqrt{\sum_{n \in \mathcal{N}} \alpha_{k_s, n} \left(\frac{a_{i, n}}{b_{i, n} + \mathbf{E}[I_{i, n}^e]} \right)^2 \frac{\mathbf{Var}[I_{i, n}^e]}{(a_{i, n} + b_{i, n} + \mathbf{E}[I_{i, n}^e])^2}} \geq R_s^{\text{rsv}}. \end{aligned} \quad (29)$$

Then, subject to $\tilde{\text{C1}}, \text{C2-4}$, the chance-constrained counterpart of (7) is

$$\min_{\alpha, \beta} \max_{\substack{\forall s \in \mathcal{S} \\ \forall k_s \in \mathcal{K}_s}} \sum_{n \in \mathcal{N}} \alpha_{k_s, n} \beta_{k_s, n}. \quad (30)$$

Remark 1: Note that since $\tilde{\text{C1}}$ is not a linear constraint, we cannot apply the conventional approaches of chance-constrained optimization, such as the Bernstein approximation [23]. Therefore, in the sequel we apply the techniques of CGP to reach a more tractable formulation.

IV. PROPOSED ITERATIVE ALGORITHM

Both $\tilde{\text{C1}}$ and C4 are non-convex and the optimization in (30) is computationally intractable due to the binary sub-carrier allocation indicator, α . In order to develop an efficient resource allocation algorithm, we first relax the elements of α to be continuous on the interval $[0, 1]$, and then decompose (30) into separate power and sub-carrier allocation problems which, while simpler than the original problem, remain challenging due to the nature of the constraints.

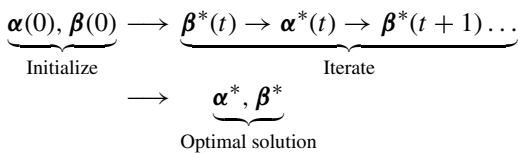
The solution to the relaxed problem may be a binary matrix, in which case it is the optimal solution to both the original and relaxed problem. In the case where it is not a binary matrix, which will be the case in general, it remains an optimal solution to the relaxed problem and provides a lower bound on the minimization. In implementation, the sub-carrier allocation can be recovered to binary via integer rounding, i.e., when the proposed iterative algorithm for sub-carrier allocation is converged, in each sub-carrier n we choose the N_{max} largest $\alpha_{k_s, n} > 0.5$ and set them to 1 while the remaining elements are set to 0.

Overall, the resource allocation algorithm is shown in Algorithm 1. We first solve for an optimal power allocation given a fixed sub-carrier allocation and then we use this power allocation to solve for an optimal sub-carrier allocation. In each step, we apply CGP using AGMA to approximate non-convex constraints with monomial functions to convert the

Algorithm 1 Iterative Power and Sub-carrier Allocation

Initialize: Set $t = t_1 = t_2 = 0$, $\beta^*(0) = [P_{\max}/N]_{K \times N}$ and $\alpha^*(0) = [\mathbf{1}]_{K \times N}$.
repeat
 Step 1: Derive power allocation matrix, $\beta^*(t)$
 repeat
 $t_1 = t_1 + 1$
 Step 1.1: Update $\Gamma_{k_s,n}(t_1)$, $\theta_{k_s}(t_1)$, $\nu_{i,n}(t_1)$, $\psi_{j,n}(t_1)$, $\rho_{i,n}(t_1)$, $\mu_{j,n}(t_1)$, $\zeta_{i,n}(t_1)$, $\Delta_{k_s,n}(t_1)$, and $\Lambda_{j,n}(t_1)$ from (41), (42), (47-51), (55), (57)
 Step 1.2: Find $\beta^*(t_1)$ according to (58) using CVX [25]
 until $\|\beta^*(t_1) - \beta^*(t_1 - 1)\| \leq \epsilon_1$
 Step 2: Derive sub-carrier allocation matrix, $\alpha^*(t)$
 repeat
 $t_2 = t_2 + 1$
 Step 2.1: Update $\tau_{k_s,n}(t_2)$, $\phi_{k_s,n}(t_2)$, $\nu_{k_s,n}(t_2)$, and $\omega_{k_s,n}(t_2)$ from (65-67), (69)
 Step 2.2: Find $\alpha^*(t_2)$ according to (70) using CVX [25]
 until $\|\alpha^*(t_2) - \alpha^*(t_2 - 1)\| \leq \epsilon_2$
until $\|\beta^*(t) - \beta^*(t - 1)\| \leq \epsilon_1$ && $\|\alpha^*(t) - \alpha^*(t - 1)\| \leq \epsilon_2$, otherwise $t = t + 1$

problem to geometric programming (GP) form which can be solved efficiently. The output of each solved GP problem is used as input to the approximation formulas used to construct the next GP problem. Thus, on each iteration of each step, we successively improve our approximation of the original problem as convex with the output of the previous iteration. This process can be represented by



For $0 < \epsilon_1, \epsilon_2 \ll 1$, the iterative procedure is stopped when

$$\|\beta^*(t) - \beta^*(t - 1)\| \leq \epsilon_1 \text{ and } \|\alpha^*(t) - \alpha^*(t - 1)\| \leq \epsilon_2$$

A. OVERVIEW OF COMPLEMENTARY GEOMETRIC PROGRAMMING

Geometric programming (GP) is a class of non-linear optimization which can be efficiently solved and has many applications in science and engineering [26]. The standard form of GP is

$$\begin{aligned} &\min_{\mathbf{x}} f_0(\mathbf{x}) \\ &\text{Subject to: } f_i(\mathbf{x}) \leq 1, \quad i = 1, 2, \dots, m \\ &\quad g_j(\mathbf{x}) = 1, \quad j = 1, 2, \dots, M \end{aligned} \quad (31)$$

for non-negative optimization variables $\mathbf{x} = [x_1, x_2, \dots, N]$ and, for $c_{j,n} > 0$ and $a_{j,n} \in \mathbb{R}$, $g_j = \prod_{n=1}^N c_{j,n} x_n^{a_{j,n}}$, $\forall j$, i.e. are monomial functions, and f_i are products of monomial terms, $\forall i$ i.e. are posynomials. The restrictions on objective

function and constraints in GP problems cannot be met for many practical problems, e.g. if f_i or g_j are posynomial or ratios of posynomials for some j , f_i are lower bounds rather than upper bounds, etc.

CGP is an approach to solve problems formulated in terms of rational functions of posynomial terms [27]. Such problems can be solved iteratively using SCA by substituting monomial approximations to convert the problem into the GP form on each iteration via the arithmetic-geometric mean inequality, i.e., for $a_j > 0$ and $d_j \geq 0$ s.t. $\sum_j d_j = 1$,

$$\sum_j a_j \geq \prod_j \left(\frac{a_j}{d_j}\right)^{d_j} \quad (32)$$

The application of (32) is called the arithmetic-geometric mean approximation (AGMA) and is used to transform functions into posynomial form. For example:

For $f_1(\mathbf{x}) : x_1 + x_2 \geq 1$, we can reformulate as

$$\frac{1}{x_1 + x_2} \leq 1 \quad (33)$$

Applying AGMA at iteration index t we approximate this constraint as

$$\left(\frac{x_1(t)}{\xi_1(t)}\right)^{\xi_1(t)} \left(\frac{x_2(t)}{\xi_2(t)}\right)^{\xi_2(t)} \leq 1 \quad (34)$$

with $\xi_k(t) = \frac{x_k(t-1)}{(x_1(t-1)+x_2(t-1))}$, $k = 1, 2$. Applying these approximations to all f_i and g_j , and the objective functions, as needed, the resulting GP problem can be solved efficiently via numerical methods. The convergence of CGP has been proven in [27] and it has been shown that the output of algorithms based on SCA of the problem to its GP form have very close performance to the optimal solution [17], [28], [29].

B. POWER ALLOCATION

With a fixed sub-carrier allocation, the optimization problem becomes

$$\min_{\beta} \max_{\substack{\forall s \in \mathcal{S} \\ \forall k_s \in \mathcal{K}_s}} \sum_{n \in \mathcal{N}} \alpha_{k_s,n} \beta_{k_s,n} \quad (35)$$

Subject to: $\tilde{C}1, C2, C4$.

C2 and C4 are in the proper GP form for fixed sub-carrier allocation, but $\tilde{C}1$ is non-convex and appropriate approximations are required. To deal with $\tilde{C}1$ we define new variables, $\mathbf{X} = [X_{k_s,n}]_{K \times N}$ and $\mathbf{Y} = [Y_{k_s}]_{K \times 1}$, and transform $\tilde{C}1$ into three constraints as follows

$$\begin{aligned} \text{C1.1: } \sum_{n \in \mathcal{N}} \alpha_{k_s,n} X_{k_s,n} + \sqrt{\frac{\epsilon_s}{1 - \epsilon_s}} Y_{k_s} &\geq R_s^{\text{rsv}}, \\ &\forall s \in \mathcal{S}, \forall k_s \in \mathcal{K}_s, \end{aligned} \quad (36)$$

$$\begin{aligned} \text{C1.2: } X_{k_s,n} &\leq \log \left(1 + \frac{a_{i,n}}{b_{i,n} + \sum_{j=1}^{i-1} 2a_{j,n} \sigma_e^2} \right), \\ &\forall s \in \mathcal{S}, \forall k_s \in \mathcal{K}_s, \forall n \in \mathcal{N}, \end{aligned} \quad (37)$$

$$\begin{aligned}
 \text{C1.3} : Y_{k_s}^2 &\leq \sum_{n \in \mathcal{N}} \alpha_{k_s, n} \left(\frac{a_{i, n}}{b_{i, n} + \sum_{j=1}^{i-1} 2a_{j, n} \sigma_e^2} \right)^2 \\
 &\times \frac{\sum_{j=1}^{i-1} 4a_{j, n}^2 \sigma_e^4}{(a_{i, n} + b_{i, n} + \sum_{j=1}^{i-1} 2a_{j, n} \sigma_e^2)^2}, \quad \forall s \in \mathcal{S}, \forall k_s \in \mathcal{K}_s.
 \end{aligned} \tag{38}$$

For simplicity of notation, let $C = \sqrt{\epsilon_s / (1 - \epsilon_s)}$, and we can re-write C1.1 as

$$\frac{R_s^{\text{rsv}}}{\sum_{n \in \mathcal{N}} \alpha_{k_s, n} X_{k_s, n} + CY_{k_s}} \leq 1. \tag{39}$$

Then, for iteration index t_1 , we can apply AGMA to approximate C1.1 with the following convex constraint

$$\begin{aligned}
 \widetilde{\text{C1.1}} : R_s^{\text{rsv}} &\times \prod_{n \in \mathcal{N}} \left(\frac{\alpha_{k_s, n} X_{k_s, n}(t_1)}{\Gamma_{k_s, n}(t_1)} \right)^{-\Gamma_{k_s, n}(t_1)} \\
 &\times \left(\frac{CY_{k_s}(t_1)}{\theta_{k_s}(t_1)} \right)^{-\theta_{k_s}(t_1)} \leq 1, \quad \forall s \in \mathcal{S}, \forall k_s \in \mathcal{K}_s,
 \end{aligned} \tag{40}$$

where

$$\Gamma_{k_s, n}(t_1) = \frac{\alpha_{k_s, n} X_{k_s, n}(t_1 - 1)}{\sum_{n \in \mathcal{N}} \alpha_{k_s, n} X_{k_s, n}(t_1 - 1) + CY_{k_s}(t_1 - 1)}, \tag{41}$$

and

$$\theta_{k_s}(t_1) = \frac{CY_{k_s}(t_1 - 1)}{\sum_{n \in \mathcal{N}} \alpha_{k_s, n} X_{k_s, n}(t_1 - 1) + CY_{k_s}(t_1 - 1)}. \tag{42}$$

To eliminate the logarithm, C1.2 can be expressed as

$$e^{X_{k_s, n}} \leq 1 + \frac{a_{i, n}}{b_{i, n} + \sum_{j=1}^{i-1} 2a_{j, n} \sigma_e^2}. \tag{43}$$

Then, approximating $e^{X_{k_s, n}}$ using the truncated Taylor series and re-arranging, we have

$$\begin{aligned}
 e^{X_{k_s, n}} &\approx \sum_{m=0}^{10} \frac{X_{k_s, n}^m}{m!} \leq \left(\frac{a_{i, n} + b_{i, n} + \sum_{j=1}^{i-1} 2a_{j, n} \sigma_e^2}{b_{i, n} + \sum_{j=1}^{i-1} 2a_{j, n} \sigma_e^2} \right) \\
 &\implies \sum_{m=0}^{10} \frac{X_{k_s, n}^m}{m!} \times \left(\frac{b_{i, n} + \sum_{j=1}^{i-1} 2a_{j, n} \sigma_e^2}{a_{i, n} + b_{i, n} + \sum_{j=1}^{i-1} 2a_{j, n} \sigma_e^2} \right) \leq 1.
 \end{aligned} \tag{44}$$

This approximation remains non-convex but can be transformed by AGMA into a convex constraint on each iteration as follows

$$\begin{aligned}
 \widetilde{\text{C1.2}} : \sum_{m=0}^{10} \frac{X_{k_s, n}^m(t_1)}{m!} \\
 &\times \left(b_{i, n}(t_1) + \sum_{j=1}^{i-1} 2a_{j, n}(t_1) \sigma_e^2 \right) \times z_{i, n}(t_1) \leq 1, \\
 &\forall s \in \mathcal{S}, \quad \forall k_s \in \mathcal{K}_s, \quad \forall n \in \mathcal{N}.
 \end{aligned} \tag{45}$$

Where, substituting from the definitions of b in (24) and applying AGMA and simplifying, we define $z_{i, n}(t_1)$ as the following convex function

$$\begin{aligned}
 z_{i, n}(t_1) &= \left(\frac{\sigma_{i, n}^2}{v_{i, n}(t_1)} \right)^{-v_{i, n}(t_1)} \times \prod_{j=1}^{i-1} \left(\frac{2a_{j, n}(t_1) \sigma_e^2}{\psi_{j, n}(t_1)} \right)^{-\psi_{j, n}(t_1)} \\
 &\times \prod_{j=i+1}^K \left(\frac{a_{j, n}(t_1)}{\rho_{j, n}(t_1)} \right)^{-\rho_{j, n}(t_1)} \times \left(\frac{a_{i, n}(t_1)}{\mu_{i, n}(t_1)} \right)^{-\mu_{i, n}(t_1)},
 \end{aligned} \tag{46}$$

with

$$v_{i, n}(t_1) = \sigma_{i, n}^2 / \zeta_{i, n}(t_1), \tag{47}$$

$$\psi_{j, n}(t_1) = 2a_{j, n}(t_1 - 1) \sigma_e^2 / \zeta_{i, n}(t_1), \tag{48}$$

$$\rho_{j, n}(t_1) = a_{j, n}(t_1 - 1) / \zeta_{i, n}(t_1), \tag{49}$$

$$\mu_{i, n}(t_1) = a_{i, n}(t_1 - 1) / \zeta_{i, n}(t_1), \tag{50}$$

and

$$\begin{aligned}
 \zeta_{i, n}(t_1) &= \sigma_{i, n}^2 + a_{i, n}(t_1 - 1) \\
 &+ \sum_{j=1}^{i-1} 2a_{j, n}(t_1 - 1) \sigma_e^2 + \sum_{j=i+1}^K a_{j, n}(t_1 - 1).
 \end{aligned} \tag{51}$$

For C1.3, we introduce auxiliary variable $\mathbf{W} = [W_{i, n}]_{K \times N}$ and define two new non-convex constraints

$$\text{C1.3.1} : Y_{k_s}^2 \leq \sum_{n \in \mathcal{N}} \alpha_{k_s, n} a_{i, n}^2 W_{i, n}, \tag{52}$$

$$\begin{aligned}
 \text{C1.3.2} : W_{i, n} \\
 &\leq \frac{\sum_{j=1}^{i-1} 4a_{j, n}^2 \sigma_e^4}{\left(b_{i, n} + \sum_{j=1}^{i-1} 2a_{j, n} \sigma_e^2 \right)^2 \left(a_{i, n} + b_{i, n} + \sum_{j=1}^{i-1} 2a_{j, n} \sigma_e^2 \right)^2}.
 \end{aligned} \tag{53}$$

We can again apply AGMA and on each iteration approximate C1.3.1 by

$$\begin{aligned}
 \widehat{\text{C1.3.1}} : Y_{k_s}^2(t_1) \\
 &\times \prod_{n \in \mathcal{N}} \left(\frac{\alpha_{k_s, n} a_{i, n}^2(t_1) W_{i, n}(t_1)}{\Delta_{k_s, n}(t_1)} \right)^{-\Delta_{k_s, n}(t_1)} \leq 1, \\
 &\forall s \in \mathcal{S}, \quad \forall k_s \in \mathcal{K}_s,
 \end{aligned} \tag{54}$$

with

$$\Delta_{k_s, n}(t_1) = \frac{\alpha_{k_s, n} a_{i, n}^2(t_1 - 1) W_{i, n}(t_1 - 1)}{\sum_{n \in \mathcal{N}} \alpha_{k_s, n} a_{i, n}^2(t_1 - 1) W_{i, n}(t_1 - 1)}, \tag{55}$$

and C1.3.2 by

$$\begin{aligned} \widetilde{\text{C1.3.2}} : & \frac{W_{i,n}(t_1)}{4\sigma_e^4} \times \left(b_{i,n}(t_1) + \sum_{j=1}^{i-1} 2a_{j,n}(t_1)\sigma_e^2 \right)^2 \\ & \times \left(a_{i,n}(t_1) + b_{i,n}(t_1) + \sum_{j=1}^{i-1} 2a_{j,n}(t_1)\sigma_e^2 \right)^2 \\ & \times \prod_{j=1}^{i-1} \left(\frac{a_{j,n}^2(t_1)}{\Lambda_{j,n}(t_1)} \right)^{-\Lambda_{j,n}(t_1)} \leq 1, \\ & \forall s \in \mathcal{S}, \quad \forall k_s \in \mathcal{K}_s, \quad \forall n \in \mathcal{N}, \end{aligned} \quad (56)$$

with

$$\Lambda_{j,n}(t_1) = \frac{a_{j,n}^2(t_1 - 1)}{\sum_{l=1}^{i-1} a_{l,n}^2(t_1 - 1)}. \quad (57)$$

Then, at each iteration t_1 , we solve

$$\begin{aligned} \min_{\beta, W, X, Y} \quad & \max_{\substack{\forall s \in \mathcal{S} \\ \forall k_s \in \mathcal{K}_s}} \sum_{n \in \mathcal{N}} \alpha_{k_s, n} \beta_{k_s, n}(t_1) \\ \text{Subject to:} \quad & \widetilde{\text{C1.1}}, \widetilde{\text{C1.2}}, \widetilde{\text{C1.3.1}}, \widetilde{\text{C1.3.2}}, \text{C2}, \text{C4}, \end{aligned} \quad (58)$$

which is in GP form and can be solved efficiently with standard convex optimization tools such as CVX [25].

C. SUB-CARRIER ALLOCATION

With a fixed power allocation, the optimization problem becomes

$$\begin{aligned} \min_{\alpha} \quad & \max_{\substack{\forall s \in \mathcal{S} \\ \forall k_s \in \mathcal{K}_s}} \sum_{n \in \mathcal{N}} \alpha_{k_s, n} \beta_{k_s, n} \\ \text{Subject to:} \quad & \widetilde{\text{C1}}, \text{C3}. \end{aligned} \quad (59)$$

C3 is in the proper GP form, but $\widetilde{\text{C1}}$ is non-convex and again needs to be transformed into GP form. For simplicity of notation, we define

$$\begin{aligned} L_{i,n} &= \log \left(1 + \frac{a_{i,n}}{b_{i,n} + \sum_{j=1}^{i-1} 2a_{j,n}\sigma_e^2} \right), \quad (60) \\ M_{i,n} &= \left(\frac{a_{i,n}}{b_{i,n} + \sum_{j=1}^{i-1} 2a_{j,n}\sigma_e^2} \right)^2 \\ & \times \frac{\sum_{j=1}^{i-1} 4a_{j,n}^2\sigma_e^4}{(a_{i,n} + b_{i,n} + \sum_{j=1}^{i-1} 2a_{j,n}\sigma_e^2)^2}, \quad (61) \end{aligned}$$

which are constants for fixed power allocation. With $C = \sqrt{\epsilon_s/(1-\epsilon_s)}$ as before, we again introduce the variable Y_{k_s} , but can now express $\widetilde{\text{C1}}$ as two non-convex constraints

$$\text{C1.4} : \sum_{n \in \mathcal{N}} \alpha_{k_s, n} L_{i,n} + CY_{k_s} \geq R_s^{\text{rsv}}, \quad \forall s \in \mathcal{S}, \quad \forall k_s \in \mathcal{K}_s, \quad (62)$$

$$\text{C1.5} : Y_{k_s}^2 \leq \sum_{n \in \mathcal{N}} \alpha_{k_s, n} M_{i,n}, \quad \forall s \in \mathcal{S}, \quad \forall k_s \in \mathcal{K}_s. \quad (63)$$

Applying AGMA, C1.4 can be approximated by a convex constraint as

$$\begin{aligned} \widetilde{\text{C1.4}} : & R_s^{\text{rsv}} \times \prod_{n \in \mathcal{N}} \left(\frac{\alpha_{k_s, n}(t_2) L_{i,n}}{\tau_{k_s, n}(t_2)} \right)^{-\tau_{k_s, n}(t_2)} \\ & \times \left(\frac{CY_{k_s}(t_2)}{\phi_{k_s}(t_2)} \right)^{-\phi_{k_s}(t_2)} \leq 1, \quad \forall s \in \mathcal{S}, \quad \forall k_s \in \mathcal{K}_s, \end{aligned} \quad (64)$$

where

$$\tau_{k_s, n}(t_2) = \alpha_{k_s, n}(t_2 - 1) L_{i,n} / v_{k_s}(t_2), \quad (65)$$

$$\phi_{k_s}(t_2) = CY_{k_s}(t_2 - 1) / v_{k_s}(t_2), \quad (66)$$

and

$$v_{k_s}(t_2) = \sum_{n \in \mathcal{N}} \alpha_{k_s, n}(t_2 - 1) L_{i,n} + CY_{k_s}(t_2 - 1). \quad (67)$$

Similarly, we can approximate C1.5 by

$$\begin{aligned} \widetilde{\text{C1.5}} : & Y_{k_s}^2(t_2) \times \prod_{n \in \mathcal{N}} \left(\frac{\alpha_{k_s, n}(t_2) M_{i,n}}{\omega_{k_s, n}(t_2)} \right)^{-\omega_{k_s, n}(t_2)} \leq 1, \\ & \forall s \in \mathcal{S}, \quad \forall k_s \in \mathcal{K}_s, \end{aligned} \quad (68)$$

where

$$\omega_{k_s, n}(t_2) = \frac{\alpha_{k_s, n}(t_2 - 1) M_{i,n}}{\sum_{n \in \mathcal{N}} \alpha_{k_s, n}(t_2 - 1) M_{i,n}}. \quad (69)$$

Then at each iteration t_2 solve,

$$\begin{aligned} \min_{\alpha, Y} \quad & \max_{\substack{\forall s \in \mathcal{S} \\ \forall k_s \in \mathcal{K}_s}} \sum_{n \in \mathcal{N}} \alpha_{k_s, n}(t_2) \beta_{k_s, n} \\ \text{Subject to:} \quad & \widetilde{\text{C1.4}}, \widetilde{\text{C1.5}}, \text{C3}, \end{aligned} \quad (70)$$

which is in GP form and can be solved efficiently with standard convex optimization tools such as CVX [25].

D. COMPLEXITY ANALYSIS

1) CONVERGENCE ANALYSIS

Problems of the form of (58) and (70) are solved using an interior point method in CVX. According to [30], the required number of iterations (Newton steps) to solve by this method is $\log(c/t^0\delta)/\log(\xi)$, where c is the total number of constraints, t^0 is the initial point used by the solver in CVX to apply the interior point method, $0 < \delta \ll 1$ is the stopping criterion, and ξ is used for updating the accuracy of the method.

For (58), the total number of constraints is $c_1 = 3K + 3KN$ and for (70) we have $c_2 = 2K + N$. Therefore, each sub-problem will converge in

$$\begin{cases} \frac{\log(c_1/(t_1^0\delta_1))}{\log(\xi_1)}, & \text{Power (58),} \\ \frac{\log(c_2/(t_2^0\delta_2))}{\log(\xi_2)}, & \text{Sub-carrier (70).} \end{cases} \quad (71)$$

iterations. With a polynomial presentation, the number of iterations (Newton steps) grows as $\mathcal{O}(\sqrt{c})$ [30], [31].

We see that power allocation (58) has higher complexity than the sub-carrier allocation (70) and is more sensitive to K and N . The number of iterations required to achieve convergence is studied further in Section V-C.

2) COMPUTATIONAL COMPLEXITY

For each iteration, a Newton step of an interior-point method costs $\mathcal{O}(c\varrho^2)$ operations where c is the number of constraints and ϱ is the number of variables [31]. For (58), the total number of variables is $\varrho_1 = 3KN + K = \mathcal{O}(KN)$ and for (70) we have $\varrho_2 = KN + K = \mathcal{O}(KN)$. Thus, assuming that the number of sub-carriers, N , is greater than the number of users, K , the total computational complexity of solving the GP problem (i.e., number of Newton steps \times required operations per step) is polynomial, namely,

$$C_{GP} = \begin{cases} \mathcal{O}(c_1^{1.5}\varrho_1^2) = \mathcal{O}(K^{3.5}N^{3.5}), & \text{Power (58),} \\ \mathcal{O}(c_2^{1.5}\varrho_2^2) = \mathcal{O}(K^2N^{3.5}), & \text{Sub-carrier (70).} \end{cases} \quad (72)$$

In addition to computations required to solve the GP problems, additional computational complexity is incurred by applying AGMA approximations. Thus, in Algorithm 1 the worst-case number of computations required to convert to the GP form using AGMA is $i_1 = 2K + 5KN + 3N$ and $i_2 = 8 + 3K + 8KN + N$, for steps 1 and 2, respectively, each of which is $\mathcal{O}(KN)$ and of lower order than C_{GP} . Thus yields an overall complexity per iteration of each of the steps in Algorithm 1 is $C_1 = C_{GP}$. The required number of iterations for Algorithm 1 to converge is studied further in Section V-C.

V. NUMERICAL RESULTS

The performance of the proposed algorithm was evaluated considering a single cell VWN with $N = 16$ sub-carriers which can each be shared by at most $N_{max} = 4$ users, supporting two slices each with $K_s = 4$ users, except where otherwise noted. In all trials we have $P_{max} = 23$ dBm, $R_s^{RSV} = R^{RSV}$, and $\epsilon_s = \epsilon$. The users are placed randomly within the BS coverage area following a uniform distribution and the channels gains are derived according to the Rayleigh fading model with $h_{k_s,n} = \chi_{k_s,n}d_{k_s}^{-\lambda}$ where $\lambda = 3$ is the path loss exponent, d_{k_s} is the distance between user k_s and the BS normalized to the radius of the coverage area, and $\chi_{k_s,n} \sim \text{Exp}(1)$. We evaluate SIC imperfections for several levels of SIC error variance with $\sigma_e^2 \in \{0.01, 0.025, 0.05, 0.10\}$.

For comparison, we present results for non-robust NOMA with both perfect and imperfect SIC, and for OMA. For non-robust NOMA we consider $\|e_{j,n}\|^2 = \sigma_e^2$ in (8), i.e., is deterministic with σ_e^2 representing the imperfect level of achieved cancellation in SIC, and $\sigma_e^2 = 0$ represents perfect SIC. For OMA, we set $N_{max} = 1$ in our algorithm to enforce orthogonality between the sub-carriers, eliminating inter-user interference. In all formulations, where no feasible solutions exists for a given channel realization, user power is set to P_{max} and all users are considered to be in outage.

A. PERFORMANCE ANALYSIS

To visualize the overall relationship between probability of outage, reserved rate, and required transmit power, Fig. 1 shows the average transmit power per user versus R^{RSV} and ϵ for $\sigma_e^2 = 0.01$ and 0.10 .

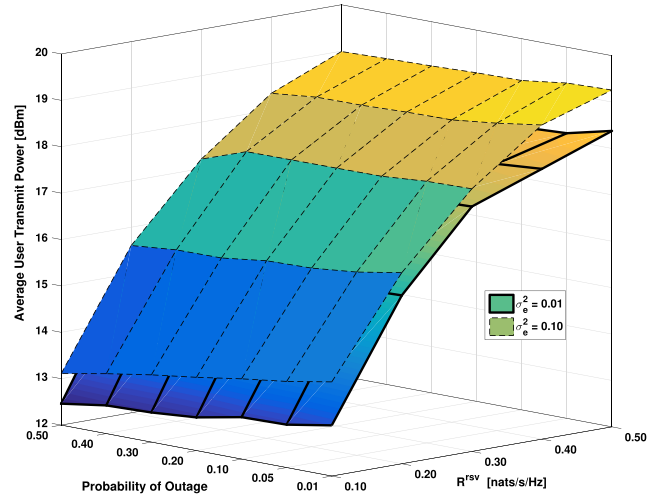


FIGURE 1. Average UE transmit power versus R^{RSV} and ϵ .

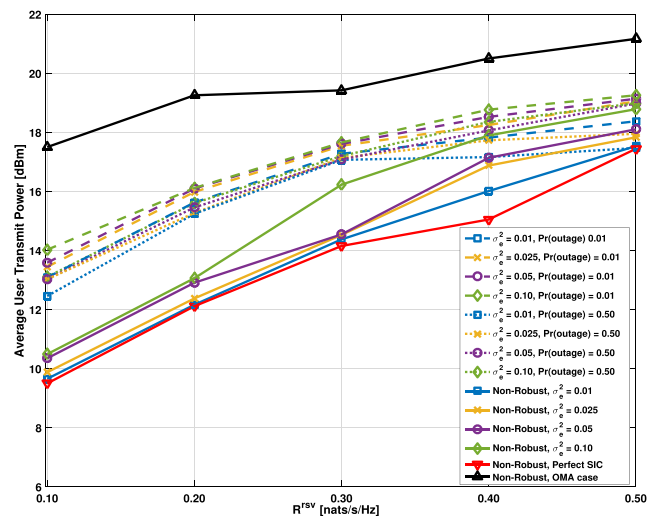


FIGURE 2. Average UE transmit power versus R^{RSV} .

Only the highest and lowest values of SIC error variance are shown in the figure for clarity. As expected, power increases with increased rate reservations and levels of SIC error and decreases with increasing probability of outage. For any specific value of ϵ , power increases sharply for increasing user rate reservations due to the decreased feasibility, and is always higher for increased SIC error. For any specific user rate reservation, lower power is required for less stringent user outage constraints as this increases the flexibility in finding a feasible solution which will provide the required maximum outage protection. The relationships over two of the three axes are depicted in the subsequent figures for specific cases.

Average transmit power versus R^{RSV} is plotted in Fig. 2 for some cases of user outage constraint, ϵ . As expected, average required transmit power increases proportionally to rate reservations, R^{RSV} , and SIC error variance, σ_e^2 , because users must transmit with higher power to achieve their

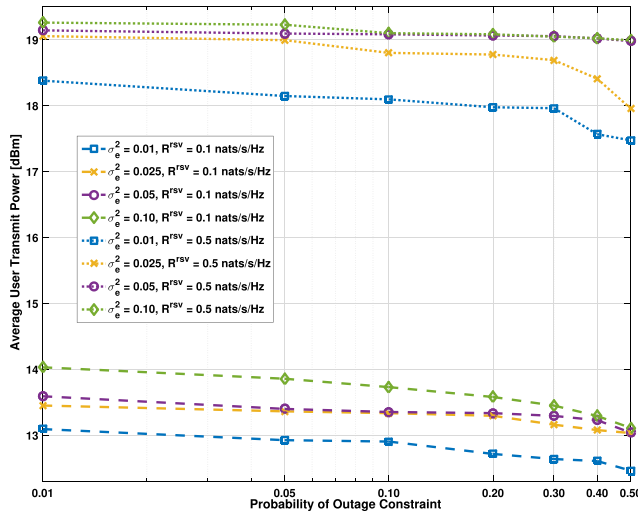


FIGURE 3. Average UE transmit power versus ϵ

desired rates. Power levels for all values of σ_e^2 are shown for $\epsilon = 0.01$ and 0.50 . There is a power trade-off for robustness and the results for non-robust power optimization are presented for comparison. For example, with $R^{\text{RSV}} = 0.2$ nats/s/Hz and $\epsilon = 0.5$, i.e., power minimization allowing up to 50% probability of user outage, 3.08 dB, 2.89 dB, 2.60 dB, and 2.55 dB higher power is allocated by the robust optimization versus non-robust, for $\sigma_e^2 = 0.01, 0.025, 0.01$, and 0.10 , respectively, and 3.28 dB higher than the perfect SIC case. However, for both robust and non-robust cases, NOMA outperforms OMA based on the inability to multiplex users on strong channels and the requirement to use weaker channels to maintain orthogonality between users and meet rate reservations. For the robust case and $R^{\text{RSV}} = 0.2$ nats/s/Hz and $\epsilon = 0.01$, OMA requires 3.63 dB, 3.28 dB, 3.16 dB, and 3.14 dB, higher power compared to NOMA, with $\sigma_e^2 = 0.01, 0.025, 0.05$, and 0.10 , respectively.

For the robust optimization, average required transmit power increases with increasing SIC error variance and rate reservations but decreases with less strict outage constraints. This can be seen clearly in Fig. 3, which plots average transmit power versus probability of outage, ϵ , for $R^{\text{RSV}} = 0.1$ nats/s/Hz and 0.5 nats/s/Hz. From the figure, we note that as the outage constraint is loosened, average power decreases but for higher levels of SIC error variance and R^{RSV} for $\epsilon \geq 0.10$ the plots converge as more users can be pushed into outage in favour of reduced user transmit power. For example, for $R^{\text{RSV}} = 0.1$ nats/s/Hz and $\epsilon = 0.10$ versus 0.01 , required transmit power decreases by 0.19 dB, 0.12 dB, 0.24 dB, and 0.30 dB, for $\sigma_e^2 = 0.01, 0.025, 0.05$, and 0.10 , respectively. For $R^{\text{RSV}} = 0.5$ nats/s/Hz and $\epsilon = 0.50$ versus 0.01 , required transmit power decreases by 0.91 dB, 0.42 dB, 0.55 dB, and 0.92 dB, for $\sigma_e^2 = 0.01, 0.025, 0.05$, and 0.10 , respectively.

Experienced user outage versus probability of outage constraint, ϵ , is plotted in Fig. 4. For clarity, only the two extreme values for SIC error variance are shown. For $\sigma_e^2 = 0.01$ and

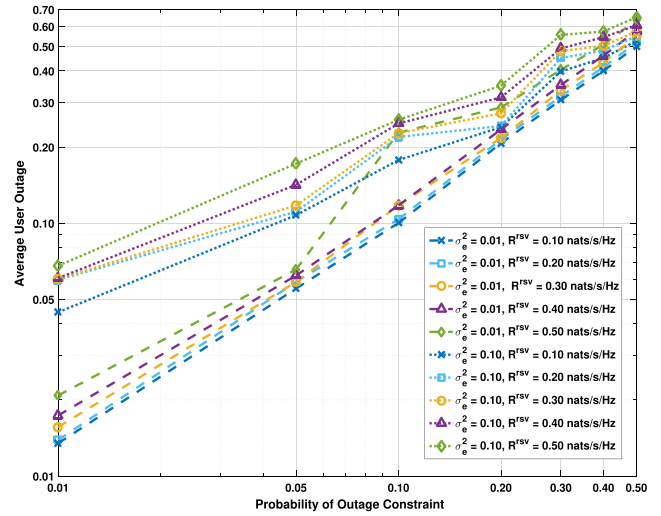


FIGURE 4. Average user outage versus ϵ .

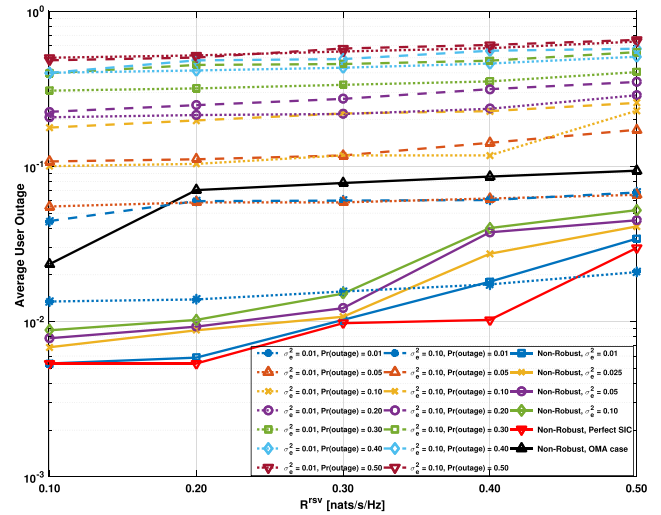


FIGURE 5. Average user outage versus R^{RSV} .

$R^{\text{RSV}} = 0.1$ nats/s/Hz, we see that the experienced outage is very close to the probability used to constrain the problem, with correlation coefficient $R = 0.9997$ and root mean square error (RMSE), as compared to the ideal values of experienced outage equal to ϵ , of 0.00322. As R^{RSV} is increased, the resulting outage increasingly differs from the constraint due to decreased feasibility of solutions but for $\sigma_e^2 = 0.01$ the worst-case correlation is for $R^{\text{RSV}} = 0.5$ nats/s/Hz at $R = 0.9771$ and RMSE of 0.0306. For higher levels of SIC error variance, feasibility of solutions can be significantly impacted over the trials conducted and we see correlation coefficients of $R = 0.9606$ and 0.7276 and RMSE of 0.0402 and 0.106, for $\sigma_e^2 = 0.10$ and $R^{\text{RSV}} = 0.1$ and 0.5 nats/s/Hz, respectively.

Experienced user outage versus rate reservation, R^{RSV} , is plotted in Fig. 5. The experienced user outage aligns very well to the probability of outage constraint, ϵ , with reduced

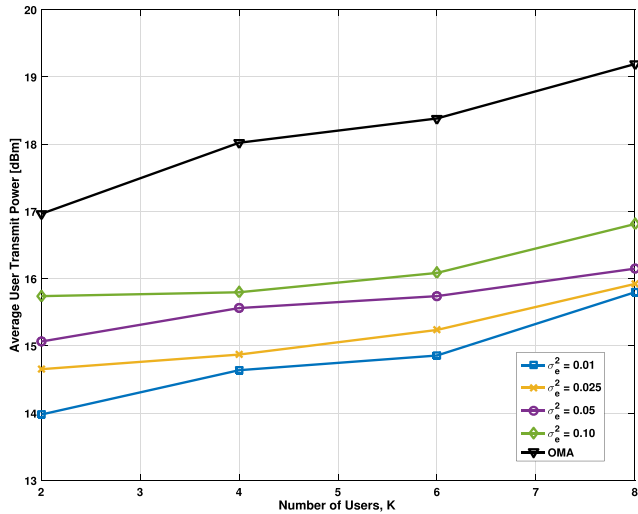


FIGURE 6. Average power versus number of users, K.

alignment as either R^{rsv} or σ_e^2 is increased, as was also seen in Fig. 4. Further, results for OMA and non-robust NOMA are presented where outage is not constrained but when no feasible solution exists for a given channel realization all users are considered to be in outage. For the OMA case, at $R^{\text{rsv}} \geq 0.20$ nats/s/Hz experienced outage is higher than the robust cases with $\epsilon = 0.01, 0.05$ and higher than all non-robust NOMA cases. Non-robust NOMA with SIC error experiences higher levels outage as the value of σ_e^2 increases but does not exceed 5.25% even for the worst case level of error.

B. SYSTEM DENSITY ANALYSIS

The ratio of users to available sub-carriers will reduce the flexibility of the system in both power and sub-carriers allocation. Under OMA, the maximum number of users which can be supported is equal to the number of available sub-carriers, and only then if a suitably strong sub-carrier is available for each user to meet their rate reservation. Under NOMA, stronger sub-carriers can be leveraged by several users in order to minimize required transmit power to meet reservations, but with increased sub-carrier sharing average transmit power will necessarily increase, especially for higher levels of SIC error variance.

For a fixed number of sub-carriers, $N = 16$, Fig. 6 plots average user transmit power as the total number of users in the system, i.e., system density, is increased, for $R^{\text{rsv}} = 0.1$ nats/s/Hz and $\epsilon = 0.01$. With more users utilizing the same number of sub-carriers, required transmit power increases at all levels of SIC error variance as sub-carrier sharing results in increased inter-user interference. For low number of users, i.e., low system density, the user rate reservations can easily be met. As the density is increased, the required power also increases but at faster rate under OMA than NOMA. For example, as system density increases for $K = 2$ to 8, transmit power under OMA increases by 2.23 dB due to

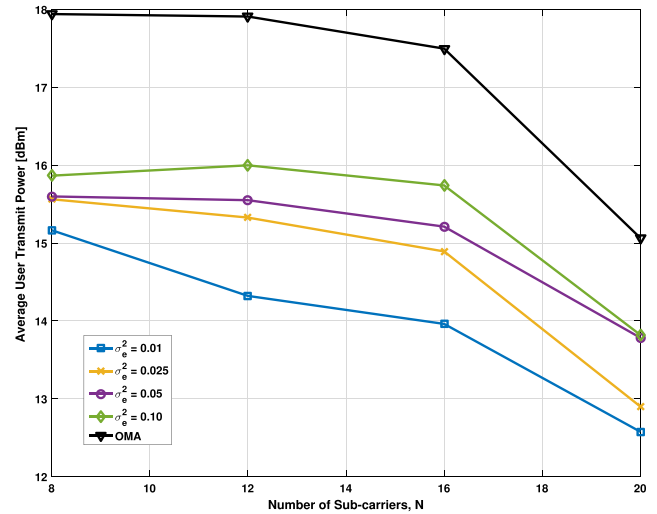


FIGURE 7. Average power versus number of sub-carriers, N.

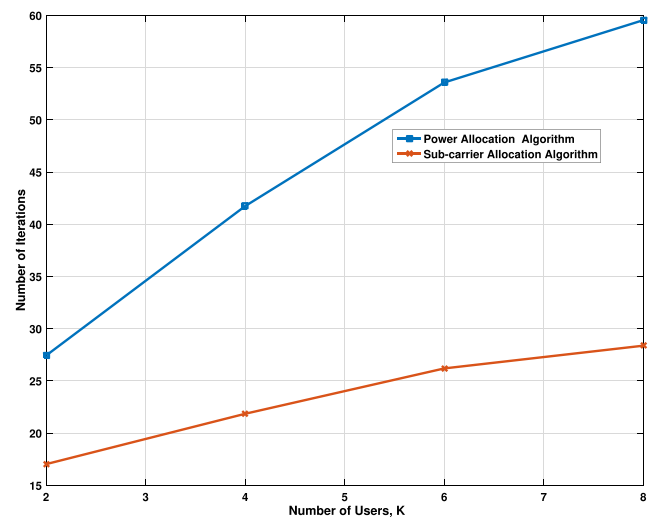


FIGURE 8. Iterations versus number of users, K.

reduced flexibility in avoiding weak channels. The increase in required power under NOMA was only 1.82 dB, 1.27 dB, 1.08 dB, and 1.07 dB, for $\sigma_e^2 = 0.01, 0.025, 0.05,$ and 0.10, respectively. In all cases, NOMA outperforms the OMA case. For example, with $K = 6$, required transmit power under NOMA is decreased compared to OMA by 3.52 dB, 3.14 dB, 2.64 dB, and 2.29 dB, for $\sigma_e^2 = 0.01, 0.025, 0.05,$ and 0.10, respectively.

For a fixed number of users, $K = 4$, Fig. 7 plots average user transmit power as the total number of sub-carriers in the system is increased, for $R^{\text{rsv}} = 0.1$ nats/s/Hz and $\epsilon = 0.01$. With an increasing number of sub-carriers, i.e., decreasing system density, required transmit power decreases for both OMA and NOMA, with reduced requirement to utilize weaker channels or multiplex and experience inter-user interference, but NOMA still outperforms OMA in all cases. For example, with $N = 12$, required transmit power under NOMA is decreased compared to OMA by 3.59 dB,

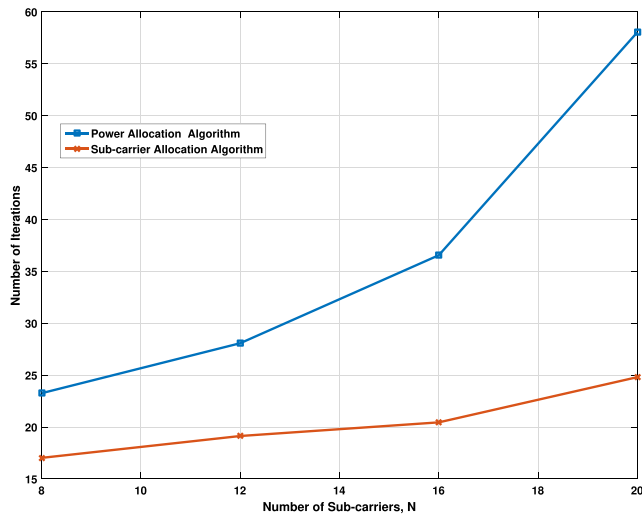


FIGURE 9. Iterations versus number of sub-carriers, N.

2.58 dB, 2.36 dB, and 1.91 dB, for $\sigma_e^2 = 0.01, 0.025, 0.05,$ and 0.10 , respectively.

C. CONVERGENCE ANALYSIS

The convergence and complexity of the proposed algorithms was studied in Section IV-D and the number of iterations for convergence for power allocation (58) and sub-carrier allocation (70) problems as a function of the number of users, K , and sub-carriers, N , are plotted in Fig. 8 and 9, respectively, for $R^{\text{TSV}} = 0.1$ nats/s/Hz and $\epsilon = 0.01$. As K and N increase the required number of iterations for each problem to converge increases. As expected, the power allocation problem is more sensitive to changes in both K and N . The analysis in Section IV-D found that the solution of the power allocation problem to be of higher complexity than the sub-carrier allocation which is confirmed in the required number of iterations depicted in Figs. 8 and 9.

VI. CONCLUSIONS

To support the changing use cases for future networks, significant improvement in system efficiencies and traffic density are required. For systems of high priority low power devices supporting critical applications, such as health and public safety monitoring, dense deployments can be expected, power is limited, and user outage is equivalent to system failure. NOMA can enable meeting these demands by allowing users to share sub-carriers via power-domain multiplexing. In this work, we have investigated robust resource allocation in UL NOMA systems subject to residual cancellation errors from imperfect SIC. We evaluated a VWN with slices comprised of low power devices serving critical applications, for which minimum achieved rate and maximum user outage must be maintained. With the goal of maximizing battery life for such devices, we first derived the probability of outage as a function of SIC error variance and then used this result to formulate a robust resource allocation problem minimizing

transmit power subject to slice and system constraints. The proposed iterative algorithm to solve the resulting non-convex and computationally intractable optimization is based on SCA and uses CGP and AGMA to transform into a convex form which can be solved efficiently at each iteration. Simulation results show the expected trade-off for robustness in terms of higher average transmit power compared to a non-robust approach. Despite this trade-off, the proposed algorithm outperforms the corresponding OMA system in terms of average user transmit power and overall system density due to the multiplexing gains available in NOMA systems.

REFERENCES

- [1] *IMT Vision—Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond*, document Rec. M.2083, Sep. 2015.
- [2] H. Tabassum, M. S. Ali, E. Hossain, M. J. Hossain, and D. I. Kim. (2016). “Non-orthogonal multiple access (NOMA) in cellular uplink and downlink: Challenges and enabling techniques.” [Online]. Available: <https://arxiv.org/abs/1608.05783>
- [3] S. M. R. Islam, N. Avazov, O. A. Dobro, and K.-S. Kwak, “Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges,” *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 721–742, 2nd Quart., 2017.
- [4] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, “Non-orthogonal multiple access (NOMA) for cellular future radio access,” in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Jun. 2013, pp. 1–5.
- [5] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, “On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users,” *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.
- [6] *Study on Downlink Multiuser Superposition Transmission (MUST) for LTE*, document TR 36.859 Version 13.1.0 Release 13, 3GPP, 2015.
- [7] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, “Optimal joint power and subcarrier allocation for MC-NOMA systems,” in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.
- [8] A. Mokdad, P. Azmi, and N. Mokari, “Radio resource allocation for heterogeneous traffic in GFDM-NOMA heterogeneous cellular networks,” *IET Commun.*, vol. 10, no. 12, pp. 1444–1455, 2016.
- [9] L. Lei, D. Yuan, and P. Värbrand, “On power minimization for non-orthogonal multiple access (NOMA),” *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2458–2461, Dec. 2016.
- [10] R. Dawadi, S. Parsaefard, M. Derakhshani, and T. Le-Ngoc, “Power-efficient resource allocation in NOMA virtualized wireless networks,” in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.
- [11] H. Hacı, H. Zhu, and J. Wang, “Performance of non-orthogonal multiple access with a novel asynchronous interference cancellation technique,” *IEEE Trans. Wireless Commun.*, vol. 65, no. 3, pp. 1319–1335, Mar. 2017.
- [12] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, “A general power allocation scheme to guarantee quality of service in downlink and uplink NOMA systems,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7244–7257, Nov. 2016.
- [13] N. Zhang, J. Wang, G. Kang, and Y. Liu, “Uplink nonorthogonal multiple access in 5G systems,” *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 458–461, Mar. 2016.
- [14] Z. Yang, W. Xu, H. Xu, J. Shi, and M. Chen, “Energy efficient non-orthogonal multiple access for machine-to-machine communications,” *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 817–820, Apr. 2017.
- [15] H. Tabassum, E. Hossain, and J. Hossain, “Modeling and analysis of uplink non-orthogonal multiple access in large-scale cellular networks using Poisson cluster processes,” *IEEE Trans. Commun.*, vol. 65, no. 8, pp. 3555–3570, Aug. 2017.
- [16] S. Shi, L. Yang, and H. Zhu, “Outage balancing in downlink nonorthogonal multiple access with statistical channel state information,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4718–4731, Jul. 2016.
- [17] P. Xu, Y. Yuan, Z. Ding, X. Dai, and R. Schober, “On the outage performance of non-orthogonal multiple access with 1-bit feedback,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6716–6730, Oct. 2016.
- [18] J. Cui, Z. Ding, and P. Fan, “A novel power allocation scheme under outage constraints in NOMA systems,” *IEEE Signal Process. Lett.*, vol. 23, no. 9, pp. 1226–1230, Sep. 2016.

- [19] A. Mokdad, M. Moltafet, P. Azmi, and N. Mokari, "Robust radio resource allocation for heterogeneous traffic in PD-NOMA-based cellular systems," in *Proc. Iranian Conf. Elect. Eng. (ICEE)*, May 2017, pp. 1796–1801.
- [20] Z. Feng, C. Qiu, Z. Feng, Z. Wei, W. Li, and P. Zhang, "An effective approach to 5G: Wireless network virtualization," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 53–59, Dec. 2015.
- [21] C. Liang and F. R. Yu, "Wireless network virtualization: A survey, some research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 358–380, 1st Quart., 2015.
- [22] A. Nemirovski and A. Shapiro, "Convex approximations of chance constrained programs," *SIAM J. Optim.*, vol. 17, no. 4, pp. 969–996, 2007.
- [23] S. Parsaeefard, A. R. Sharafat, and N. Mokari, *Robust Resource Allocation in Future Wireless Networks*. New York, NY, USA: Springer, 2017.
- [24] D. Bertsimas and M. Sim, "The price of robustness," *Oper. Res.*, vol. 52, no. 1, pp. 35–53, Feb. 2004.
- [25] M. Grant and S. Boyd. (Mar. 2017). *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.1*. [Online]. Available: <http://cvxr.com/cvx>
- [26] M. Chiang, "Geometric programming for communication systems," *Commun. Inf. Theory*, vol. 2, nos. 1–2, pp. 1–154, Jul. 2005.
- [27] M. Avriel and A. C. Williams, "Complementary geometric programming," *SIAM J. Appl. Math.*, vol. 19, no. 1, pp. 125–141, Jul. 1970.
- [28] M. Chiang, C. W. Tan, D. P. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.
- [29] S. Parsaeefard, R. Dawadi, M. Derakhshani, and T. Le-Ngoc, "Joint user-association and resource-allocation in virtualized wireless networks," *IEEE Access*, vol. 4, pp. 2738–2750, 2016.
- [30] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
- [31] A. Nemirovski. *Interior Point Polynomial Time Methods in Convex Programming*. [Online]. Available: http://www2.isye.gatech.edu/~nemirovs/Lect_IPM.pdf



SAEEDAH PARSAEEFARD (S'09–M'14) received the B.Sc. and M.Sc. degrees from the Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, in 2003 and 2006, respectively, and the Ph.D. degree in electrical and computer engineering from Tarbiat Modares University, Tehran, Iran, in 2012. She was a Post-Doctoral Research Fellow with the Telecommunication and Signal Processing Laboratory, Department of Electrical and Computer Engineering, McGill University, Montréal, Canada. From 2010 to 2011, she was a Visiting Ph.D. Student with the Department of Electrical Engineering, University of California at Los Angeles, Los Angeles, USA. She is currently an invited Lecturer (Assistant Professor) with the Computer Engineering Department, Amirkabir University of Technology (Tehran Polytechnic) and a Faculty Member with the Iran Telecommunication Research Center. Her current research interests include resource management in software-defined networking, Internet-of-Things and fifth generation of wireless networks, and applications of robust optimization theory and game theory in resource allocation and management in wireless networks.



DANIEL TWEED (S'13) received the B.Sc. degree (Hons.) in computer engineering from the University of Manitoba, Winnipeg, Canada, in 2015. He is currently pursuing the M.Eng. degree with the Department of Electrical and Computer Engineering, McGill University, Montréal, Canada. His research interests include multi-radio access technology coexistence, power-domain and hybrid non-orthogonal multiple access, radio resource management for wireless networks, applications of optimization theory to wireless networks, Internet-of-Things and machine-to-machine communication for future networks, and software-defined wireless network architectures and resource allocation techniques.



MAHSA DERAKHSHANI (S'10–M'13) received the B.Sc. and M.Sc. degrees from the Sharif University of Technology, Tehran, Iran, in 2006 and 2008, respectively, and the Ph.D. degree from McGill University, Montréal, Canada, in 2013, all in electrical engineering. From 2013 to 2015, she was a Post-Doctoral Research Fellow with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada, and a Research Assistant with the Department of Electrical and Computer Engineering, McGill University. From 2015 to 2016, she was an Honorary NSERC Post-Doctoral Fellow with the Department of Electrical and Electronic Engineering, Imperial College London. She is currently a Lecturer (Assistant Professor) in digital communications with the Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University. Her research interests include radio resource management for wireless networks, software-defined wireless networking, applications of convex optimization and game theory for communication systems, and spectrum sensing techniques in cognitive radio networks. Dr. Derakhshani received the John Bonsall Porter Prize, the McGill Engineering Doctoral Award, the Fonds de Recherche du Québec-Nature et Technologies and Natural Sciences and Engineering Research Council of Canada Post-Doctoral Fellowships.



THO LE-NGOC (F'97) received the B.Eng. degree (Hons.) in electrical engineering and the M.Eng. degree from McGill University, Montréal, QC, Canada, in 1976 and 1978, respectively, and the Ph.D. degree in digital communications from the University of Ottawa, Canada, in 1983. From 1977 to 1982, he was a Research and Development Senior Engineer with Spar Aerospace Ltd., Sainte-Anne-de-Bellevue, Canada, where he was involved in the development and design of satellite communications systems. From 1982 to 1985, he was the Engineering Manager with the Radio Group, Department of Development Engineering, SRTelecom Inc., Saint-Laurent, QC, Canada, where he developed the new point-to-multipoint DA-TDMA/TDM Subscriber Radio System SR500. From 1985 to 2000, he was a Professor with the Department of Electrical and Computer Engineering, Concordia University, Montréal, Canada. Since 2000, he has been with the Department of Electrical and Computer Engineering, McGill University. His research interest is in the area of broadband digital communications. He is a Fellow of the Engineering Institute of Canada, the Canadian Academy of Engineering, and the Royal Society of Canada. He was a recipient of the 2004 Canadian Award in telecommunications research, and the IEEE Canada Fessenden Award 2005. He is the Canada Research Chair (Tier I) on Broadband Access Communications.

...