Spatial Transmitter Density Allocation for Frequency-Selective Wireless Ad Hoc Networks

Songtao Lu, *Member, IEEE*, and Zhengdao Wang, *Fellow, IEEE*

***Abstract*— We consider a network of pairs of nodes that perform ad hoc simultaneous communications over frequency-selective channels. We assume that the whole frequency band is divided into a number of subbands, and each transmitter can only use one subband. Assuming that the network is geometrically infinite, we use the transmission capacity (TC) as a measure of the network throughput. We consider the problem of allocating nodes to the subbands so that the total TC is maximized, under the constraint of a fixed total spatial node density. The optimization problem turns out to be nonconvex. We investigate the detailed structure of the functions involved in the optimization and identify a set of properties of the optimal transmitter density over the subbands. An iterative resource allocation scheme with low complexity is derived to obtain the global optimal solution of the TC maximization problem. The solution can be loosely interpreted as a water-filling solution for a nonconvex optimization problem. Based on numerical simulations, it is shown that the optimal solution obtained through the theoretical analysis is consistent with the one obtained through an exhaustive search, which reveals that the outage probability and the total spatial transmitter density are the keys to determining the network TC.**

***Index Terms*— Transmission capacity, frequency-selective networks, nonconvex optimization, Lambert function.**

I. Introduction

TRANSMISSION capacity (TC) is a useful metric to measure the throughput of wireless communication networks. It is defined as the number of successful transmissions per unit area under a target outage probability [2], [3]. Stochastic geometry is often employed in TC analysis to model the spatial distribution of communication nodes and their connections. For example, the limits of successful transmission data rates with increasing number of users and number of base stations (BS) in a cellular network have been quantified in the problem of wireless network densification [4]. The transmission throughput can be improved by optimizing system parameters. Recently, stochastic geometry is also employed in some emerging applications of wireless communications and the effectiveness of the methodology has been demonstrated with both numerical and real experiments for various wireless networks, such as underwater networks [5], general cellular networks [6], [7], vehicle-to-vehicle (V2V) safety communications [8], energy harvesting device-to-device (D2D) communications [9], millimeter wave (mm-wave) ad hoc networks [10], [11], etc.

For frequency-flat networks, the TC framework [6] can provide a quantitative network-level performance analysis for spectrum management. The tradeoff between bandwidth and signal-to-interference plus noise ratio (SINR) in the spatial random ad hoc network was studied in [12], where the TC was maximized with respect to the spatial transmitter density under a fixed rate requirement. The coverage probability maximization problem in downlink cellular networks was proposed in [13] and solved with respect to optimizing transmit power, BS density and transmit power density. In heterogeneous wireless networks, the sum of the individual TC over multiple bands for D2D communications was optimized in terms of the user density in flat fading channels [14], where the TC maximization problem was a convex problem within a certain range of the outage probability and therefore solved easily. Layered transmission strategies were proposed in [15], where there were two tiers of BSs considered: macro-tier and femto-tier. Each type of the BSs was modeled as a homogeneous Poisson point process (PPP) independently and occupied a number of subbands over the whole transmission spectrum either deterministically or randomly. Then, the outage probability versus the proportion of the frequency resource allocated between two tiers was analyzed numerically. Also, the optimal allocation of the transmitter density in spectrum sharing systems was proposed in [2] to maximize the TC over the shared spectrum. In general, there is already a body of results in the literature about the performance analysis and resource allocation in terms of system parameters (e.g., transmission power, transmitter/user density, target outage probability, retransmission rate, etc.) for homogeneous and heterogeneous networks; see e.g., [6], [15]–[18], etc.

For frequency-selective networks, the channel outage probability varies with frequency. Such outage behavior in underwater networks was discussed from a physical layer perspective in [19] and [20], which takes both the transmission burstiness and path loss properties into account. Furthermore, due to

Manuscript received February 19, 2018; revised September 23, 2018; accepted November 4, 2018. Date of publication November 27, 2018; date of current version January 8, 2019. This work was supported in part by NSF under Grant 1711922. This paper was presented in part at the IEEE Wireless Communications and Networking Conference, New Orleans, LA, USA, 9–12 Mar. 2015 [1]. The associate editor coordinating the review of this paper and approving it for publication was H. Pishro-Nik. *(Corresponding author: Zhengdao Wang.)*

S. Lu is with the Department of Electrical and Computer Engineering, University of Minnesota Twin Cities, Minneapolis, MN 55455 USA (e-mail: lus@umn.edu).

Z. Wang is with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA (e-mail: zhengdao@iastate.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2018.2881707

the effect of long time delay and busy terminals, the packets collision probability in underwater communication channel is very high, which was verified by both experiments and simulations [21]. Spectrum sharing schemes under frequency-selective channels were also examined in [22], where several independent mobile networks share and/or partially share the same spectrum resource. In this work, the effect of the frequency-selective fading on the signal-to-interference ratio was investigated. With the analytical outage probability, the network throughput was studied in [5] for underwater communications, where the theoretical model was verified with a ray tracing simulator. Leveraging resources, e.g., spatial-reuse, with some elaborated protocols is one of the promising ways of improving the network throughput, such as underwater acoustic networks, where the field experiment has been conducted and the results are shown in [23].

Transmitting packets over multiple frequencies for frequency-selective channels is a typical practice. Maximizing the spectral utilization for such transmissions is an important optimization problem. Insight on the TC maximization over frequency-selective communication networks provides guidance for designing practical scheduling protocols, especially under a constraint of a total spatial density of transmitters. Unfortunately, there are few works on the tradeoff between the spatial density of transmitters and outage probability for frequency-selective pathloss. In [5], the optimal allocation of the transmitter density for each frequency band was given for maximizing the throughput of an underwater communication work without the outage probability consideration. Currently, there is no analytical result on the optimal transmitter density allocation across multiple frequency subbands, for a given total spatial transmitter density over multiple frequencies and the outage probability constraint for each subband.

A. Contributions

In this paper, we focus on the performance analysis of a wireless network over frequency-selective channels, quantifying the TC limit of the network with partitioning the spatial node density into different frequency subbands. Given the spatial transmitter density and outage probability constraint, the network throughput is maximized by allocating the transmitter density in each subband. The problem is formulated as an optimization problem which turns out to be nonconvex. According to Karush-Kuhn-Tucker (KKT) conditions, the local optimal points are obtained in terms of the Lambert function. Through studying the properties of the solutions, it is interesting that most of local optimal points are excluded such that the complexity of finding the global optimal solution is reduced significantly. There is a line of work that focuses on solving the nonlinear sum-of-ratios problems. The proposed problem can be also solved by some iterative algorithms with guaranteed convergence to the global optimal solution. For example, an algorithm proposed in [24] is based on solving a sequence of convex programming problems, shorten as the SCP algorithm in this paper. The optimality of the solution obtained through our proposed algorithm is also verified with numerical examples and compared with the SCP algorithm.

This paper gives a solution for the maximization of the network TC, in the sense as defined in [3], under a total transmitter density constraint in wideband networks where the constraint can be from zero to infinity (i.e., the density of the network is from sparse to extremely dense). The main contributions of the work are summarized as follows:

1) Given the total transmitter density in a network, the transmission strategy over multiple subbands is proposed where the network TC is defined and the network TC maximization is formulated as a nonconvex optimization problem.
2) Based on the convexity property of the objective function and properties of the KKT conditions, the complexity of finding the global optimal solution for the network TC maximization is significantly reduced, where only one-dimensional search in terms of the Lambert function is involved.
3) Through the theoretical analysis, the solution of solving the TC maximization problem is provided (in almost closed-form), which optimally allocates the transmitter density to frequency bins under the constraints of the total transmitter density and outage probability.

In addition, we reveal some properties of the Lambert function (we will see them later in Lemma 2 and Lemma 3) which are used in the proof of the TC maximization problem. To the best of our knowledge, these properties have not been identified before.

B. Organization

The rest of the paper is organized as follows. A system model is developed in Section II. The network TC is discussed in Section III. The optimal solutions based on the KKT conditions are given in Section IV. The properties of the optimal solutions are analyzed further in Section V where the challenges of the problems are also introduced. Section VI studies how to reduce the complexity of finding the global optimal solution. The complete solutions of this optimization problem are presented in Section VII. Numerical simulation results are reported in Section VIII and finally conclusions are drawn in Section IX.

II. System Model

We assume that the transmitters in a wireless network are spatially distributed according to a homogeneous PPP with density λ in an infinite space, denoted as $\Pi(\lambda) = \{\boldsymbol{X}_t\}$, where $\boldsymbol{X}_t \in \mathbb{R}^2$ represents the location of the transmitter t. Each transmitter has its corresponding desired receiver at a fixed distance R away. We assume that the signaling overhead between the transmitter and receiver is small and can be exchanged perfectly. The transmitters send the packets with equal power P independently and there is no coordination involved.

The channel characteristics are modeled in two parts, path loss and fading. The channel path loss in frequency-selective channels is denoted as $h(R, f_o)$ at frequency f_o. The channel response of small fading from the transmitter to the receiver, i.e., H, has a complex Gaussian distribution with zero mean

and unit variance. Since the network is infinitely large, we let the reference receiver be located at the origin. The received SINR of this receiver at frequency f_o is

$$\mathsf{SINR}(f_o) = \frac{\frac{P|H|^2}{h(R,f_o)}}{N_0(f_o) + I(f_o)} \tag{1}$$

where $N_0(f_o)$ denotes the power spectral density (PSD) of noise, and $I(f_o)$ is the aggregate interference PSD received at the origin, i.e.,

$$I(f_o) = \sum_{\boldsymbol{X}_t \in \Pi(\lambda)\setminus X_0} \frac{P|H_t|^2}{h(\|\boldsymbol{X}_t\|, f_o)} \tag{2}$$

where $\|\boldsymbol{X}_t\|$ represents the distance from $\boldsymbol{X}_t$ to the origin, H_t denotes the corresponding channel fading, and $\boldsymbol{X}_0$ is the position of the reference transmitter. Define the successful transmission probability (or coverage probability) P_s as the probability that the received SINR is above the predefined threshold β, which is written as [19, eq. (17)] (see also [25, eq. (9)])

$$P_s\big(\mathsf{SINR}(f_o) > \beta\big) = \exp\left(-\frac{\beta h(R,f_o) N_0(f_o)\Delta f}{P}\right) \mathcal{L}_I(\beta h(R,f_o)) \tag{3}$$

where $\mathcal{L}_I$ is the Laplace transform of the probability density function (PDF) of the interference, Δf denotes the bandwidth of each subband. For a given β, we have

$$P_s(\lambda) = \exp\{-V(R,f_o)\lambda - U(R,f_o)\} \tag{4}$$

where $U(R,f_o) = \beta h(R,f_o) N_0(f_o)\Delta f/P$ and $V(R,f_o)$ can be obtained based on which the path loss model is used. For example, the expressions of $V(R,f_o)$ in frequency-selective underwater communication scenarios and mm-wave channels have been shown at [19, eq. (9)] (see also [5, eq. (5)]) and [26, eq. (13)]. As a result, our proposed transmitter density optimization is relevant to performance improvement in ad hoc networks operating in these channel conditions. Such models have been considered in other resource allocation problems in [2], [13], and [14].

III. Problem Formulation

TC is defined by the number of successful transmissions per unit area under a target outage probability ϵ and expressed by

$$\tau(\lambda) = b\lambda P_s(\lambda), \quad p_o(\lambda) \le \epsilon \tag{5}$$

where $\epsilon \in [0,1]$ denotes the target probability, $p_o(\lambda) = 1 - P_s(\lambda)$ represents the outage probability, $b = B\log_2(1+\beta)$ is the supportable ergodic rate (see, e.g., [5], [27]) and B is the bandwidth. There is a tradeoff between the density of transmitters and the individual link quality. Given β, we know that b is a constant which is ignored in the rest of this paper.

For the wideband communication networks, the total bandwidth is divided into N subbands, where the bandwidth of each subband is small enough such that the communication can be considered as narrowband transmission, such as in underwater communication networks [5]. We define the set $\mathcal{N} = \{1,\ldots,N\}$. The carrier frequency of the ith subband is denoted as $f_{o,i}, i \in \mathcal{N}$. For each subband, the transmitters and receivers are separated by a fixed distance $R_i, i \in \mathcal{N}$ and use the same transmit power due to fairness. The nodes are also distributed according to a PPP, which is denoted as $\Pi(\lambda_i)$. Let λ_i represent the transmitter density at the ith subband. We also assume that each transmitter only occupies one subband, which is valid in practical applications, such as in frequency-selective orthogonal frequency division multiple access (OFDMA) systems [28]. Multiple users are assigned in different subbands such that the best beam pattern at different subcarriers can be adopted by each user.

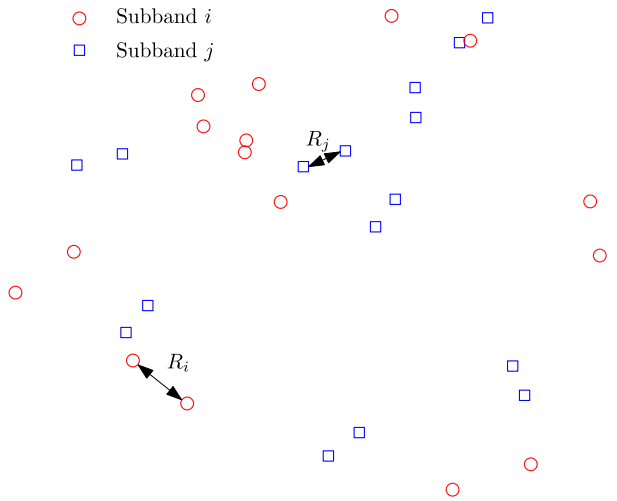

Fig. 1. Example that the system i occupies frequency $f_{o,i}$ and the system j uses frequency $f_{o,j}$.

We remark that the frequency-selective model considered in this work is referred to as ultrawide band channels in [29], since the path loss at each subchannel is dependent on subband frequency [30], [31]. Applications of this model include underwater communications and hybrid multiple-regulatory-bands radio systems. For typical wireless communication systems, where the carrier frequency is much larger than the bandwidth, the path loss within the communication band is essentially constant. For such cases, transmitter density allocation does not offer gains compared to uniform allocation. However, when the multiple subbands operate over different distances R_i (or transmit power), e.g., in multi-tier networks [32], the path loss factor can still be different, rendering the transmitter density allocation consideration valid.

Thus, for each R_i there is a corresponding λ_i at the ith subband, which is defined as system i, where $\Pi(\lambda_i) \cap \Pi(\lambda_j) = \emptyset, i \ne j, i, j \in \mathcal{N}$. In Fig. 1, we show an example there are only two systems i and j which are distinguished by red circles and blue squares.

A. Network Transmission Capacity

Assume that the outage probabilities over all subbands are the same. The network TC is the sum of the individual throughput across all subbands under the target outage probability $\epsilon \in [0,1]$ on each subband, which is defined by

$$f(\lambda_1,\ldots,\lambda_N) = \sum_{i=1}^{N} \tau_i(\lambda_i), \quad p_{o,i}(\lambda_i) \le \epsilon, \ \forall i \tag{6}$$

where the throughput at the ith subband is $\tau_i(\lambda_i) = \lambda_i p_{s,i}(\lambda_i)$, the corresponding outage probability is $p_{o,i}(\lambda_i) = 1 - p_{s,i}(\lambda_i)$ and the successful transmission probability is

$$p_{s,i}(\lambda_i) = \exp\left(-V(R_i, f_{o,i})\lambda_i - U(R_i, f_{o,i})\right). \quad (7)$$

If the impact of the transmitter density over multiple subbands dominates, the interference limited regime is considered, i.e., the case when $\frac{P}{N_0(f_o)} \to \infty$. In this situation, the noise part, i.e., $U(R_i, a(f_{o,i}))$, is negligible such as in 5G mm-wave cellular networks [33], [34]. The successful transmission probability becomes $p_{s,i}(\lambda_i) = \exp\left(-V(R_i, f_{o,i})\lambda_i\right)$, and the corresponding network TC is

$$f(\lambda_1, \ldots, \lambda_N) = \sum_{i=1}^{N} \lambda_i \exp\left(-V(R_i, f_{o,i})\lambda_i\right), \quad \text{where} \quad p_{o,i}(\lambda_i) \le \epsilon, \quad \forall i. \quad (8)$$

The maximum TC can be obtained through setting $\partial f(\lambda_1, \ldots, \lambda_i, \ldots, \lambda_N)/\partial \lambda_i = 0$ if $p_{o,i}(\lambda_i) \le \epsilon$. Otherwise, the optimal λ_i should be $-\ln(1-\epsilon)/V(R_i, f_{o,i})$. Consequently, the optimal transmitter density is given by

$$\lambda_i^* = \min\left\{-\frac{\ln(1-\epsilon)}{V(R_i, f_{o,i})}, \frac{1}{V(R_i, f_{o,i})}\right\}, \quad \forall i. \quad (9)$$

If the noise is not negligible, this case is of interest in the noise limited regime. The corresponding network TC becomes

$$f(\lambda_1, \ldots, \lambda_N) = \sum_{i=1}^{N} \widetilde{\lambda}_i \exp\left(-D(R_i, f_{o,i})\widetilde{\lambda}_i\right), \quad \text{where} \quad p_{o,i}(\widetilde{\lambda}_i) \le \widetilde{\epsilon}, \quad \forall i \quad (10)$$

with transmitter density $\widetilde{\lambda}_i \triangleq \lambda_i \exp(U(R_i, f_{o,i}))$, $D(R_i, f_{o,i}) \triangleq V(R_i, f_{o,i}) \exp(U(R_i, f_{o,i}))$, and $\widetilde{\epsilon} \triangleq \min_i \{\epsilon \exp(U(R_i, f_{o,i})), 1\}$. Then, problem (10) can be still formulated in the form of problem (8). Note that when $\widetilde{\epsilon} = 1$, the problem is naturally reduced to the network throughput maximization problem (a special case of the TC maximization problem). Which regime a network is operating is dependent on the specific factors or scenarios, such as the BS density in 5G mm-wave cellular networks [13]. In the following, we use the notations of problem (8) to illustrate the problem of network TC maximization under the spatial transmitter density constraint.

B. Network Transmission Capacity With Spatial Transmitter Density Constraint

The total transmitter density λ_T is fixed in physical scenarios, so the maximum TC may not be achieved since every user is scheduled with the equal priority under the assumption of the homogeneity of the network. For each λ_T, the maximum network TC needs to be analyzed in terms of the transmitter density over the wideband channel according to both the path loss at different frequencies and distances of the desired links. We will try to find appropriate transmitter density λ_i at each frequency $f_{o,i}$ such that $f(\lambda_1, \ldots, \lambda_N)$ is maximized under a given total spatial transmitter density constraint.

The optimization problem can be formulated as

$$\underset{\lambda_i, \forall i}{\text{maximize}} \ f(\lambda_1, \ldots, \lambda_N)$$

$$\text{subject to} \ \sum_{i=1}^{N} \lambda_i = \lambda_T, \quad (11a)$$

$$0 \le \lambda_i \le -\frac{\ln(1-\epsilon)}{V(R_i, f_{o,i})}, \quad \forall i \quad (11b)$$

where λ_T is a constant, predetermined by the number of transmitters per unit area in a network. The outage probability constraint is considered in (11b), meaning that $p_{s,i}(\lambda_i) = \exp(-V(R_i, f_{o,i})\lambda_i) \ge 1 - \epsilon$. Since the objective function (8) is nonconvex with respect to λ_i, there are many maxima or minima that make the optimization difficult.

IV. Optimal Solutions

In this section, the solutions of this nonconvex optimization problem are provided. Letting $x_i \triangleq \lambda_i$ and $a_i \triangleq V(R_i, f_{o,i})$, we obtain the optimization problem as

$$\mathbb{OP}\ 1: \ \underset{x_i, \forall i}{\text{maximize}} \ f(x) = \sum_{i=1}^{N} x_i \exp(-a_i x_i)$$

$$\text{subject to} \ \sum_{i=1}^{N} x_i = \lambda_T, \quad (12a)$$

$$0 \le x_i \le \gamma_i, \quad \forall i \quad (12b)$$

where $a_i \ge 0$, $\gamma_i = -\frac{\ln(1-\epsilon)}{a_i}$, and $\boldsymbol{x} \triangleq [x_1, \ldots, x_N] \in \mathbb{R}^N$. It is a resource allocation problem and also arisen in the estimation of proportions [35].

Note that constraints (12a) and (12b) may be incompatible with each other. Hence, there is a requirement for λ_T, which is

$$0 < \lambda_T \le \sum_{i=1}^{N} \gamma_i, \quad (13)$$

meaning that $\lambda_T \le -\lambda_C \ln(1-\epsilon)$ where $\lambda_C \triangleq \sum_{i=1}^{N} 1/a_i$.

A. KKT Conditions

Since the function $f_i(x_i) = x_i e^{-a_i x_i}$ is a nonconvex function, we can only get stationary points by the KKT conditions of $\mathbb{OP}$ 1, which are denoted as $\boldsymbol{x}^* \in \mathcal{S}$, where $\boldsymbol{x}^* \triangleq [x_1^*, \ldots, x_N^*] \in \mathbb{R}^N$ and set $\mathcal{S}$ denotes all locally optimal points, where the global optimal solution of $\max f(\boldsymbol{x})$ is

$$\boldsymbol{x}_o^* = \arg\max_{\boldsymbol{x}^* \in \mathcal{S}} f(\boldsymbol{x}^*). \quad (14)$$

First, we introduce Lagrange multipliers $\boldsymbol{q} \in \mathbb{R}^N$ for the inequality constraints $x_i \ge 0$, $\boldsymbol{\mu} \in \mathbb{R}^N$ for the inequality constraints $x_i \le \gamma_i$, and a multiplier $\nu \in \mathbb{R}$ for the equality constraint $\sum_{i=1}^{N} x_i = \lambda_T$. Then, we can obtain the Lagrangian

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{q}, \boldsymbol{\mu}, \nu) = -\sum_{i=1}^{N} x_i e^{-a_i x_i} + \nu\left(\sum_{i=1}^{N} x_i - \lambda_T\right) - \boldsymbol{q}^T \boldsymbol{x} + \sum_{i=1}^{N} \mu_i (x_i - \gamma_i) \quad (15)$$

and the KKT conditions [35] for $i = 1, \ldots, N$ as follows,

$$x_i^* \geq 0, \quad x_i^* \leq \gamma_i, \quad \boldsymbol{q}^* \succeq 0, \quad \boldsymbol{\mu}^* \succeq 0, \quad \sum_{i=1}^{N} x_i^* = \lambda_T,$$

$$q_i^* x_i^* = 0, \quad \mu_i^*(x_i^* - \gamma_i) = 0, \tag{16a}$$

$$-(e^{-a_i x_i^*} - a_i x_i^* e^{-a_i x_i^*}) - q_i^* + \mu_i^* + \nu^* = 0. \tag{16b}$$

Substituting (16b) to the complementary slackness conditions (16a), we have

$$x_i^* \left(\nu^* + \mu_i^* - (e^{-a_i x_i^*} - a_i x_i^* e^{-a_i x_i^*})\right) = 0, \tag{17}$$

$$(x_i^* - \gamma_i) \left(\nu^* - q_i^* - (e^{-a_i x_i^*} - a_i x_i^* e^{-a_i x_i^*})\right) = 0. \tag{18}$$

From (17), we know that either $x_i^* = 0$ or

$$x_{i,k}^* = -\frac{\mathcal{W}_k((\nu^* + \mu_i^*)e) - 1}{a_i} > 0, \quad k = 0, 1, 2 \tag{19}$$

where $\mu_i^* \geq 0$, and $\mathcal{W}_k(z), k = 0, 1, 2$, $z \geq -e^{-1}$, are the Lambert function [36], i.e., the inverse function of $x \exp(x)$. When $z > 0$, the Lambert function is a single-valued function, denoted by $\mathcal{W}_0(z)$. When $-1/e \leq z \leq 0$, the Lambert function is a double-valued function which includes two branches, denoted by $\mathcal{W}_1(z)$ and $\mathcal{W}_2(z)$ respectively.

In order to simplify the KKT conditions further, we need to give a definition of regions of each x_i for the convenience of discussion.

B. Definition of Regions

Based on the properties of the Lambert function, we define three regions for x_i as follows.

1) Definition of Region 0: When $0 \leq x_i \leq 1/a_i, \forall i$, there is a unique value when $\nu \in [0, 1]$, which is

$$x_{i,0}(\nu) \triangleq -\frac{1}{a_i}\left(\mathcal{W}_0(\nu e) - 1\right). \tag{20}$$

We define this interval as Region 0.

2) Definition of Region I: From the properties of the Lambert function [36], when $-e^{-2} \leq \nu < 0$, we know that

$$\frac{1}{a_i} < x_{i,1}(\nu) \triangleq \frac{\rho_1(\nu)}{a_i} \leq \frac{2}{a_i}, \tag{21}$$

where $\rho_1(\nu) \triangleq -(\mathcal{W}_1(\nu e) - 1)$ within $(1, 2]$ is uncorrelated with a_i, and this interval is defined as Region I. The subscript 1 indicates that the ith variable is in Region I.

3) Definition of Region II: Similarly, for the other branch of the Lambert function, we have

$$x_{i,2}(\nu) \triangleq \frac{\rho_2(\nu)}{a_i} > \frac{2}{a_i}, \tag{22}$$

where $\rho_2(\nu) \triangleq -(\mathcal{W}_2(\nu e) - 1)$ within $(2, \infty]$, and this part is defined as Region II. The subscript 2 indicates that the ith variable is in Region II.

Based on the above definitions, it is obvious that $\rho_2(\nu) > \rho_1(\nu)$ and $x_{i,2}(\nu) > x_{i,1}(\nu)$ when $-e^{-2} \leq \nu < 0$.

The regions (0, I and II) in terms of x_i are depicted in Fig. 2 and the corresponding regions in terms of ν are illustrated in Fig. 3.

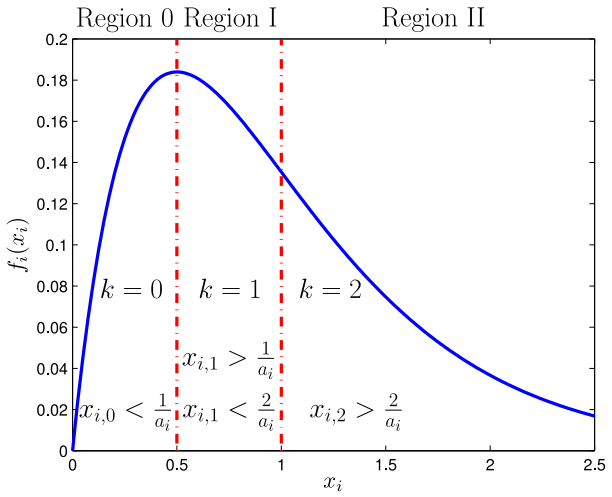

Fig. 2. Regions of the solutions from the KKT conditions, where $a_i = 2$.

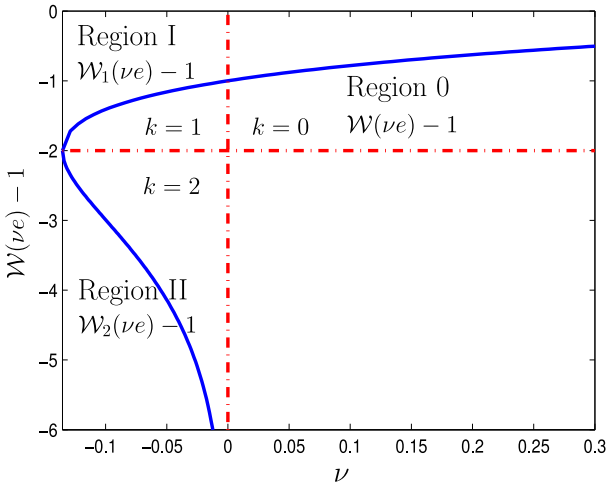

Fig. 3. Regions of the solutions in terms of the Lambert function.

C. Simplifying the KKT Conditions

For the slackness condition (17), we note that when $x_i = 0$, the increasing speed of $f_i(x_i)$ achieves the maximum for all i which is $df_i(x_i)/dx_i = 1$ when $x_i = 0$. Thus, we have $x_i^* \neq 0$. Consequently, the expression of x_i^* is shown in (19).

Based on the previous definitions for the regions, the optimal solutions can be simplified and given in Theorem 1.

Theorem 1: The KKT conditions of $\mathbb{OP}$ 1 should be $0 \leq x_i^* \leq \gamma_i$,

$$\sum_{i=1}^{N} x_i^* = \lambda_T, \tag{23a}$$

$$x_{i,2}^* = -\frac{\mathcal{W}_2((\nu^* + \mu_i^*)e) - 1}{a_i}, \quad \text{when } x_{i,2}^* = \gamma_i, \tag{23b}$$

$$x_{i,k}^* = -\frac{\mathcal{W}_k(\nu^* e) - 1}{a_i}, \quad k = 0, 1, 2 \quad \text{otherwise.} \tag{23c}$$

where $\mu_i^* > 0$ and $-1/e^2 \leq \nu^* \leq 1$.

Proof: See Appendix. B. ■

Remark 1:

(i) For the case when $x_{i,2}^* = \gamma_i$, $x_{i,2}^*$ does not change as ν^* changes when we solve (23b) with dual variable μ_i^*,

since it is independent with other variables. Hence, we can focus on the case when $x^*_{i,2} \neq \gamma_i$.

(ii) If $x^*_{i,k} = -\frac{\mathcal{W}_k(\nu^* e)-1}{a_i} = \gamma_i, \quad k = 0, 1$, we have $\mathcal{W}_k(\nu^* e) - 1 = \ln(1-\epsilon)$, which is independent of a_i. This result indicates that $x^*_{i,k}, \forall i$ are equal to γ_i at the same time when $\lambda_T = -\lambda_C \ln(1-\epsilon)$.

Therefore, based on (i) and (ii), we will only consider the case when $x^*_i < \gamma_i, \forall i$ in Section V and Section VI.

V. Properties of the KKT Solutions

In general, finding the global optimal solution for a nonconvex problem is not easy and the corresponding complexity is exponential with respect to the problem dimension. However, through observing the characteristics of the KKT conditions, we know that there is a relationship among $x^*_i, \forall i$. Then it is possible to design an algorithm such that this problem can be solved with low computational complexity. First, we study the relationships of $x^*_i, \forall i$ as follows.

A. Relationship Among Variables

From the KKT conditions, the two variables x^*_i and x^*_j, where $i, j \in \mathcal{N}, i \neq j, \forall i, j$, keep a certain relationship, i.e.,

$$\frac{x^*_{i,k}}{x^*_{j,k}} = \frac{a_j}{a_i}, \quad i \neq j, i, j \in \mathcal{N}, k = 0, 1, 2. \tag{24}$$

Hence, every $x^*_{i,k}$ can be expressed by

$$x^*_{i,k} = \frac{a_1}{a_i} x^*_{1,k} = \xi_i x^*_{1,k} \tag{25}$$

where $\xi_i \triangleq a_1/a_i$.

B. Closed Form of the KKT Solutions

If $x^*_i, \forall i$ are from the same branch of the Lambert function, then according to (23a) we obtain

$$\sum_{i=1}^{N} x^*_i = x^*_{1,k} \sum_{i=1}^{N} \xi_i = \lambda_T. \tag{26}$$

It is obvious that the closed form optimal solutions can be given directly, which are

$$\begin{aligned} x^*_{1,k} &= \frac{\lambda_T}{\sum_{i=1}^{N} \xi_i}, \\ x^*_{j,k} &= \xi_j x^*_{1,k}, \quad j = 2, \ldots, N. \end{aligned} \tag{27}$$

Remark 2: If $x_i, \forall i$ are within Region 0, the objective function is concave, since $d^2 f_i(x_i)/d^2 x_i = a_i(a_i x_i - 2)e^{-a_i x_i} < 0$ when $x_i < 1/a_i$ and $f(x)$ is separable. In this region, $\mathbb{OP}$ 1 is a convex optimization problem and therefore there is a unique value for each x^*_i given ν^*, which is the global optimal solution of $\mathbb{OP}$ 1.

When $\nu^* = 0$, i.e., $\sum_{i=1}^{N} x^*_i = \sum_{i=1}^{N} 1/a_i = \lambda_C$, the objective function achieves the maximum value. If $\lambda_T > \lambda_C$, the variables must be within Region I or Region II, which is a complicated case. The reason is that the Lambert function is a double-valued function in these regions. We will illustrate this point in the next section.

C. Challenges of the Problem

There are two solutions for each $x^*_i(\nu)$ when $\nu < 0$, so the permutation of the possible combination of $x^*_{i,1}$ and $x^*_{j,2}$ where $i \in \mathcal{A}, j \in \mathcal{A}^c, \mathcal{A} \in \mathcal{N}$ should be $\mathcal{O}(2^N)$ in (23a). In this case it results in a large amount of computational complexity as N increases, since we have to verify every possible combination to figure out the global optimal solution. If we can exclude some solutions whose corresponding objective functions are always local maxima or minima, the complexity of finding the global solution can be decreased significantly. The idea is to reduce the cardinality of set $\mathcal{S}^\dagger \in \mathcal{S}$, where set $\mathcal{S}^\dagger$ includes all the candidates of the global optimal solution.

In the next section, we will consider the properties of the possible solutions that are given by the KKT conditions when $x^*_i \geq 1/a_i$.

VI. Cardinality of Set $\mathcal{S}^\dagger$

In this section, we use two steps to reduce the cardinality of set $\mathcal{S}^\dagger$ in terms of N from exponential to a constant.

A. Reducing the Cardinality of Set $\mathcal{S}^\dagger$ From $\mathcal{O}(2^N)$ to $\mathcal{O}(N)$

Lemma 1: In the global optimal solution x^*, there is at most one optimal variable in Region II.

Proof: See Appendix. A. ∎

According to Lemma 1, it is clear that the number of possible combinations of $x^*_{i,1}$ and $x^*_{j,2}$ is reduced from $\mathcal{O}(2^N)$ to $\mathcal{O}(N)$ where $i \in \mathcal{A}, j \in \mathcal{A}^c, \mathcal{A} \in \mathcal{N}$.

Remark 3: In the proof of Lemma 1, only the convexity of the objective function is considered, so the idea of the proof is not just restricted for this problem. The conclusion of Lemma 1 can be extended to more general problems where the objective function has a similar convexity of the one in $\mathbb{OP}$ 1. To be more specific, the statement is as the following,

Corollary 1: Consider problem

$$\begin{aligned} &\underset{x_i, \forall i}{\text{maximize}} \sum_{i=1}^{N} g_i(x_i) \\ &\text{subject to} \sum_{i=1}^{N} x_i = c \end{aligned}$$

where $c \geq 0$ is a constant, $g_i(x_i), \forall i$ are convex when $x_i \geq x^\dagger_i$ where $x^\dagger_i$ are some constants. Then, in the global optimal solution, there is at most one variable that is strictly greater than $x^\dagger_i$.

Proof: The proof of this corollary basically is the same as the proof of Lemma 1, just by replacing the transition point $2/a_i$ by $x^\dagger_i$ and λ_T by c. ∎

Remark 4: Lemma 1 also implies that there is another case that all the variables are in Region I.

Next, we need to figure out which variable should be in Region II.

B. Reducing the Cardinality of Set $\mathcal{S}^\dagger$ From $\mathcal{O}(N)$ to $\mathcal{O}(1)$

If we know which x^*_i should be in Region II, the algorithm can be simplified further. Based on Lemma 1, in this section

we consider the case that for $x^* \in \mathcal{S}^\dagger$ there is one x_i^* in Region II while other ones are in Region I.

First, we consider the simplest case where there are only two variables. The result in this case will be very useful to discuss the N variables case.

1) Two Variables: Before the discussion, we need to introduce two lemmas which give the inherent properties of the Lambert function as follows,

Lemma 2: The absolute value of the derivative of $\mathcal{W}_2(\nu)$ is larger than the one of $\mathcal{W}_1(\nu)$ where $-e^{-1} < \nu < 0$.

Proof: See Appendix. C. ■

Lemma 3: The absolute values of the derivative of $\rho_1(\nu)$ and $\rho_2(\nu)$ are arbitrarily close at $\nu = -e^{-1}+d\nu$ when $dv \to 0$.

Proof: See Appendix. D. ■

Then, we consider the case which there are only two variables, i.e., $x = [x_1, x_2]$. Assume $a_1 < a_2$. When $x_i^* < \gamma_i$, $i = 1, 2$, $\mathbb{OP}$ 1 is reduced to

$$\begin{aligned} \mathbb{OP}\ 2: \ & \text{maximize } x_1 e^{-a_1 x_1} + x_2 e^{-a_2 x_2} \\ & \text{subject to } x_1 + x_2 = \lambda_T, x_1 \geq 0, x_2 \geq 0. \end{aligned} \quad (28)$$

Then there are two cases if one solution is in Region I and the other is in Region II.

- Case 1: Define the function

$$\Sigma_1(\nu) \triangleq x_{1,1}(\nu) + x_{2,2}(\nu) = \frac{\rho_1(\nu)}{a_1} + \frac{\rho_2(\nu)}{a_2}. \quad (29)$$

There exists a ν' such that $\Sigma_1(\nu') = \lambda_T$ and the corresponding objective function is denoted as

$$J_1 = \frac{\rho_1'}{a_1} e^{-\rho_1'} + \frac{\rho_2'}{a_2} e^{-\rho_2'} \quad (30)$$

where $\rho_1' \triangleq -(\mathcal{W}_1(\nu' e) - 1)$ and $\rho_2' \triangleq -(\mathcal{W}_2(\nu' e) - 1)$. Applying Lemma 3 we have the relationship: $-d\rho_1/d\nu = d\rho_2/\nu$ when $\nu \to -e^{-2}$ and $a_1 < a_2$. The derivative of $\Sigma_1(\nu)$ is

$$\begin{aligned} \frac{d\Sigma_1(\nu)}{d\nu} &= -\frac{1}{a_1}\left|\frac{d\rho_1(\nu)}{d\nu}\right| + \frac{1}{a_2}\left|\frac{d\rho_2(\nu)}{d\nu}\right| \\ &= \left(\frac{1}{a_2} - \frac{1}{a_1}\right)\left|\frac{d\rho(\nu)}{d\nu}\right| < 0 \end{aligned} \quad (31)$$

where $|d\rho(\nu)/d\nu|$ represents the absolute value of the derivative of $\rho_1(\nu)$ and $\rho_2(\nu)$. Therefore, $\Sigma_1(\nu)$ decreases first.

Applying Lemma 2 we know that as ν increases, $\Sigma_1(\nu)$ increases. After getting back to $2/a_1 + 2/a_2$, $\Sigma_1(\nu)$ can go to infinity.

Remark 5: $\Sigma_1(\nu)$ is decreasing first from $\lambda_C = 1/a_1 + 1/a_2$ and increasing to infinity when $\nu \in (-e^{-2}, 0)$.

- Case 2: Define the function

$$\Sigma_2(\nu) \triangleq x_{1,2}(\nu) + x_{2,1}(\nu) = \frac{\rho_2(\nu)}{a_1} + \frac{\rho_1(\nu)}{a_2}. \quad (32)$$

Similarly, there exits a ν'' such that $\Sigma_2(\nu'') = \lambda_T$. Then, the corresponding objective function is

$$J_2 = \frac{\rho_1''}{a_2} e^{-\rho_1''} + \frac{\rho_2''}{a_1} e^{-\rho_2''} \quad (33)$$

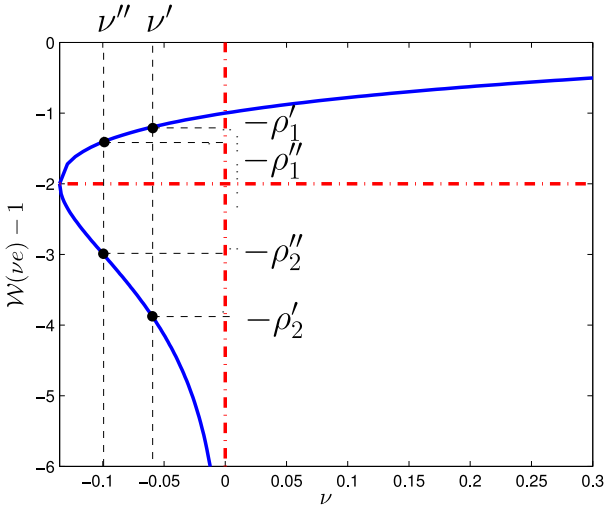

Fig. 4. The relationship between ν' and ν'' for $a_1 < a_2$.

where $\rho_1'' \triangleq -(\mathcal{W}_1(\nu'' e) - 1)$ and $\rho_2'' \triangleq -(\mathcal{W}_2(\nu'' e) - 1)$. Meanwhile, based on Lemma 2 and $a_1 < a_2$, we have

$$\frac{d\Sigma_2(\nu)}{d\nu} = -\frac{1}{a_2}\left|\frac{d\rho_1(\nu)}{d\nu}\right| + \frac{1}{a_1}\left|\frac{d\rho_2(\nu)}{d\nu}\right| > 0. \quad (34)$$

Thus, $\Sigma_2(\nu)$ is always greater than $2/a_1 + 2/a_2$.

Remark 6: $\Sigma_2(\nu)$ is monotonically increasing from $\lambda_C = 1/a_1 + 1/a_2$ when $\nu \in (-e^{-2}, 0)$.

Lemma 4: For $\mathbb{OP}$ 2, under constraint $\Sigma_1(\nu') = \Sigma_2(\nu'') = \lambda_T$, we have $\nu'' < \nu'$ when $\nu', \nu'' \in (-e^{-2}, 0)$.

Proof: See Appendix. E. ■

In Fig. 4, we show the relationship between ν'' and ν' under constraint $\Sigma_1(\nu') = \Sigma_2(\nu'') = \lambda_T$.

With the help of Lemma 4, we can obtain the relationship between J_1 and J_2 by the following lemma.

Lemma 5: For $\mathbb{OP}$ 2, when $\lambda_T = 2/a_1 + 2/a_2$ and $a_1 < a_2$, $J_1 > J_2$ always holds.

Proof: See Appendix. F. ■

If we can know $J_1 > J_2$ or $J_1 < J_2$ for all $\Sigma_1(\nu') = \Sigma_2(\nu'') = \lambda_T$, then the smaller one can be excluded from set $\mathcal{S}^\dagger$. Based on Lemma 4 and Lemma 5, we will have Lemma 6.

Lemma 6: For $\mathbb{OP}$ 2 as λ_T increases from $\lambda_T = 2/a_1 + 2/a_2$ to infinity, $J_1 > J_2$ is always true, where the corresponding $[x_{1,1}^*, x_{2,2}^*] = [\rho_1'/a_1, \rho_2'/a_2]$.

Proof: See Appendix. G. ■

Thus, for $\mathbb{OP}$ 2, the solution $[x_{1,2}^*, x_{2,1}^*] = [\rho_2''/a_1, \rho_1''/a_2]$ can be excluded from $\mathcal{S}^\dagger$. This conclusion is also very useful for the case where there are N variables as follows.

2) N Variables: With Lemma 1 and Lemma 6 in mind, we have the following theorem:

Theorem 2: For $\mathbb{OP}$ 1 if $x_i^*, \forall i$ come from both Region I and Region II, set $\mathcal{S}^\dagger$ involves $N-1$ variables in Region I and only one in Region II with $a_{\max}$, where the corresponding $a_{\max} = \max\{a_i\}, \forall i$.

Proof: See Appendix. H. ■

From Theorem 2, if $x_i^*, \forall i$ come from both Region I and Region II, set $\mathcal{S}^\dagger$ only involves one possible combination in terms of both $x_{i,1}$ and $x_{i,2}$. In this case, we define a function

TABLE I
OPTIMAL ALLOCATION SCHEME

$\lambda_T \le \lambda_C$	Case 1: We have $f_{\max} = f_0$.
$\lambda_C \le \lambda_T \le 2\lambda_C$	Case 2: $\lambda_C \le \lambda_T \le 2\lambda_C$, we have f_{I}. Case 3: $\Sigma^{\dagger}_{\min} < \lambda_T < 2\lambda_C$ Step 1: Find $\Sigma^{\dagger}_{\min}$ and $\widetilde{\nu}$. Step 2: Get $f^{(1)}_{\text{I,II}}$ and $f^{(2)}_{\text{I,II}}$ by (36). Step 3: Obtain $f_{\text{I,II}} = \max\left\{f^{(1)}_{\text{I,II}}, f^{(2)}_{\text{I,II}}\right\}$. We have $f_{\max} = \max\{f_{\text{I}}, f_{\text{I,II}}\}$.
$\lambda_T \ge 2\lambda_C$	Case 3: We obtain $f_{\max} = f_{\text{I,II}}$ within $[\widetilde{\nu}, 0)$ by (36).

as the following, i.e.,

$$\Sigma^{\dagger}(\nu) \triangleq x_{j,2}(\nu) + \sum_{i=1, i\neq j}^{N} x_{i,1}(\nu) \tag{35}$$

where $x_{j,2}$ is with $a_{\max}$. Then, the bisection algorithm can be adopted to find ν^* such that $\Sigma^{\dagger}(\nu^*) = \lambda_T$.

VII. ALGORITHM IMPLEMENTATION

In this section, we provide an algorithm of obtaining the global optimal solution of $\mathbb{OP}$ 1 and discuss about the complexity of the proposed method.

A. For the Case $x_i^* < \gamma_i$

According to Section VI, we know that set $\mathcal{S}^{\dagger}$ only involves three cases.

1) All variables are within Region 0, when $\lambda_T \le \lambda_C$. There is a closed-form solution according to (27). The maximum value is denoted as f_0.
2) All variables are within Region I, where the function $\Sigma^{\ddagger}(\nu) \triangleq \sum_{i=1}^{N} x_{i,1}(\nu) \in [\lambda_C, 2\lambda_C], \nu \in [-e^{-2}, 0)$. This case is similar as Case 1, where the maximum value is denoted as f_{I}.
3) All variables are within Region I except one of them within Region II.
   According to the relationships of $x^*_{i,k}, k = 1, 2$ in (25), we have $x^*_{i,1} = \frac{a_1}{a_i} x^*_{1,1} = \xi_i x^*_{1,1}$ and $x^*_{i,2} = \xi_i x^*_{1,2}$. Therefore, after some manipulations, we know that (35) can be reduced further as

$$\Sigma^{\dagger}(\nu) = \rho_2(\nu)\frac{1}{a_{\max}} + \rho_1(\nu) \sum_{i=1, i\neq j}^{N} \frac{1}{a_i}. \tag{36}$$

   The goal is still to find the ν^* such that $\Sigma^{\dagger}(\nu^*) = \lambda_T$. However, based on Remark 5 there may be two solutions that both satisfy this constraint when $\lambda_T \le 2\lambda_C$, since $\Sigma^{\dagger}_{\min} \le 2\lambda_C$ where $\Sigma^{\dagger}_{\min}$ is the minimum value of function $\Sigma^{\dagger}(\nu), \nu \in [-e^{-2}, 0)$.
   Therefore, we need to
   (i) Find out $\Sigma^{\dagger}_{\min}$ and $\widetilde{\nu}$ by the bisection algorithm, where $\Sigma^{\dagger}(\widetilde{\nu}) = \Sigma^{\dagger}_{\min}$.
   (ii) Implement the bisection algorithm in intervals $[-e^{-2}, \widetilde{\nu}]$ and $[\widetilde{\nu}, 0)$ using (36) separately to find out the two possible solutions: $\widetilde{\nu}^{(1)}$ and $\widetilde{\nu}^{(2)}$, where $\Sigma^{\dagger}(\widetilde{\nu}^{(1)}) = \Sigma^{\dagger}(\widetilde{\nu}^{(2)}) = \lambda_T$, and give the corresponding objective values: $f^{(1)}_{\text{I,II}}$ and $f^{(2)}_{\text{I,II}}$.
   (iii) Finally, the maximum value of the objective function is simply given by $f_{\text{I,II}} = \max\{f^{(1)}_{\text{I,II}}, f^{(2)}_{\text{I,II}}\}$.

Remark 7: When $0 \le x_i \le 1/a_i, \forall i$ (Region I), the optimization problem $\mathbb{OP}$ 1 is strongly convex according to Remark 2, which implies $\Sigma^{\dagger}_{\min} > \lambda_C$.

We summarize the whole optimal allocation strategy as in Table I, where the global optimal value is denoted by $f_{\max}$.

B. For the Case $\exists i$, $x_i^* = \gamma_i$

There are two cases as follows.

1) If $x^*_{i,k} = \gamma_i, \quad k = 0, 1$, then $x^*_i = \gamma_i, \forall i$.
2) If $x^*_{i,2} = \gamma_i$, then the ith variable is simply ignored from the optimization process, since it serves as a constant. Then, we just need to consider the other variables.

C. Complexity of the Algorithm

Since the explicit solution from (27) for f_0 and f_{I} can be calculated directly, and the complexity of computing the solution from (36) for $f_{\text{I,II}}$ is at the order of one-dimensional search algorithm in terms of the Lambert function. Actually, the Lambert function has been used in many resource allocation problems in wireless networks [37], [38] and the function can be computed very efficiently. In summary, the total complexity is still proportional to the one-dimensional search algorithm with the Lambert function. The one-dimensional search can be interpreted as search for a water-filling solution, although the problem solved is non-convex, as opposed to convex as in classical water-filling problems.

VIII. NUMERICAL RESULTS

In this section, we numerically calculate and optimize the network TC of an example for underwater acoustic communication systems. The channel path loss in underwater environment is dependent on carrier frequency and given by $h(R_i, f_{o,i}) = R_i^{\alpha} a(f_{o,i})^{R_i}$ at frequency $f_{o,i}$ [30], where α is the path loss exponent (spreading factor) and $a(f_{o,i})^{R_i}$ models part of channel gain due to absorption. $V(R_i, a(f_{o,i}))$ is derived as in [5, eq. (6)], i.e.,

$$V(R_i, f_{o,i}) = c_d \int_0^{\infty} \frac{dr^{d-1}}{1 + \frac{r^{\alpha} a(f_{o,i})^r}{\beta R_i^{\alpha} a(f_{o,i})^{R_i}}} dr$$

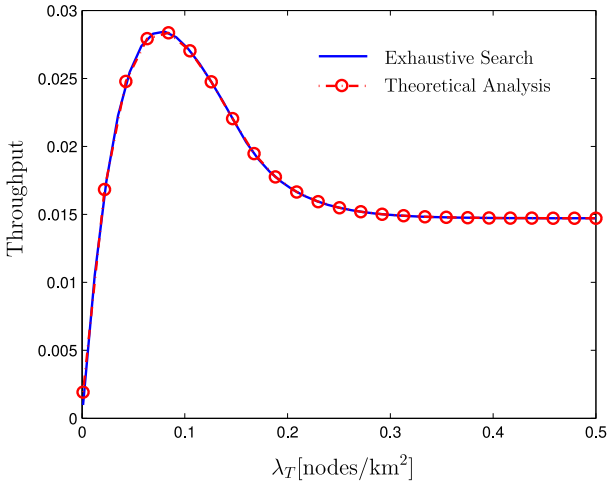

Fig. 5. Comparison between theoretical result and exhaustive searching in terms of λ_T, where there are two subbands, $f_{o,1} = 20$kHz and $f_{o,2} = 50$kHz.

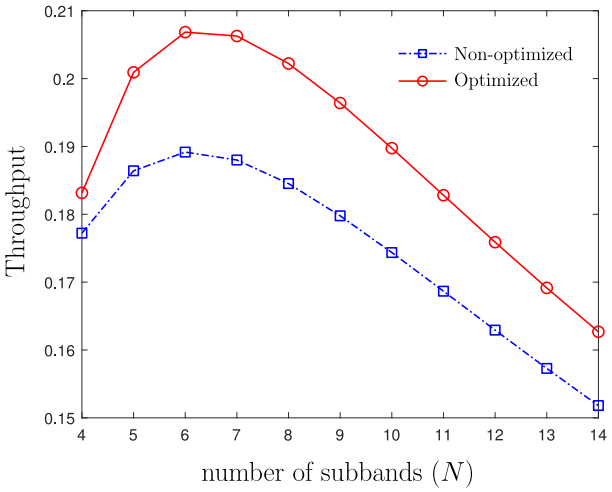

Fig. 7. Throughput versus N where $R_i = 1000, \forall i$, and $\lambda_T = 1.8$[nodes/km^2].

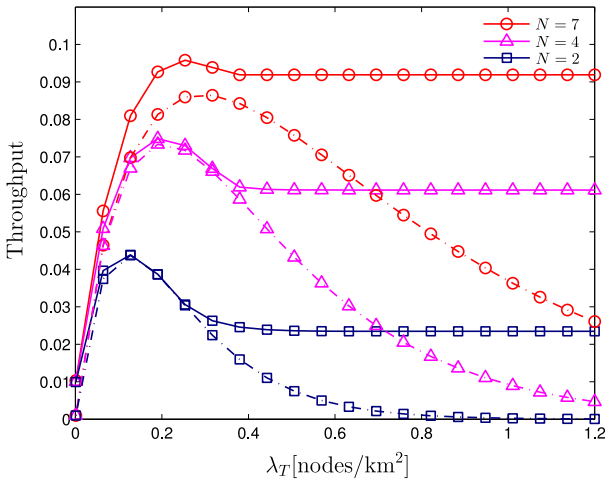

Fig. 6. Throughput versus λ_T with different number of subbands. Dashed lines: the transmitter density is equally allocated in each subband, Solid lines: the results of optimized λ_i.

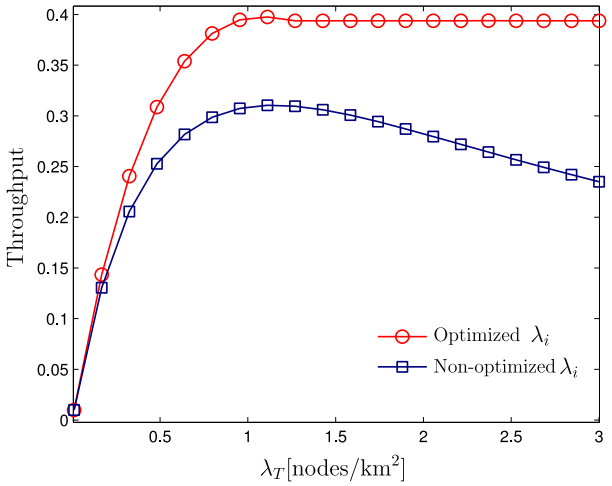

Fig. 8. Throughput versus λ_T when $R_i, i = 1, \ldots, N$ are different in each subband, where $N = 20$.

where $c_d = \text{Vol}(B_d(0,1))$ is the volume of the d-dimensional unit ball. The parameters are chosen as $\alpha = 1.5$, $\beta = 10$dB, $R_i = 1000$m, $d = 3$, $a(f_{o,i})$ are adopted according to [5] and carrier frequency $f_{o,i}, \forall i$ are uniformly selected from 20kHz to 50kHz where $i = 1, \ldots, N$.

A. Throughput Analysis

We first verify the proposed transmitter density allocation scheme as shown in Table I, where the constraint of an outage probability is ignored. In Fig. 5, the results from exhaustive search are compared with that obtained by the proposed algorithm, which shows the consistency between the numerical simulations and theoretical analysis.

From Fig. 6, it is obvious that the optimal allocation scheme for each subband provides higher throughput gain compared with the non-optimized one that only equally allocates the spatial density of transmitters over the different frequency subbands. As λ_T becomes large, it can be observed that the throughput decreases if $\lambda_i, i = 1, \ldots, N$ are equal in each subband. However, in the optimal allocation strategy, the optimized throughput does not change for large enough λ_T, since the redundant users are allocated into the worst channel while the other users are sharing with the rest of the frequency subbands. The results show clearly that the optimal scheme yields a higher throughput. Also, we can observe that as the number of subbands increases, the difference between the optimized one and non-optimized one becomes large.

Fig. 7 shows that the network throughput increases and then decreases as the total number of subbands increases, where the total bandwidth (20kHz) is fixed (from $f_{o,1} = 20$kHz to $f_{o,N} = 40$kHz) and the whole bandwidth is divided into N subbands. It can be observed that when N is increased, the effect of interference in each subband becomes weaker but the available bandwidth is reduced, resulting in a tradeoff of the network throughput shown in Fig. 7.

Furthermore, $h(R_i, f_{o,i})$ is also dependent on R_i. In Fig. 8, $R_i > 0, i = 1, \ldots, N$ are randomly generated with

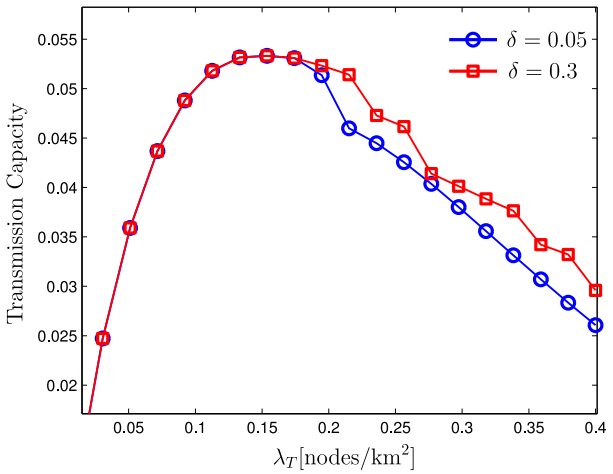

Fig. 9. TC versus λ_T with different outage probability constraints, where $N = 5$ and $\epsilon = 1 - \exp(-(\lambda_T + \delta)/\lambda_C)$.

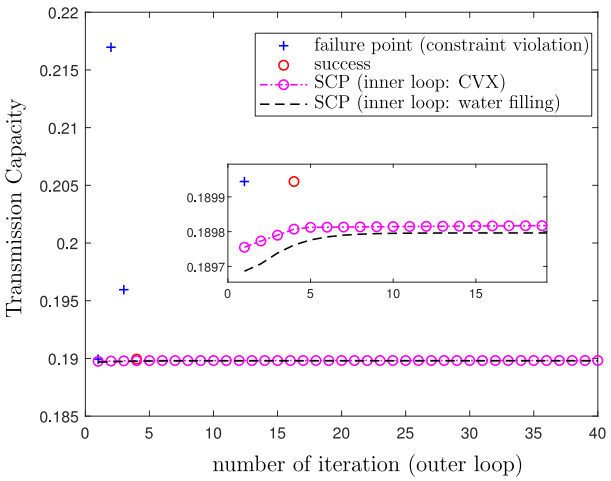

Fig. 10. Comparisons between the proposed algorithm and SCP [24], where $N = 5$, $\epsilon = 1 - \exp(-(\lambda_T + \delta)/\lambda_C)$, $\lambda_T = 1.5$ and $\delta = 0.3$.

mean 1000m and variance 100m, which follows a truncated Gaussian distribution. This figure compares the optimized throughput and the non-optimized one, which shows significant gains that the optimized method provides. Similar as in Fig. 6, when λ_T is large, the optimized algorithm scarifies the outage probability of the worst channel so that the total throughput is very high.

B. TC Analysis

From Fig. 6 and Fig. 8, we can see that the throughput may not be a good enough metric of studying the network volume, when the node density is large. The outage probability also plays important roles in measuring the quality of the transmission. In order to observe the changes of TC as λ_T increases to a large value, we consider the case where ϵ is large such that the outage probability can satisfy the constraint. In Fig. 9, based on (13) we assume $\lambda_T + \delta = -\lambda_C \ln(1 - \epsilon)$ where δ is a number. We observe that when λ_T is large, TC will be decreased since λ_i has to satisfy the constraints of both total transmitter density and outage probability. Also, it can be seen that TC is higher for a larger δ since the system can undertake a higher outage probability. These results are consistent with both theory and intuition. It is not applicable to allocate the transmitter density equally for each subband, because some λ_i may make the outage probability violate the constraint. It is still of interest when the outage probability is high, due to the possibility of using automatic repeat request (ARQ) mechanism [39]–[41]. Note that our analyzed probability is for one transmission, rather than for multiple repeated transmissions (due to ARQ).

Although the optimal TC is given, TC can be further increased if more advanced transmission strategies are adopted. For instance, when λ_T is small, if the transmitters can send packets over multiple subbands, TC may be increased; when λ_T is large, i.e., dense wireless networks, interference management needs to be considered [42], [43], e.g., interference cancelation, such that the equivalent transmitter density can be reduced [44].

C. Algorithms Comparison

In this section, we give more details about how the proposed algorithm works and also compare the solutions obtained by the proposed algorithm and SCP (denoted as MN in [24]) shown in Fig. 10. Consider $N = 5$, $\boldsymbol{a} = [2, 6, 8, 10, 12]$ and constraint $\lambda_T = 1.5$[nodes/km^2], $\epsilon = 1 - \exp(-(\lambda_T + \delta)/\lambda_C)$, and $\delta = 0.3$. We can know that $\lambda_C \leq \lambda_T \leq 2\lambda_C$, which is a relatively complicated case. At the first iteration, after running the proposed algorithm shown in Table I, we can get the optimal solution without the outage probability constraint. We need to check whether $x^*_{i,2}, \forall i$ is greater than $\gamma_i, \forall i$. In this case, $x^*_{5,2} > \gamma_5$, meaning that $\Sigma^\dagger(\nu)$ is not large enough when the outage constraint is considered. Then, we set $x^*_5 = \gamma_5$, since this variable has achieved the boundary (or the 5th channel has been saturated). In Fig. 10, we mark this case as blue "+" in the legend. Continue implementing the proposed algorithm and checking the boundary points. We have the same situations as in the first iteration until the 4th one. In this case, there are three points that achieve the boundary and the other two are within region I, where the final solution is marked as "○" in the red color. The total time consumed by the proposed algorithm is 0.22s.

We also implement the SCP algorithm, which contains both inner and outer loops. The inner loop of SCP solves a convex optimization problem, which can be solved either using the CVX package [35] or using a water-filling algorithm [45]. We plot the outer loop convergence behavior of the SCP algorithm in terms of the iteration. It can be observed that the algorithm converges very fast, even the first iteration of SCP has already provided a good enough solution. For the CVX-based solution, it takes 3.93s on solving the convex problem in the first inner loop. If the inner loop is solved by the water-filling algorithm, it only takes 0.07s for the first iteration, where we set the error tolerance as 10^{-8}. Based on these simulations, SCP provides a viable alternative to solve the proposed problem iteratively. A major reason for preferring our proposed algorithm is that the problem is solved in the dual domain, which involves only one-dimensional search,

whereas SCP solves the problem in the primal domain where the solution quality and running time rely heavily on the convex optimization solvers. Theoretically, the SCP algorithm is an Oracle algorithm, where we assume the optimization problem in the inner loop is solved without any error (or with an infinite number of iterations in general). However, the derived optimal solution only involves one-dimensional search and hence does not have such a requirement.

IX. Conclusions

In this paper, the network TC of an ad hoc network was optimized over frequency-selective channels by allocating the transmitter densities in different frequency subbands. The problem was formulated as a nonconvex problem under the constraints of a total transmitter density and any outage probability. The global optimal solutions of the nonconvex problem were given, where an iterative resource allocation scheme using the bisection algorithm was proposed. Simulation results showed that the TC with the optimized transmitter density is consistent with the one obtained through exhaustive search as the total density of the transmitters increases. When the total transmitter density is either low or high, more advanced transmission techniques could be considered as the future work, such as the transmitters can send the packets over a subset of the whole spectrum or interference management is adopted.

Appendix

A. Proof of Lemma 1

Proof: Considering function $f_i(x_i) = x_i e^{-a_i x_i}$, we know that the transition point between the concave and convex parts is at $x_i = 2/a_i$. When $1/a_i \le x_i \le 2/a_i$, function $f(x_i)$ is concave. When $x_i \ge 2/a_i$, function $f(x_i)$ is convex.

We need to prove the following: it is impossible that there are two and more solutions in $\mathcal{S}^\dagger$ which are in Region II. Consider a procedure as follows:

1) Select any two x_i^* from x^* which are $x_i^* > 2/a_i$ and $x_j^* > 2/a_j$ where $i \neq j$ and $a_i \neq a_j$.

2) Fixing other variables, we have $x_i^* + x_j^* = \widehat{\lambda}_T$. Because $f_i(x_i)$ and $f_j(x_j)$ are both convex in this region, $f_i(x_i) + f_j(x_j)$ is also convex under the constraint $\widehat{\lambda}_T$.

3) According to the convexity of the objective function, the maximum point of $f_i(x_i)+f_j(x_j)$ is located at the boundary of the convex set, i.e., the maximum point is located either at $[\widetilde{x}_i, \widetilde{x}_j] = [2/a_i, \widehat{\lambda}_T - 2/a_i]$ or $[\widetilde{x}_i, \widetilde{x}_j] = [\widehat{\lambda}_T - 2/a_j, 2/a_j]$, which means the previous points (x_i^* and x_j^*) are not the global optimal solution.

4) Update points $[x_i^*, x_j^*]$ by $[\widetilde{x}_i, \widetilde{x}_j]$ that has the largest objective value, and go to step 1).

Since the objective function is separable, it is clear that $f(x)$ during this process is monotonically increasing until at most one solution is greater than $2/a_i$ after enumerating all $x_i^* > 2/a_i, i \in \mathcal{N}$, meaning that there is at most one solution which is in Region II. ∎

B. Proof of Theorem 1

Proof: In slackness condition (18), there are two cases,

1) If $x_{i,k}^* \neq \gamma_i$:
we have $x_{i,k}^* = -\frac{\mathcal{W}_k((\nu^* - q_i^*)e)-1}{a_i}$ where $k = 0, 1, 2$. Combining with (19) and $\mu_i^* \ge 0$ and $q_i^* \ge 0$, we conclude that $\mu_i^* = q_i^* = 0$ and

$$x_{i,k}^* = -\frac{\mathcal{W}_k(\nu^* e) - 1}{a_i} < \gamma_i, \tag{37}$$

2) If $x_{i,k}^* = \gamma_i$:
we have

$$x_{i,k}^* = -\frac{\mathcal{W}_k((\nu^* + \mu_i^*)e) - 1}{a_i} = \gamma_i. \tag{38}$$

It is interesting to see that

$$\mathcal{W}_k((\nu^* + \mu_i^*)e) - 1 = \ln(1 - \epsilon), \tag{39}$$

which is uncorrelated with a_i, meaning that if $x_{i,k}^* = \gamma_i$ and $x_{j,k}^* = \gamma_j$ where $i \neq j$, then $\mu_i^* = \mu_j^*$. Based on the KKT conditions, $x_i^*, \forall i$ also need to satisfy $\sum_{i=1}^N x_i^* = \lambda_T$ with ν^* and μ_i^*.
In (23a), if $x_i^*, \forall i$ are from the same region, then $\nu^* + \mu_i^*$ in (19) can be considered as one variable. Consequently, x_i^* has the same form as (37).
In (23a) if $x_i^*, \forall i$ are from different regions, then we need to discuss the two cases as follows,

a) In (23a) if there are both $x_{i,2}^*$ and $x_{j,1}^*$, where $i \neq j$, then $x_{i,2}^* = \gamma_i$ and $x_{j,1}^* < \gamma_j$ hold. In this case, there exits dual variable μ_i^* that ensures $x_{i,2}^* = \gamma_i$ and dual variable ν^* that keeps $\sum_{i=1}^N x_i^* = \lambda_T$. Therefore, we only need to consider $x_{j,1}^* < \gamma_j$, since all $x_{i,2}^* = \gamma_i$ are fixed.

b) Except case a): for example, consider combination $x_{i,0}^*$ and $x_{j,1}^*$ in (23a). We know $x_{i,0}^* < x_{j,1}^*$ where $i \neq j$, since $\mu_i^* > 0$. Because $x_{i,0}^* = \gamma_i$, then we know $x_{j,1}^* > \gamma_j$ based on (38), which contradicts the constraint $x_j^* \le \gamma_j$. Therefore, we conclude $\mu_i^* = 0$. The other combinations can be also easily verified in the same way and show the same result.

In conclusion, expression form (37) of x_i^* is used frequently in the proof, since there is only one case that (38) is active which is when $x_{i,2}^* = \gamma_i$.

∎

C. Proof of Lemma 2

We need to prove the derivative of branch $\mathcal{W}_2(\nu)$ is bigger than $\mathcal{W}_1(\nu)$, i.e., $d\mathcal{W}_2(\nu)/d\nu > d\mathcal{W}_1(\nu)/d\nu$ when $-e^{-1} < \nu < 0$.

Proof: According to [36], we have

$$\mathcal{W}_1'(\nu) \triangleq \frac{d\mathcal{W}_1(\nu)}{d\nu} = \frac{\mathcal{W}_1(\nu)}{\nu(1 + \mathcal{W}_1(\nu))}, \tag{40}$$

$$\mathcal{W}_2'(\nu) \triangleq \frac{d\mathcal{W}_2(\nu)}{d\nu} = \frac{\mathcal{W}_2(\nu)}{\nu(1 + \mathcal{W}_2(\nu))}. \tag{41}$$

Since $-1 < \mathcal{W}_1(\nu) < 0$ and $\mathcal{W}_2(\nu) < -1$, we know that the signs of the derivatives of the two branches are inverse. Consequently, the difference between the absolute values of the two derivatives is just the sum of (40) and (41).

Also, we know that when $\nu = 0$ $\mathcal{W}_1'(\nu)$ is 1 and $\mathcal{W}_2'(\nu)$ approaches to infinity [36], i.e., their sum is greater than 0. Based on the fact that $\mathcal{W}_1'(\nu)$ and $\mathcal{W}_2'(\nu)$ are continuous functions, we only need to prove that there is no ν such that the sum of their derivatives equals zero when $-e^{-1} < \nu < 0$.

Assuming they were equal, we obtain the relationship

$$\mathcal{W}_1(\nu) + \mathcal{W}_2(\nu) + 2\mathcal{W}_1(\nu)\mathcal{W}_2(\nu) = 0, \tag{42}$$

which means

$$\mathcal{W}_1(\nu) = -\frac{\mathcal{W}_2(\nu)}{1 + 2\mathcal{W}_2(\nu)}. \tag{43}$$

According to the definition of the Lambert function, we have

$$\mathcal{W}_1(\nu)e^{\mathcal{W}_1(\nu)} = \mathcal{W}_2(\nu)e^{\mathcal{W}_2(\nu)}. \tag{44}$$

After some algebraic manipulations, we arrive at

$$\mathcal{W}_1(\nu) - \mathcal{W}_2(\nu) = \ln\frac{\mathcal{W}_2(\nu)}{\mathcal{W}_1(\nu)}. \tag{45}$$

Substituting (43) into (45), we obtain

$$-\mathcal{W}_2(\nu)\frac{2(\mathcal{W}_2(\nu)+1)}{2\mathcal{W}_2(\nu)+1} = \ln\left(-(1+2\mathcal{W}_2(\nu))\right). \tag{46}$$

We define an auxiliary variable $z \triangleq -(1+2\mathcal{W}_2(\nu)), z > 1$. Eq. (46) is reduced to

$$\frac{z^2-1}{2z} = \ln z. \tag{47}$$

Finally, we only need to check whether the function $g(z) \triangleq (z^2-1)/2z - \ln z$, $z > 1$ is always greater than 0. Taking the derivative of $g(z)$, we have

$$g'(z) = \frac{1}{2}(1 + \frac{1}{z^2}) - \frac{1}{z} > 0, \tag{48}$$

so $g(z)$ is a monotonically increasing function. Consequently, the minimum point is 0 when $z = 1$. Thus, equality (42) doesn't hold, which contradicts the assumption. It means that $\mathcal{W}_1(\nu) + \mathcal{W}_2(\nu)$ is always decreasing. Therefore, the absolute value of $\mathcal{W}_2'(\nu)$ is bigger than the one of $\mathcal{W}_1'(\nu)$. ■

D. Proof of Lemma 3

Proof: From the proof of Lemma 2, we know that when $\nu = -1/e$, then (42) holds, meaning that $\mathcal{W}_1'(\nu) = \mathcal{W}_2'(\nu)$. However, there is no definition for the derivative of the Lambert function when $\nu = -1/e$. Instead, considering $\nu = -1/e + d\nu$ when $d\nu \to 0$, we have

$$\begin{aligned}\lim_{d\nu\to 0}\frac{\mathcal{W}_1'(\nu)}{\mathcal{W}_2'(\nu)} &= \lim_{d\nu\to 0}\frac{(1+\mathcal{W}_2(\nu))\mathcal{W}_1(\nu)}{(1+\mathcal{W}_1(\nu))\mathcal{W}_2(\nu)} \\ &\overset{(a)}{=} \lim_{d\nu\to 0}\frac{\mathcal{W}_2'(\nu)\mathcal{W}_1(\nu) + (1+\mathcal{W}_2(\nu))\mathcal{W}_1'(\nu)}{\mathcal{W}_1'(\nu)\mathcal{W}_2(\nu) + (1+\mathcal{W}_1(\nu))\mathcal{W}_2'(\nu)} \\ &= \lim_{d\nu\to 0}\frac{\mathcal{W}_2'(\nu)}{\mathcal{W}_1'(\nu)}\end{aligned} \tag{49}$$

where (a) is according to the L'Hôpital's rule. Therefore, we conclude that

$$\lim_{d\nu\to 0}|\mathcal{W}_1'(\nu)| = \lim_{d\nu\to 0}|\mathcal{W}_2'(\nu)|. \tag{50}$$

■

E. Proof of Lemma 4

Proof: First, we can have the difference between the corresponding sums of the two variables, which is

$$\begin{aligned}\Sigma_1(\nu) - \Sigma_2(\nu) &= \frac{\rho_1(\nu)}{a_1} + \frac{\rho_2(\nu)}{a_2} - \left(\frac{\rho_1(\nu)}{a_2} + \frac{\rho_2(\nu)}{a_1}\right) \\ &= (\rho_1(\nu) - \rho_2(\nu))\left(\frac{1}{a_1} - \frac{1}{a_2}\right) < 0.\end{aligned} \tag{51}$$

Second, based on Remark 5 and Remark 6, we know that ν' needs to increase in order to increase $\Sigma_1(\nu')$. In this way, it is only possible that $\Sigma_1(\nu') = \Sigma_2(\nu'') = \lambda_T$. In conclusion, $\nu'' < \nu'$ under constraint $\Sigma_1(\nu') = \Sigma_2(\nu'') = \lambda_T$. ■

F. Proof of Lemma 5

Proof: When $\lambda_T = 2/a_1 + 2/a_2$, we have $\rho_1 = \rho_2 = 2$. Based on Remark 5 and Remark 6, there are only totally two KKT points. One is located at $[x_{1,2}'', x_{2,1}''] = [2/a_1, 2/a_2]$ and another one is $[x_{1,1}', x_{2,2}'] = [\rho_1'/a_1, \lambda_T - \rho_1'/a_1]$. Since they are KKT points, the first derivatives of these two points are both equal to zero. However, the second derivatives are different at these two points, which are

1) point $[x_{1,1}', x_{2,2}']$:

The second derivative of function $f_i(x_i)$ is

$$g_i(x_i) = a_i(2 - a_i x_i)e^{-a_i x_i}. \tag{52}$$

When $x_1 \le 2/a_1$, we know $g_1(x_1) \le 0$; when $x_2 \ge 2/a_2$, we have $g_2(x_2) \ge 0$, and vice versa. Therefore, there is no intersection point between $g_1(x_1)$ and $g_2(x_2)$ except the point: $x_1 = 2/a_1$ and $x_2 = 2/a_2$. Hence, we conclude that the second derivative of the objective function at point $[x_{1,1}', x_{2,2}']$ is not equal to zero.

2) point $[x_{1,2}'', x_{2,1}'']$:

Consider the convexity of function $f_i(x_i)$, since the transition point is $x_i = 2/a_i$, the second derivative is also zero at point $[x_{1,2}'', x_{2,1}'']$.

Consider the facts that $f_i(x_i) \ge 0$, $f_i(0) = 0$, and $f_i(x_i)$ is concave when $0 \le x_i \le 1/a_i$, and there are only two points whose first derivatives are zero. It can be concluded that point $[x_{1,2}'', x_{2,1}'']$ is a local optimal point. Consequently, we know that the corresponding function J_2 at this point is smaller than J_1, i.e., $J_1 > J_2$. ■

G. Proof of Lemma 6

Proof: Getting rid of the equality constraint in $\mathbb{OP}$ 2, the objective function can be simply written as

$$J = x_1 e^{-a_1 x_1} + (\lambda_T - x_1)e^{-a_2(\lambda_T - x_1)}.$$

Taking the derivative of J in terms of λ_T, we have

$$\begin{aligned}\frac{dJ}{d\lambda_T} &= e^{-a_2(\lambda_T - x_1)} - a_2(\lambda_T - x_1)e^{-a_2(\lambda_T - x_1)} \\ &= (1 - a_2 x_2)e^{-a_2 x_2}.\end{aligned} \tag{53}$$

Taking the derivative of J in terms of λ_T and substituting two KKT points $x_{2,2}' = \rho_2'/a_2$ and $x_{2,1}'' = \rho_1''/a_2$

into (53) respectively, we can obtain

$$\frac{dJ}{d\lambda_T}|_{x'_{2,2}} = (1-\rho'_2)e^{-\rho'_2} = \mathcal{W}_2(\nu' e)e^{-\mathcal{W}_2(\nu' e)-1} = \nu'$$
$$\frac{dJ}{d\lambda_T}|_{x''_{2,1}} = (1-\rho''_1)e^{-\rho''_1} = \mathcal{W}_1(\nu'' e)e^{-\mathcal{W}_1(\nu'' e)-1} = \nu''.$$

Based on Lemma 4 we know that the decreasing speed of objective function J at point $x'_{2,2}$ is less than the one at $x''_{2,1}$. In addition, from Lemma 5 we can conclude that $J_1 > J_2$ is always true for $\lambda_T \geq 2/a_1 + 2/a_2$ when $a_1 < a_2$. ■

H. Proof of Theorem 2

Proof: Consider a KKT point x^*. There exits a ν'' such that there is one variable in Region II, say $x^*_{j,2}$ and the left $N-1$ variables are all in Region I according to Lemma 1. Consider one variable $x^*_{k,1}$ (Region I). Assume $a_k > a_j$. Then, the KKT point is $[x^*_{1,1}, \ldots, x^*_{j,2}, \ldots, x^*_{k,1}, \ldots, x^*_{1,N}]$. With fixing all variables except the two variables indexed by j,k, the optimization problem with N variables is reduced to $\mathbb{OP}$ 2. The KKT point can be expressed by

$$\begin{aligned}\boldsymbol{x}^* &= [x^*_{1,1}, \ldots, x^*_{j,2}, \ldots, x^*_{k,1}, \ldots, x^*_{1,N}] \\ &= [\rho''_1/a_1, \ldots, \rho''_2/a_j, \ldots, \rho''_1/a_k, \ldots, \rho''_1/a_N]. \quad (54)\end{aligned}$$

Applying the result of Lemma 6, we know that a better point (with a higher objective value) should be

$$\begin{aligned}\boldsymbol{x}' &= [x^*_{1,1}, \ldots, x^*_{j,1}, \ldots, x^*_{k,2}, \ldots, x^*_{1,N}] \\ &= [\rho''_1/a_1, \ldots, \rho'_1/a_j, \ldots, \rho'_2/a_k, \ldots, \rho''_1/a_N], \quad (55)\end{aligned}$$

which implies that point $\boldsymbol{x}^*$ is not the global solution of $\mathbb{OP}$ 1. Note that point $\boldsymbol{x}'$ is not a KKT point of $\mathbb{OP}$ 1, since $\nu' \neq \nu''$. After enumerating all $\{a_i\}$, it is concluded that only the variable with $a_{\max}$ should be in Region II. ■

REFERENCES

[1] S. Lu and Z. Wang, "Throughput maximization over frequency-selective communication networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, New Orleans, LA, USA, Mar. 2015, pp. 2008–2013.

[2] J. Lee, J. G. Andrews, and D. Hong, "Spectrum-sharing transmission capacity," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3053–3063, Sep. 2011.

[3] S. Weber, J. G. Andrews, and N. Jindal, "An overview of the transmission capacity of wireless networks," *IEEE Trans. Commun.*, vol. 58, no. 12, pp. 3593–3604, Dec. 2010.

[4] J. G. Andrews, X. Zhang, G. D. Durgin, and A. K. Gupta, "Are we approaching the fundamental limits of wireless network densification?" *IEEE Commun. Mag.*, vol. 54, no. 10, pp. 184–190, Oct. 2016.

[5] K. Stamatiou, P. Casari, and M. Zorzi, "The throughput of underwater networks: Analysis and validation using a ray tracing simulator," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1108–1117, Mar. 2013.

[6] S. Weber and J. G. Andrews, "Transmission capacity of wireless networks," *Found. Trends Netw.*, vol. 5, nos. 2–3, pp. 109–281, Jan. 2012.

[7] R. K. Ganti and M. Haenggi, "Asymptotics and approximation of the SIR distribution in general cellular networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2130–2143, Mar. 2016.

[8] Z. Tong, H. Lu, M. Haenggi, and C. Poellabauer, "A stochastic geometry approach to the modeling of DSRC for vehicular safety communication," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 5, pp. 1448–1458, May 2016.

[9] H. H. Yang, J. Lee, and T. Q. S. Quek, "Heterogeneous cellular network with energy harvesting-based D2D communication," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1406–1419, Feb. 2016.

[10] J. G. Andrews, T. Bai, M. N. Kulkarni, A. Alkhateeb, A. K. Gupta, and R. W. Heath, Jr., "Modeling and analyzing millimeter wave cellular systems," *IEEE Trans. Commun.*, vol. 65, no. 1, pp. 403–430, Jan. 2017.

[11] X. Yu, J. Zhang, M. Haenggi, and K. B. Letaief, "Coverage analysis for millimeter wave networks: The impact of directional antenna arrays," *IEEE J. Select. Areas Commun.*, vol. 35, no. 7, pp. 1498–1512, Jul. 2017.

[12] N. Jindal, J. G. Andrews, and S. Weber, "Bandwidth partitioning in decentralized wireless networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5408–5419, Dec. 2008.

[13] T. Samarasinghe, H. Inaltekin, and J. S. Evans, "On optimal downlink coverage in poisson cellular networks with power density constraints," *IEEE Trans. Commun.*, vol. 62, no. 4, pp. 1382–1392, Apr. 2014.

[14] Z. Liu, T. Peng, H. Chen, and W. Wang, "Optimal D2D user allocation over multi-bands under heterogeneous networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2012, pp. 1339–1344.

[15] L. Wu, Y. Zhong, W. Zhang, and M. Haenggi, "Scalable transmission over heterogeneous networks: A stochastic geometry analysis," *IEEE Trans. Veh. Technol*, vol. 66, no. 2, pp. 1845–1859, Feb. 2017.

[16] S. P. Weber, X. Yang, J. Andrews, and G. de Veciana, "Transmission capacity of wireless ad hoc networks with outage constraints," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4091–4102, Dec. 2005.

[17] K. Huang, V. Lau, and Y. Chen, "Spectrum sharing between cellular and mobile ad hoc networks: Transmission-capacity trade-off," *IEEE J. Select. Areas Commun.*, vol. 27, no. 7, pp. 1256–1267, Sep. 2009.

[18] X. Lin, R. Ratasuk, A. Ghosh, and J. G. Andrews, "Modeling, analysis, and optimization of multicast device-to-device transmissions," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4346–4359, Aug. 2014.

[19] S. Lu, Z. Wang, Z. Wang, and S. Zhou, "Underwater acoustic network with random access: A physical layer perspective," in *Proc. 8th ACM Int. Conf. Underwater Netw. Syst.*, Kaohsiung, Taiwan, Nov. 2013, no. 34.

[20] S. Lu, Z. Wang, Z. Wang, and S. Zhou, "Throughput of underwater wireless ad hoc networks with random access: A physical layer perspective," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6257–6268, Nov. 2015.

[21] L. Pu *et al.*, "Impact of real modem characteristics on practical underwater MAC design," in *Proc. OCEANS*, May 2012, pp. 1–6.

[22] J. Kibiłda, N. J. Kaminski, and L. A. DaSilva, "Radio access network and spectrum sharing in mobile networks: A stochastic geometry perspective," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2562–2575, Apr. 2017.

[23] R. Diamant, P. Casari, F. Campagnaro, and M. Zorzi, "Leveraging the nearâŁ"far effect for improved spatial-reuse scheduling in underwater acoustic networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1480–1493, Mar. 2017.

[24] Y.-C. Jonga. (2012). *An Efficient Global Optimization Algorithm for Nonlinear Sum-of-Ratios Problem*. [Online]. Available: http://www.optimizationonline.org

[25] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009.

[26] J. Kokkoniemi and J. Lehtomäki, and M. Juntti, "Stochastic geometry analysis for mean interference power and outage probability in TH_z networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3017–3028, May 2017.

[27] C.-H. Liu and J. G. Andrews, "Ergodic spatial throughput of wireless ad hoc networks with Markovian fading channels," in *Proc. Int. Symp. Model. Optim. Mobile, Ad Hoc Wireless Netw (WiOpt)*, May 2011, pp. 366–371.

[28] S. Park and R. W. Heath, Jr., "Frequency selective hybrid precoding in millimeter wave OFDMA systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, p. 1–6.

[29] A. Lozano and D. Porrat, "Non-peaky signals in wideband fading channels: Achievable bit rates and optimal bandwidth," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 246–257, Jan. 2012.

[30] C. Polprasert, J. A. Ritcey, and M. Stojanovic, "Capacity of OFDM systems over fading underwater acoustic channels," *IEEE J. Ocean. Eng.*, vol. 36, no. 4, pp. 514–524, Oct. 2011.

[31] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3029–3056, Sep. 2015.

[32] H. ElSawy, E. Hossain, and M. Haenggi, "Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 996–1019, 3rd Quart., 2013.

[33] M. Rebato, J. Park, P. Popovski, E. D. Carvalho, and M. Zorzi, "Stochastic geometric coverage analysis in mmWave cellular networks with a realistic channel model," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–6.
[34] M. Rebato, M. Mezzavilla, S. Rangan, F. Boccardi, and M. Zorzi, "Understanding noise and interference regimes in 5G millimeter-wave cellular networks," in *Proc. 22nd Eur. Wireless Conf.*, May 2016, pp. 1–5.
[35] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge, U. K.: Cambridge Univ. Press, 2004.
[36] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the lambertW function," *Adv. Comput. Math.*, vol. 5, no. 1, pp. 329–359, Dec. 1996.
[37] M. K. Hanawal, E. Altman, and F. Baccelli, "Stochastic geometry based medium access games in wireless ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 11, pp. 2146–2157, Dec. 2012.
[38] F. Brah, A. Zaidi, J. Louveaux, and L. Vandendorpe, "On the Lambert-W function for constrained resource allocation in cooperative networks," *EURASIP J. Wireless Commun. Netw.*, vol. 2011, p. 19, Dec. 2011.
[39] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the Gaussian collision channel," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1971–1988, Jul. 2001.
[40] R. Comroe and D. Costello, "ARQ schemes for data transmission in mobile radio systems," *IEEE J. Sel. Areas Commun.*, vol. 2, no. 4, pp. 472–481, Jul. 1984.
[41] A. Rajanna, I. Bergel, and M. Kaveh, "Performance analysis of rateless codes in an ALOHA wireless Ad Hoc network," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6216–6229, Nov. 2015.
[42] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge, U. K.: Cambridge Univ. Press, 2005.
[43] Y. Tian, S. Lu, and C. Yang, "Macro-pico amplitude-space sharing with optimized Han-Kobayashi coding," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4404–4415, Oct. 2013.
[44] C.-H. Liu and J. G. Andrews, "Ergodic transmission capacity of wireless Ad Hoc networks with interference management," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2136–2147, Jun. 2012.
[45] C. Xing, Y. Jing, S. Wang, S. Ma, and H. V. Poor. (2018). "New viewpoint and algorithms for water-filling solutions in wireless communications." [Online]. Available: https://arxiv.org/abs/1808.01707

Songtao Lu (S'10–M'18) received the Ph.D. degree in electrical and computer engineering from Iowa State University, Ames, IA, USA, in 2018. He is currently a Post-Doctoral Associate with the Department of Electrical and Computer Engineering, University of Minnesota Twin Cities, Minneapolis, MN, USA. His primary research interests include wireless communications, optimization, and machine learning. He received the Graduate and Professional Student Senate Research Award from Iowa State University in 2015, the Research Excellence Award from the Graduate College, Iowa State University, in 2017, and the Student Travel Award from the 20th International Conference on Artificial Intelligence and Statistics in 2017.

Zhengdao Wang (S'00–M'02–SM'08–F'16) received the B.S. degree in electronic engineering and information science from the University of Science and Technology of China in 1996, the M.Sc. degree in electrical and computer engineering from the University of Virginia, Charlottesville, VA, USA, in 1999, and the Ph.D. degree in electrical and computer engineering from the University of Minnesota, Minneapolis, MN, USA, in 2002. He is currently with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA. His interests are in the areas of signal processing, communications, information theory, and machine learning. He was a co-recipient of the IEEE Signal Processing Magazine Best Paper Award in 2003, the IEEE Communications Society Marconi Paper Prize Award in 2004, and the EURASIP Journal on Advances in Signal Processing Best Paper Award in 2009. He served as an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE SIGNAL PROCESSING LETTERS, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE Signal Processing Society Online Video Library, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and *ZTE Communications*.