Maximizing Spatial α-Fairness in Multi-Tier Multi-Rate Spatial Aloha Networks

Yujie Xu, *Student Member, IEEE,* Ben Liang, *Fellow, IEEE,*
Gary Boudreau, *Senior Member, IEEE,* and S. Hossein Seyedmehdi

***Abstract*—We consider the maximization of α-fair utility in a generalized spatial Aloha network consisting of multiple tiers of transmitter-receiver (T-R) pairs each forming a Poisson bipolar process. The tiers are distinguished by transmission power and the T-R distance. Multi-rate communication between the T-R pairs is facilitated by multiple received signal-to-interference ratio (SIR) thresholds. We aim to optimize the transmission probability of each tier. This results in a complex non-convex optimization problem due to intra-tier and cross-tier interference. We propose a solution termed Minorize-Maximization with Tier Separation (MMTS), through designing an iterative sequence of lower bound problems that can be decomposed into tier-separable one-dimensional convex optimization problems and solved efficiently. Specific solutions are derived for the cases $0 \leq \alpha < 1$, $\alpha = 1$, and $\alpha > 1$. We show the convergence of MMTS to the objective value of a Karush-Kuhn-Tucker (KKT) point of the original problem and further identify several conditions under which it finds the global optimum. Numerical results demonstrate the near optimality of MMTS and substantial performance advantage over existing alternatives.**

***Index Terms*—Spatial Aloha networks, utility maximization, transmission probability, minorize-maximization, tier separation.**

I. Introduction

Direct transmission between devices in proximity is a well promoted paradigm to allow ad hoc network access and to increase wireless spectrum utilization [1]. In particular, device-to-device communication has become an important aspect of next-generation wireless standardization [2]. One of the main challenges in direct communication among devices is how to allocate the common wireless spectrum in an efficient and fair manner over a large-scale network, such as in the Internet-of-Things environment.

Utility maximization in random access networks has received much attention in the literature. The authors of [3] first studied the problem of maximizing network α-fair utility for $\alpha > 1$ using the protocol model. They showed that the problem can be recast as a convex optimization problem and proposed a distributed scheduling based on Lagrangian dual decomposition. Then [4] further proposed a method based on coordinate descent to solve the utility maximization problem by observing that the problem is convex in the transmission probability of each T-R pair separately for all α values. In addition, [5] studied the utility maximization problem in the single signal-to-interference-plus-noise ratio (SINR)-threshold model instead of the protocol model, and show that the single SINR threshold model yields higher throughput. However, these works all required that the exact position of each node is known, which may be difficult to obtain. This inspires researchers to study random access networks with a random topology defined only by its statistics.

In the presence of a large number of direct communication pairs, acquiring the exact topology information is a prohibitive task, especially in networks with high mobility. Thus, research on wireless network models with random topology has drawn much attention. In the celebrated *spatial Aloha* model [6], only statistical information of the topology is available. Each transmitter in the network randomly transmits following the slotted Aloha medium access control (MAC) protocol. Both the Poisson point process (PPP) and Poisson bipolar process (PBP) are commonly used to model the location of transmitters and receivers in spatial Aloha networks. In the PBP model, the transmitters form a PPP, and each transmitter is paired with a dedicated receiver at some distance away.

An interesting design problem in spatial Aloha networks is to optimize the transmission probability. This is a challenging problem, often with a non-convex objective function due to signal interference. In [6]–[11], this problem is studied where all transmitters are assumed to use the same transmission probability if the exact location of nodes is unknown. This model is suitable only when the network is uniform. In many practical scenarios, the transmitters may have different powers and the T-R distance may be different for different T-R pairs, so that the transmitters should use different transmission probabilities. Some previous studies have addressed this problem in static Aloha networks [3]–[5], [12]–[14], but none of them allows randomness in the network topology.

Furthermore, most existing works assume single-rate communication either based on a single received SIR threshold [5]–[11], such that the data rate is $\log(1 + \mathrm{Th})$ if the received SIR is above some threshold Th, and is zero otherwise; or based on the "protocol model" [3], [4], [12]–[14] wherein the data rate is some fixed term if the nearby transmitters do not transmit. In terms of physical implementation, both cases correspond to the usage of only a single modulation-coding scheme at the transmitter. Such a model simplifies mathematical analysis but has limited application in more sophisticate multi-rate systems.

In this work, we extend the spatial Aloha model of [7]–[9]

Y. Xu and B. Liang are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, M5S 3G4, Canada (e-mail: {yujiexu, liang}@ece.utoronto.ca). G. Boudreau and S. H. Seyedmehdi are with Ericsson Canada, Ottawa, ON, K2K 2V6, Canada (e-mail: {gary.boudreau, hossein.seyedmehdi}@ericsson.com).

This work was supported in part by research grants from Ericsson and the Natural Sciences and Engineering Research Council of Canada.

to the multi-tier, multi-rate regime. Each tier of the network is defined by the power of the transmitters and the T-R distance. We consider both intra-tier and inter-tier interference. We also accommodate multi-rate communication through multiple received SIR thresholds. We aim to optimize the transmission probability of each tier, to maximize a general α-fair utility function.

Our main contributions are as follows:

- We first derive a closed-form expression of the average throughput of T-R pairs in each tier, which takes into account the random location of multi-tier interferers and multiple received SIR thresholds. This is then used in the formulation of an optimization problem to maximize the network-wide spatial α-fair utility, which is generally non-convex.
- We propose a computationally efficient iterative algorithm, termed Minorize-Maximization with Tier Separation (MMTS), to address this optimization problem. By exploring the partial-convexity and partial-concavity of the objective function when $0 \leq \alpha \leq 1$ and $\alpha > 1$, respectively, we develop special lower bounds to the α-fair objective, which we dynamically update in each iteration through solving an optimization sub-problem. Furthermore, the lower bounds are designed so that these sub-problems can be decomposed into one-dimensional convex optimization problems that are separable according to T-R tiers, which drastically reduces the computational complexity.
- We show that MMTS converges to the objective value of a KKT point of the optimization problem. We further provide various sufficient conditions under which the KKT point is a global optimizer. Numerical evaluation results demonstrate that MMTS is near optimal over a wide range of parameter settings, and it substantially outperforms existing alternatives.

The rest of the paper is organized as follows. In Section II, we summarize the related work. In Section III, we present the system model and formulate our optimization problem. In Sections IV, we derive the average throughput of communication pairs in different tiers. In Section V and VI, we present MMTS and discuss its convergence and optimality, respectively. In Section VII, we present numerical evaluation results. Conclusions are drawn in Section VIII.

II. Related Works

There has been a large amount of research into ad hoc, device-to-device, or direct-transmission networks that employ the Aloha MAC protocol [3]–[21]. Among them, [3]–[5], [12]–[14] consider a fixed transmitter/receiver topology, [6], [17]–[21] use the PPP model, [7], [10], [11], [15], [16] use the PBP model, while [8], [9] use the PBP model with partial topology information. The networks in the latter three groups are commonly termed spatial Aloha networks. In this section, we briefly review works in optimizing transmission probabilities in spatial aloha networks. We further describe the Minorize-Maximization framework and its application in communication systems.

TABLE I: Table of Notations

Notation	Description
N	Number of tiers
λ_n	Intensity of PBP of transmitters in tier n
R_n	T-R distance in tier n
P_n	Transmission power of transmitter in tier n
p_n	Transmission probability of transmitter in tier n
γ	Pathloss exponent
L	Number of SIR thresholds to modulate the received signal
T_l	lth SIR threshold
α	Fairness index in the utility function

A. Spatial Aloha Networks

Several studies consider the optimization of transmission probability in spatial Aloha [6]–[11]. However, in these works all transmitters are assumed to use the same transmission probability when the exact location of nodes is unknown. In our work, we design different transmission probabilities for different tiers of the network based on T-R distance and transmission power. Our numerical results show that this can lead to substantial performance improvement.

Furthermore, all of [6]–[11] use a simple single-rate communication model based on a single received SIR threshold, with [7] further considering the Shannon-rate upper bound, while in this work we allow multiple SIR thresholds for multi-rate communication between each T-R pair. Finally, the performance objectives of these works are narrow: [6], [7], [10], [11] focus on the sum throughput, while [8], [9] concern the log utility. All of these objectives are special cases of the general α-fair utility in this work.

B. Minorize-Maximization Framework

The Minorize-Maximization (MM) framework is *not a specific algorithm*, but a general approach to construct an algorithm. It has been used to efficiently solve some non-linear optimization problems in wireless networking design [22], [23]. One of its advantages is the avoidance of matrix inversion in many cases, which reduces computational complexity. The key design to apply MM in maximization problems is to develop appropriate lower bounds for the objective that are iteratively updated. Examples include Taylor series expansion [22] and log functions [23]. In this work, we develop unique lower bounds specific to α-fairness in spatial Aloha networks, for different values of α. Unlike general MM-based heuristics, these lower bounds are designed to be iteratively updated to converge to the objective value of a KKT point of our utility maximization problem. To our best knowledge, this is the first time such a solution has been constructed under the general MM framework.

III. System Model and Problem Formulation

In this section, we describe multi-tier, multi-rate spatial aloha network. We further consider the fairness in the network, and formulate the problem of maximizing the α-fair utility of the network. Important notations throughout this paper are summarized in Table I.

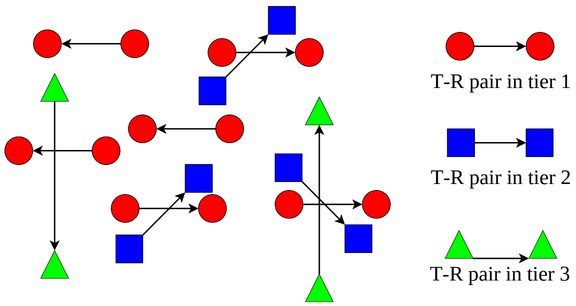


Fig. 1: Multi-tier spatial Aloha network

A. Multi-tier Spatial Aloha Network

Consider a spatial random network in two-dimensional Euclidean space consisting of multiple simple T-R pairs with fixed transmission power, communicating over a shared channel as illustrated in Fig.1. The T-R pairs are differentiated into N tiers, defined by their transmission power and T-R distance. Each tier independently forms a PBP, i.e., the tier n transmitters form a PPP with intensity λ_n, denoted by Φ_n^{Tx}, and each transmitter is associated with a receiver that is uniformly randomly located on a circle of fixed radius R_n centered at the transmitter. We observe that by this definition of the PBP, the tier n receivers also form a PPP with intensity λ_n, since they are i.i.d. marks of Φ_n^{Tx}.

Due to the large number of T-R pairs in the network, we assume that each pair only acquires its own T-R distance, transmission power, and spatial statistics of the interfering T-R pairs. The T-R pairs employ the slotted Aloha MAC protocol due to the lack of topology information [7]. The transmission probability of each tier n transmitter is denoted by p_n. Then, the independent thinning property of a PPP implies that the *active* transmitters in tier n is a PPP with intensity $p_n\lambda_n$. We denote by P_n the fixed transmission power of active tier n transmitters.

Note that this N-tier spatial Aloha network model is general. For example, if in practice the power of transmitters in a tier becomes unequal, this tier can be further split into multiple tiers such that the above definition of a tier is upheld. The total number of tiers can be set large enough based on the required precision of system modeling and analysis. In this work, our analysis proceeds assuming that the tiers are already given.

B. Multiple SIR Thresholds and Average Throughput

The received power of the receiver located at $\mathbf{y}$ from the transmitter located at $\mathbf{x}$ is given by

$$P_{\mathbf{xy}} = \frac{P_{\mathbf{x}} h_{\mathbf{xy}}}{|\mathbf{x}-\mathbf{y}|^{\gamma}}, \tag{1}$$

where $P_{\mathbf{x}}$ is the transmission power, $h_{\mathbf{xy}}$ is the channel power gain under Rayleigh fading, and γ is the path loss exponent where $\gamma > 2$. We assume that $h_{\mathbf{xy}}$ is i.i.d. with unit mean and independent of Φ_n^{Tx} for all n.

Because of spatial stationarity, we may focus on an arbitrary T-R pair in tier n, termed the *typical pair*. We further assume the transmitter and the receiver in the typical pair are situated at $\mathbf{x}_{n0}$ and origin $\mathbf{0}$, respectively. Then, the SIR of the receiver in the typical pair is given by

$$\mathrm{SIR}_n = \frac{\frac{P_n h_{\mathbf{x}_{n0}\mathbf{0}}}{{R_n}^{\gamma}}}{\sum_{k=1,k\neq n}^{N}\sum_{\mathbf{x}\in\Phi_k^{\mathrm{Tx}}}\frac{P_k h_{\mathbf{x0}}}{|\mathbf{x}|^{\gamma}} + \sum_{\mathbf{x}\in\Phi_n^{\mathrm{Tx}}\setminus\{\mathbf{x}_{n0}\}}\frac{P_n h_{\mathbf{x0}}}{|\mathbf{x}|^{\gamma}}},$$

where $\frac{P_n h_{\mathbf{x}_{n0}\mathbf{0}}}{{R_n}^{\gamma}}$ is the received power at the DRx of the typical pair, $\sum_{\mathbf{x}\in\Phi_n^{\mathrm{Tx}}\setminus\{\mathbf{x}_{n0}\}}\frac{P_n h_{\mathbf{x0}}}{|\mathbf{x}|^{\gamma}}$ is the interference from other T-R pairs in the same tier as the typical pair, and $\sum_{k=1,k\neq n}^{N}\sum_{\mathbf{x}\in\Phi_k^{\mathrm{Tx}}}\frac{P_k h_{\mathbf{x0}}}{|\mathbf{x}|^{\gamma}}$ is the interference from the T-R pairs in the other tiers.

We assume the system is interference limited. The receivers use L SIR thresholds to demodulate the received signal, denoted by $\mathcal{T} = \{T_1, \cdots, T_L\}$. Without loss of generality, we assume that $T_i < T_j$ if $i < j$. We further assume that the transmission rate is r_l when a T-R pair's SIR is between T_l and T_{l+1}, and r_L when the SIR is larger than or equal to T_L. And we set $\mathsf{r}_0 = 0$ for notational convenience. Since $T_l > T_{l-1}$, we have $\mathsf{r}_l > \mathsf{r}_{l-1}$. In the above, we have normalized the channel bandwidth to one.

The average throughput of the typical pair in tier n is given by

$$\begin{aligned} r_n &= p_n\left[\sum_{l=1}^{L-1}\mathbb{P}(T_l \le \mathrm{SIR}_n < T_{l+1})\mathsf{r}_l + \mathbb{P}(\mathrm{SIR}_n \ge T_L)\mathsf{r}_L\right] \\ &= p_n\sum_{l=1}^{L} a_l\mathbb{P}(\mathrm{SIR}_n \ge T_l), \end{aligned} \tag{2}$$

where $a_l = \mathsf{r}_l - \mathsf{r}_{l-1}$.

C. Problem Statement

Similar to [3]–[5], we define the spatial α-fair utility as

$$U(\mathbf{p}) = \begin{cases} \sum_{n=1}^{N}\frac{(\lambda_n r_n)^{1-\alpha}}{1-\alpha} & \text{if } \alpha\neq 1, \\ \sum_{n=1}^{N}\log(\lambda_n r_n) & \text{if } \alpha = 1, \end{cases} \tag{3}$$

where $\mathbf{p} = [p_n]_{N\times 1}$, and formulate our main optimization problem as

$$\mathcal{P}: \quad \max_{\mathbf{p}} \quad U(\mathbf{p}) \tag{4}$$

$$\text{s.t.} \quad p_{n,\min} \le p_n \le p_{n,\max}, \quad 1\le n\le N, \tag{5}$$

where $p_{n,\max}$ and $p_{n,\min}$ are the maximum and minimum transmission probabilities of transmitters in tier n. Furthermore, we define $\mathfrak{P}$ as the feasible set of problem $\mathcal{P}$.

Thus, we can adjust the fairness among different tiers by tuning the value of α. It should be noted that, since the T-R pairs in each tier have the same T-R distance, transmission power, and interference statistics, they have the same expected average throughput. Therefore, although the objective of Problem $\mathcal{P}$ is formulated as a sum of utility over tiers, tuning the value of α can also adjust the fairness among individual T-R pairs. Generally, when α is set larger, maximizing the α-fair utility allows T-R pairs in unfavorable transmission conditions (e.g., long T-R distance or lower transmission power) to obtain more throughput, and thus the system becomes more fair. Specifically, when $\alpha = 0$, $U(\mathbf{p})$ degrades to the sum throughput over all T-R pairs; when $\alpha = 1$, maximizing $U(\mathbf{p})$ leads to the celebrated proportional fairness; and when $\alpha \to \infty$, maximizing $U(\mathbf{p})$ leads to max-min fairness.

The challenges of problem $\mathcal{P}$ are two-fold. First, the average throughput of communication pairs in different tiers (i.e., r_n) needs to be derived for the multi-tier, multi-rate scenario. Second, $\mathcal{P}$ is non-convex in most cases because of the non-concavity of its objective function. Thus, conventional convex solvers do not apply.

IV. Average Throughput Derivation

In this section, we derive a closed-form expression of the average throughput of the typical T-R pair. The probability of the event, that the received SIR of the receiver in the typical T-R pair is no less than T, denoted by SIR_n, is given by

$$\begin{aligned}\mathbb{P}(\text{SIR}_n \geq T) &= \mathbb{P}\left(h_{\mathbf{x}_{n0}\mathbf{0}} \geq \frac{TI{R_n}^{\gamma}}{P_n}\right)\\ &\overset{(a)}{=} \mathbb{E}_I\left[\exp\left(-\frac{TI{R_n}^{\gamma}}{P_n}\right)\right],\end{aligned} \tag{6}$$

where I is the sum intra-tier and inter-tier interference received by the receiver in the typical T-R pair, given by $I = \sum_{k=1,k\neq n}^{N}\sum_{\mathbf{x}\in\Phi_k^{\text{Tx}}}\frac{P_k h_{\mathbf{x}\mathbf{0}}}{|\mathbf{x}|^{\gamma}} + \sum_{\mathbf{x}\in\Phi_n^{\text{Tx}}\setminus\{\mathbf{x}_{n1}\}}\frac{P_n h_{\mathbf{x}\mathbf{0}}}{|\mathbf{x}|^{\gamma}}$, and (a) is based on the distribution of $h_{\mathbf{x}_{n0}\mathbf{0}}$ and its independence of I.

As shown in Appendix A, the Laplace transform of I is given by

$$\mathbb{L}_I(s) = \prod_{j=1}^{N}\exp\left(-p_j\lambda_j\pi(sP_j)^{\frac{2}{\gamma}}\Gamma(1-\frac{2}{\gamma})\Gamma(1+\frac{2}{\gamma})\right). \tag{7}$$

Substituting (7) into (6), we have

$$\mathbb{P}(\text{SIR}_n \geq T) = \exp\left(-\sum_{j=1}^{N}p_j\lambda_j{R_n}^2\left(\frac{P_j}{P_n}\right)^{2/\gamma}C\right), \tag{8}$$

where $C = \pi T^{\frac{2}{\gamma}}\Gamma(1-\frac{2}{\gamma})\Gamma(1+\frac{2}{\gamma})$.

Substituting (8) into (2) and simplifying, we find the average throughput of the typical T-R pair:

$$r_n = p_n\sum_{l=1}^{L}a_l\exp\left(-m_{nl}\sum_{j=1}^{N}p_j\lambda_jP_j^{'}\right), \tag{9}$$

where $C_l = \pi T_l^{\frac{2}{\gamma}}\Gamma(1-\frac{2}{\gamma})\Gamma(1+\frac{2}{\gamma})$, $P_n^{'} = {P_n}^{2/\gamma}$, and $m_{nl} = \frac{{R_n}^2C_l}{P_n^{'}}$. Note that even though C_l contains the gamma function, it is a positive constant, so the right-hand side of (9) is in closed-form with respect to $\mathbf{p}$. This contributes to the tier-separability and efficiency of MMTS.

V. Minorize-Maximization with Tier Separation

In this section, we present MMTS to solve problem $\mathcal{P}$. We first develop three lower bounds for objective (4), for $0 \leq \alpha < 1$, $\alpha = 1$, and $\alpha > 1$. Then, we develop lower bound problems for these three cases, and show how they can be decomposed to multiple one-dimensional convex optimization problems that are separable according to the T-R tiers, which can be solved either in closed-form or otherwise efficiently. Finally, we explain how these tier-separable solutions can be employed in a recursive MM framework to address the original problem $\mathcal{P}$.

A. Lower Bound Problem for $0 \leq \alpha < 1$

Though objective (4) is not concave, it has a special *partially-convex* structure when $0 \leq \alpha < 1$. We take advantage of this special structure and develop a lower bound as stated in Lemma 1.

Lemma 1: When $0 \leq \alpha < 1$, for all $\mathbf{p} = [p_n]_{N\times1}, \mathbf{p}^{t-1} = [p_n^{t-1}]_{N\times1} \in \mathfrak{P}$,

$$\begin{aligned}&\frac{\left(\lambda_np_n\sum_{l=1}^{L}a_l\exp(-m_{nl}\sum_{j=1}^{N}P_j^{'}\lambda_jp_j)\right)^{1-\alpha}}{1-\alpha}\\ &\geq f_n^{t-1} - d_n^{t-1}\sum_{j=1}^{N}P_j^{'}\lambda_j(p_j - p_j^{t-1}) + e_n^{t-1}\log\frac{p_n}{p_n^{t-1}},\end{aligned} \tag{10}$$

where

$$\begin{aligned}d_n^{t-1} &= \frac{\left(\lambda_np_n^{t-1}\right)^{1-\alpha}\sum\limits_{l=1}^{L}a_lm_{nl}\exp(-m_{nl}\sum\limits_{j=1}^{N}P_j^{'}\lambda_jp_j^{t-1})}{\left(\sum\limits_{l=1}^{L}a_l\exp(-m_{nl}\sum\limits_{j=1}^{N}P_j^{'}\lambda_jp_j^{t-1})\right)^{\alpha}},\\ e_n^{t-1} &= \left(\lambda_np_n^{t-1}\sum_{l=1}^{L}a_l\exp(-m_{nl}\sum_{j=1}^{N}P_j^{'}\lambda_jp_j^{t-1})\right)^{1-\alpha},\\ f_n^{t-1} &= \frac{\left(\lambda_np_n^{t-1}\sum\limits_{l=1}^{L}a_l\exp(-m_{nl}\sum\limits_{j=1}^{N}P_j^{'}\lambda_jp_j^{t-1})\right)^{1-\alpha}}{1-\alpha},\end{aligned} \tag{11}$$

and the equality holds if $\mathbf{p} = \mathbf{p}^{t-1}$.

Proof: See Appendix B. ■

From Lemma 1, a lower bound of objective (4) is

$$\begin{aligned}\tilde{U}(\mathbf{p},\mathbf{p}^{t-1}) = \sum_{n=1}^{N}\Big(&-\sum_{j=1}^{N}d_j^{t-1}P_n^{'}\lambda_n(p_n - p_n^{t-1})\\ &+ e_n^{t-1}\log\frac{p_n}{p_n^{t-1}} + f_n^{t-1}\Big),\end{aligned}$$

for all $\mathbf{p}, \mathbf{p}^{t-1} \in \mathfrak{P}$, and $\tilde{U}(\mathbf{p},\mathbf{p}^{t-1}) = U(\mathbf{p})$ when $\mathbf{p} = \mathbf{p}^{t-1}$. Therefore, when $0 \leq \alpha < 1$, for any given $\mathbf{p}^{t-1} \in \mathfrak{P}$, we may consider the following optimization problem that lower bounds $\mathcal{P}$.

$$\mathcal{P}_{\text{LB}}^1: \quad \max_{\mathbf{p}} \quad \tilde{U}(\mathbf{p},\mathbf{p}^{t-1}) \tag{12}$$

$$\text{s.t.} \quad p_{n,\min} \leq p_n \leq p_{n,\max},\ 1 \leq n \leq N. \tag{13}$$

Remark: Note that the right-hand side (RHS) of (10) has been designed to be separable with respect to the optimization variables $\{p_n\}$. This enables our decomposition of $\mathcal{P}_{\text{LB}}^1$ into N separable one-dimensional convex optimization problem according to the T-R tiers. The nth problem, for $1 \le n \le N$, is given by

$$\mathcal{P}_{\text{LB}}^{1,n}: \quad \max_{p_n} \quad \tilde{U}_n(p_n) \tag{14}$$
$$\text{s.t.} \quad p_{n,\min} \le p_n \le p_{n,\max}, \tag{15}$$

where

$$\tilde{U}_n(p_n) = -\sum_{j=1}^{N} d_j^{t-1} P_n^{'} \lambda_n (p_n - p_n^{t-1}) + e_n^{t-1} \log \frac{p_n}{p_n^{t-1}} + f_n^{t-1}.$$

Problem $\mathcal{P}_{\text{LB}}^{1,n}$ has a closed-form optimal solution as stated in Lemma 2, whose proof is given in Appendix C.

Lemma 2: Problem $\mathcal{P}_{\text{LB}}^{1,n}$ is convex in p_n, and its optimal solution is

$$\tilde{p}_n^* = \left[\frac{e_n^{t-1}}{P_n^{'} \lambda_n \sum_{j=1}^{N} d_j^{t-1}} \right]_{p_{n,\min}}^{p_{n,\max}}, \tag{16}$$

where $[x]_{x_{\min}}^{x_{\max}} = \min\{\max\{x, x_{\min}\}, x_{\max}\}$.

Since $\mathcal{P}_{\text{LB}}^{1}$ is separable into convex problems, itself is also a convex problem. This facilitates our analysis of the convergence of MMTS as described in Section VI.

B. Lower Bound Problem for $\alpha = 1$

When $\alpha = 1$, the objective of $\mathcal{P}$ is

$$U(\mathbf{p}) = \sum_{n=1}^{N} \log(\lambda_n p_n) + \log(\sum_{l=1}^{L} a_l \exp(-m_{nl} \sum_{j=1}^{N} P_j^{'} \lambda_j p_j).$$

We note that $\log(\sum_{l=1}^{L} a_l \exp(-m_{il} \sum_{j=1}^{N} x_j))$, for all l and $a_l > 0$, is convex in $\mathbf{x} = [x_j]_{N \times 1}$ based on the convexity of the *LogSumExp* function [24]. Hence, for all $\mathbf{p} = [p_n]_{N \times 1}, \mathbf{p}^{t-1} = [p_n^{t-1}]_{N \times 1} \in \mathfrak{P}$, we have

$$\begin{aligned} &\log(\sum_{l=1}^{L} a_l \exp(-m_{nl} \sum_{j=1}^{N} P_j^{'} \lambda_j p_j)) \\ &\ge -g_n^{t-1} \sum_{j=1}^{N} P_j^{'} \lambda_j (p_j - p_j^{t-1}) + h_n^{t-1}, \end{aligned} \tag{17}$$

where

$$\begin{aligned} g_n^{t-1} &= \frac{\sum_{l=1}^{L} a_l m_{nl} \exp(-m_{nl} \sum_{j=1}^{N} P_j^{'} \lambda_j p_j^{t-1})}{\sum_{l=1}^{L} a_l \exp(-m_{nl} \sum_{j=1}^{N} P_j^{'} \lambda_j p_j^{t-1})}, \\ h_n^{t-1} &= \log \left(\sum_{l=1}^{L} a_l \exp(-m_{nl} \sum_{j=1}^{N} P_j^{'} \lambda_j p_j^{t-1}) \right), \end{aligned} \tag{18}$$

and the equality holds if $\mathbf{p} = \mathbf{p}^{t-1}$.

Hence, a lower bound of objective (4) is given by

$$\begin{aligned} \hat{U}(\mathbf{p}, \mathbf{p}^{t-1}) = &-\sum_{j=1}^{N} \sum_{n=1}^{N} g_j^{t-1} P_n^{'} \lambda_n (p_n - p_n^{t-1}) \\ &+ \sum_{n=1}^{N} \log(\lambda_n p_n) + \sum_{n=1}^{N} h_n^{t-1}, \end{aligned}$$

for all $\mathbf{p}, \mathbf{p}^{t-1} \in \mathfrak{P}$, and $\hat{U}(\mathbf{p}, \mathbf{p}^{t-1}) = U(\mathbf{p})$ when $\mathbf{p} = \mathbf{p}^{t-1}$. Therefore, when $\alpha = 1$, for any given $\mathbf{p}^{t-1} \in \mathfrak{P}$, we may consider the following optimization problem that lower bounds $\mathcal{P}$.

$$\mathcal{P}_{\text{LB}}^{2}: \quad \max_{\mathbf{p}} \quad \hat{U}(\mathbf{p}, \mathbf{p}^{t-1}) \tag{19}$$
$$\text{s.t.} \quad p_{n,\min} \le p_n \le p_{n,\max}, 1 \le n \le N. \tag{20}$$

Similarly to the case $1 \le \alpha < 1$, the RHS of (17) has been designed to be separable with respect to the optimization variables $\{p_n\}$. We can decompose $\mathcal{P}_{\text{LB}}^{2}$ into N separable one-dimensional convex optimization problem according to the T-R tiers. The nth problem, for $1 \le n \le N$, is given by

$$\mathcal{P}_{\text{LB}}^{2,n}: \quad \max_{p_n} \quad \hat{U}_n(p_n) \tag{21}$$
$$\text{s.t.} \quad p_{n,\min} \le p_n \le p_{n,\max}, \tag{22}$$

where

$$\hat{U}_n(p_n) = -\sum_{j=1}^{N} g_j^{t-1} P_n^{'} \lambda_n (p_n - p_n^{t-1}) + \log(\lambda_n p_n) + h_n^{t-1}.$$

Problem $\mathcal{P}_{\text{LB}}^{2,n}$ has a closed-form optimal solution as stated in Lemma 3.

Lemma 3: Problem $\mathcal{P}_{\text{LB}}^{2,n}$ is convex in p_n, and its optimal solution is

$$\hat{p}_n^* = \left[\frac{1}{P_n^{'} \lambda_n \sum_{j=1}^{N} g_j^{t-1}} \right]_{p_{n,\min}}^{p_{n,\max}}. \tag{23}$$

C. Lower Bound Problem for $\alpha > 1$

When $\alpha > 1$, We take advantage of a special *partially-concave* structure of objective (4) and develop the following lower bound as stated in Lemma 4. For notational convenience, we define $\alpha' = (N+1)(1-\alpha)$.

Lemma 4: When $\alpha > 1$, for all $\mathbf{p} = [p_n]_{N \times 1}, \mathbf{p}^{t-1} = [p_n^{t-1}]_{N \times 1} \in \mathfrak{P}$,

$$\begin{aligned} &\frac{\left(\lambda_n p_n \sum_{l=1}^{L} a_l \exp(-m_{nl} \sum_{j=1}^{N} P_j^{'} \lambda_j p_j)\right)^{1-\alpha}}{(1-\alpha)} \\ &\ge \frac{r_n^{t-1} {p_n}^{\alpha'}}{\alpha'} + \sum_{j=1}^{N} \sum_{l=1}^{L} \frac{s_{njl}^{t-1} \exp(-\alpha' m_{nl} P_j^{'} \lambda_j p_j)}{\alpha'}, \end{aligned} \tag{24}$$

where

$$\begin{aligned} r_n^{t-1} =& {\lambda_n}^{1-\alpha} {p_n^{t-1}}^{N(\alpha-1)} \left(\sum_{l=1}^{L} a_l \exp(-m_{nl} \sum_{j=1}^{N} P_j^{'} \lambda_j p_j^{t-1}) \right)^{1-\alpha} \\ s_{njl}^{t-1} =& \frac{a_l \exp\left(-m_{nl} \left(\sum_{k=1}^{N} P_k^{'} \lambda_k p_k^{t-1} - \alpha' P_j^{'} \lambda_j p_j^{t-1}\right)\right)}{\left(\lambda_n p_n^{t-1}\right)^{\alpha-1} \left(\sum_{l'=1}^{L} a_{l'} \exp(-m_{nl'} \sum_{k=1}^{N} P_k^{'} \lambda_k p_k^{t-1}) \right)^{\alpha}}, \end{aligned} \tag{25}$$

and the equality holds if $\mathbf{p} = \mathbf{p}^{t-1}$.

Proof: See Appendix D. ■

From Lemma 4, a lower bound of objective (4) is

$$\begin{aligned} &\bar{U}(\mathbf{p}, \mathbf{p}^{t-1}) \\ &= \sum_{n=1}^{N} \left(\frac{r_n^{t-1} p_n^{\alpha'}}{\alpha'} + \sum_{j=1}^{N} \sum_{l=1}^{L} s_{jnl}^{t-1} \frac{\exp(-\alpha' m_{jl} P_n^{'} \lambda_n p_n)}{\alpha'} \right), \end{aligned}$$

for all $\mathbf{p}, \mathbf{p}^{t-1} \in \mathfrak{P}$, and $\bar{U}(\mathbf{p}, \mathbf{p}^{t-1}) = U(\mathbf{p})$ when $\mathbf{p} = \mathbf{p}^{t-1}$. Therefore, when $\alpha > 1$, for any given $\mathbf{p}^{t-1} \in \mathfrak{P}$, we may consider the following optimization problem that lower bounds $\mathcal{P}$.

$$\mathcal{P}_{\mathrm{LB}}^{3}: \quad \max_{\mathbf{p}} \quad \bar{U}(\mathbf{p}, \mathbf{p}^{t-1}) \tag{26}$$

$$\text{s.t.} \quad p_{n,\min} \le p_n \le p_{n,\max}, 1 \le n \le N. \tag{27}$$

Again, similarly to the other cases of α, the RHS of (24) has been designed to be separable with respect to the optimization variables $\{p_n\}$. We can decompose P_{LB}^{3} into N separable one-dimensional convex optimization problem according to the T-R tiers. The nth problem, for $1 \le n \le N$, is given by

$$\mathcal{P}_{\mathrm{LB}}^{3,n}: \quad \max_{p_n} \quad \bar{U}_n(p_n) \tag{28}$$

$$\text{s.t.} \quad p_{n,\min} \le p_n \le p_{n,\max}, \tag{29}$$

where

$$\bar{U}_n(p_n) = \frac{r_n^{t-1} p_n^{\alpha'}}{\alpha'} + \sum_{j=1}^{N} \sum_{l=1}^{L} s_{jnl}^{t-1} \frac{\exp(-\alpha' m_{jl} P_n^{'} \lambda_n p_n)}{\alpha'}.$$

Problem $\mathcal{P}_{\mathrm{LB}}^{3,n}$ is convex as stated in Lemma 5.

Lemma 5: Optimization problem $\mathcal{P}_{\mathrm{LB}}^{3,n}$ is convex in p_n.

Proof: Recall that $\frac{r_n^{t-1}}{\alpha'} < 0$ and $\frac{s_{jnl}^{t-1}}{\alpha'} < 0$ when $\alpha > 1$. $f(x) = x^{\beta}$ is convex in x when $x > 0$ and $\beta < 0$, and $f(x) = \exp(ax)$ is convex in x, for all $a, x \in \mathbb{R}$. Therefore, the objective $\bar{U}_n(p_n)$ is concave in p_n. Thus, the problem $\mathcal{P}_{\mathrm{LB}}^{3,n}$ is convex with linear constraints. ■

Remark: Though we cannot find a closed-form solution to $\mathcal{P}_{\mathrm{LB}}^{3,n}$, a global optimizer can still be efficiently computed by methods such as bi-section search.

D. MMTS Algorithm Description and Complexity

Based on the above lower bounds, we design MMTS to solve probelm $\mathcal{P}$. The main idea of MMTS is to solve the lower bound problems iteratively until convergence. Specifically, for $0 \le \alpha < 1$, in iteration t, we compute the parameters in the lower bound problem, $\{d_n^{t-1}\}$ and $\{e_n^{t-1}\}$, by (11), and obtain the transmission probability of each tier by (16). For $\alpha = 1$, in iteration t, we compute the parameters in the lower bound problem, $\{g_n^{t-1}\}$, by (18), and obtain the transmission probability of each tier by (23). For $\alpha > 1$, in iteration t, we compute the parameters in the lower bound problem, $\{r_n^{t-1}\}$ and $\{s_{njl}^{t-1}\}$, by (25), and obtain the transmission probability of each tier by solving optimization problem $\mathcal{P}_{\mathrm{LB}}^{3,n}$, for $1 \le n \le N$, via bi-section search. We thus iteratively update the transmission probability of each tier in this way until convergence.

The pseudo code of MMTS is presented in Algorithm 1. In each iteration of MMTS, the lower bound problem either has a straightforward closed-form solution for $0 \le \alpha \le 1$, or is otherwise separable into N *one-dimensional* convex optimization problems, which can be solved efficiently by bi-section search. Thus the computational complexity in each iteration is low. Specifically, when $0 \le \alpha < 1$, the main computation in each iteration is the calculation of parameters $\{d_n^{t-1}\}$ and $\{e_n^{t-1}\}$ related to the sub-problems, since the solution of problems in each iteration is in closed-form. The computation complexity of this is $O(NL)$, where N is the number of tiers, and L is the number of SIR thresholds. When $\alpha = 1$, similar to the $0 \le \alpha < 1$ case, the main computation in each iteration is the calculation of parameters $\{g_n^{t-1}\}$, the computation complexity of which is also $O(NL)$. When $\alpha > 1$, the computation in each iteration consists of two parts, the computation of parameters $\{s_{njl}^{t-1}\}$ and $\{r_n^{t-1}\}$ with computational complexity $O(N^2L)$, and the computation of gradient in bisection search to solve the sub-problems with computational complexity $O(N^2L)$. In summary, the computational complexity in each iteration is $O(NL)$ for $0 \le \alpha \le 1$ and $O(N^2L)$ for $\alpha > 1$.

Furthermore, such decomposition of the original problem into N sub-problems allows distributed computation by each tier, separately from the other tiers. This can be implemented either by some elected representative for each tier, or by each individual nodes in each tier. In one possible distributed implementation scenario, a representative first collects the intensity of each tier. This information can be collected by some central node, such as a base station, and sent to each representative. Alternatively, the representative of each tier collects the local intensity information in its neighborhood, and exchange such information among all representatives, so that they can estimate the intensity of all tiers. Then, in each iteration of MMTS, each representative solves the sub-problem for its own tier. Specifically, each representative runs lines 5-21 in Algorithm 1 and sends the updated transmission probability of its tier to the representatives in other tiers. The information exchange can be realized, for example, via the methods in [25], [26].

VI. Convergence and Optimality of MMTS

In this section, we discuss the convergence and optimality of MMTS. First, as stated in Theorem 6, we observe that MMTS always converges to the objective value of a KKT point of problem $\mathcal{P}$.[1]

Theorem 6: MMTS converges to the objective value of a KKT point of optimization problem $\mathcal{P}$.

Proof: See Appendix E. ■

A KKT point can be a global optimizer, a local optimizer, or even a saddle point. In the following, we present some sufficient conditions under which MMTS converges to the global optimum.

Define $\varrho_{il,\max} = \exp(-m_{il} \sum_{k=1}^{N} P_k^{'} \lambda_k p_{k,\min})$, and $\varrho_{il,\min} = \exp(-m_{il} \sum_{k=1}^{N} P_k^{'} \lambda_k p_{k,\max})$. Define $o_{i,\max} = \sum_{l=1}^{L} a_l \varrho_{il,\max}$, $o_{i,\min} = \sum_{l=1}^{L} a_l \varrho_{il,\min}$, $\xi_{i,\max} = \sum_{l=1}^{L} a_l m_{il} \varrho_{il,\max}$, $\xi_{i,\min} = \sum_{l=1}^{L} a_l m_{il} \varrho_{il,\min}$, $\delta_{i,\max} = \sum_{l=1}^{L} a_l {m_{il}}^2 \varrho_{il,\max}$, and $\psi_{i,\max} = (\lambda_i p_{i,\max} o_{i,\max})^{1-\alpha}$. Further define,

$$\omega_{\min} = \sum_{j=1}^{N} \frac{(\lambda_j p_{j,\min})^{1-\alpha} \xi_{j,\min}}{{o_{j,\max}}^{\alpha}},$$

$$V_{ii} = \max\{ \frac{{o_{i,\max}}^{1-\alpha}}{w_{\min} P_i^{'} (\lambda_i p_{i,\min})^{\alpha}}$$

[1] The KKT conditions are necessary conditions for global or local optimality.

Algorithm 1 MMTS for Solving $\mathcal{P}$

Input: $\{T_l\}$, $\{\mathsf{r}_l\}$ $\{P_n\}$, $\{R_n\}$, $\{\lambda_n\}$, $\{p_{n,\min}\}$, $\{p_{n,\max}\}$, α, γ, N, L.
Output: $\mathbf{p}$.

Compute $a_l = \mathsf{r}_l - \mathsf{r}_{l-1}$, $C_l = \pi T_l^{\frac{2}{\gamma}} \Gamma(1-\frac{2}{\gamma})\Gamma(1+\frac{2}{\gamma})$, $P_n^{'} = {P_n}^{2/\gamma}$, $m_{nl} = \frac{{R_n}^2 C_l}{P_n^{'}}$, for $1 \le n \le N$, $1 \le l \le L$, and $\alpha' = (N+1)(1-\alpha)$.
Pick initial point $\mathbf{p}^0 = [p_n^0]_{N\times 1}$ in $\mathfrak{P}$, and set $t = 0$.
repeat
  $t = t + 1$.
  if $0 \le \alpha < 1$ **then**
    Compute $\{d_n^{t-1}\}$ and $\{e_n^{t-1}\}$ by (11), and set $p_n^t = \left[\frac{e_n^{t-1}}{P_n^{'}\lambda_n \sum_j d_j^{t-1}}\right]_{p_{n,\min}}^{p_{n,\max}}$, for $1 \le n \le N$.
  else if $\alpha = 1$ **then**
    Compute $\{g_n^{t-1}\}$ by (18), and set $p_n^t = \left[\frac{1}{P_n^{'}\lambda_n \sum_j g_j^{t-1}}\right]_{p_{n,\min}}^{p_{n,\max}}$, for $1 \le n \le N$.
  else
    Compute $\{r_n^{t-1}\}$ and $\{s_{njl}^{t-1}\}$ by (25).
    for $n \in \{1, ..., N\}$ **do**
      $p_{\text{lower}} = p_{n,\min}$ and $p_{\text{upper}} = p_{n,\max}$.
      repeat
        $p_{\text{mid}} = (p_{\text{lower}} + p_{\text{upper}})/2$.
        $g = r_n^{t-1} {p_{\text{mid}}}^{\alpha'-1} - \sum_{j=1}^{N}\sum_{l=1}^{L} s_{jnl}^{t-1} m_{jl} P_n^{'} \lambda_n \exp(-\alpha' m_{jl} P_n^{'} \lambda_n p_{\text{mid}})$.
        if $g > 0$ **then**
          $p_{\text{lower}} = p_{\text{mid}}$.
        else
          $p_{\text{upper}} = p_{\text{mid}}$.
        end if
      until convergence
      Set $p_n^t = p_{\text{mid}}$.
    end for
  end if
until convergence
return $\mathbf{p}^t = [p_n^t]_{N\times 1}$

$$+ \psi_{i,\max} \sum_{j'=1}^{N} \frac{(\lambda_{j'} p_{j',\max})^{1-\alpha}}{{w_{\min}}^2 {o_{j',\min}}^{\alpha}} \frac{\delta_{j',\max} - \frac{\alpha {\xi_{j',\min}}^2}{o_{j',\max}}}{1-\alpha},$$

$$\frac{(\lambda_i p_{i,\max})^{1-\alpha} \xi_{i,max}}{w_{\min} {o_{i,\min}}^{\alpha}} + \psi_{i,\max} \frac{(\lambda_i p_{i,\min})^{-\alpha} \xi_{i,\max}}{{w_{\min}}^2 P_i^{'} {o_{i,\min}}^{\alpha}}\},$$

$$V_{ij} = \max\{\psi_{i,\max} \sum_{j'=1}^{N} \frac{(\lambda_{j'} p_{j',\max})^{1-\alpha}}{{w_{\min}}^2 {o_{j',\min}}^{\alpha}} \frac{\delta_{j',\max} - \frac{\alpha {\xi_{j',\min}}^2}{o_{j',\max}}}{1-\alpha},$$

$$\frac{(\lambda_i p_{i,\max})^{1-\alpha} \xi_{i,max}}{w_{\min} {o_{i,\min}}^{\alpha}} + \psi_{i,\max} \frac{(\lambda_j p_{j,\min})^{-\alpha} \xi_{j,\max}}{{w_{\min}}^2 P_j^{'} {o_{j,\min}}^{\alpha}}\}, \ \forall j \neq i.$$

Theorem 7: For $0 \le \alpha < 1$, if

$$(1-\alpha)^2 \sum_{i=1}^{N}\sum_{j=1}^{N} {V_{ij}}^2 < 1, \tag{30}$$

then MMTS converges to the global optimum of problem $\mathcal{P}$.

Proof: See Appendix F. ∎

Based on Theorem 7, we have a corollary as stated in Corollary 8. First, we define $\varrho_{1,\max} = \max_{i,l} \varrho_{il,\max}$, $\varrho_{1,\min} = \min_{i,l} \varrho_{il,\min}$, $\varrho_{2,\max} = \max_{i,l} m_{il}\varrho_{il,\max}$, $\varrho_{2,\min} = \min_{i,l} m_{il}\varrho_{il,\min}$, $\varrho_{3,\min} = \varrho_{2,\min}\sum_{j=1}^{N}\min\{1, \lambda_j p_{j,\min}\}$, and $\varrho_{4,\max} = \max_{i,l} {m_{il}}^2 \varrho_{il,\max}$. We further define

$$\tilde{\beta} = \sum_{i=1}^{N} \frac{\frac{1}{\varrho_{3,\min}}}{P_i^{'} \min\{1, \lambda_i p_{i,\min}\}} + N\frac{\max\{1, \lambda_i p_{i,\max}\}\varrho_{2,\max}}{\varrho_{1,\min}\varrho_{3,\min}},$$

$$\check{\beta} = N\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\max\{1, \lambda_i p_{i,\max}\}\max\{1, \lambda_j p_{j,\max}\rho_{1,\max}\}\varrho_{4,\max}}{\varrho_{1,\min}{\varrho_{3,\min}}^2},$$

$$\dot{\beta} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\varrho_{2,\max}\max\{1, \lambda_j p_{j,\max}\rho_{1,\max}\}}{\varrho_{1,\min}{\varrho_{3,\min}}^2 P_i^{'}\min\{1, \lambda_i p_{i,\min}\}}.$$

Corollary 8: For $1 - \frac{1}{\tilde{\beta}} < \alpha < 1$, if

$$\mathsf{r}_L < \left(\frac{1-(1-\alpha)\tilde{\beta}}{(\check{\beta} + (1-\alpha)\dot{\beta})}\right)^{\frac{1}{2(1-\alpha)}},$$

then MMTS converges to the global optimum.

Proof: See Appendix G. ∎

Remark: Corollary 8 asserts that for $1 - \frac{1}{\tilde{\beta}} < \alpha < 1$, when the maximum transmission rate r_L is sufficiently small, MMTS is optimal. Furthermore, larger minimum transmission probability of each tier, i.e., $p_{n,\min}$, for all n, or smaller maximum transmission probability of each tier, i.e., $p_{n,\max}$, for all n, can increase the range of α where MMTS is optimal.

For $\alpha = 1$, we have the following sufficient condition as stated in Theorem 9 .

Theorem 9: For $\alpha = 1$, if

$$T_L < \left(\frac{1}{N^2} + 1\right)^{\frac{\gamma}{4}} T_1,$$

then MMTS converges to the global optimum.

Proof: See Appendix H. ∎

Remark: Theorem 9 asserts that, for $\alpha = 1$, i.e., when proportional fairness is our optimization objective, if the maximum SIR threshold T_L, and the minimum SIR threshold T_1, are sufficiently close, the number of tiers N is sufficiently small, or the pathloss exponent γ is sufficiently large, then MMTS is globally optimal.

For $\alpha > 1$, we have another sufficient condition as stated in Theorem 10. We define $o_{i,\max}^{'} = \max\{1, \lambda_i p_{i,\max} o_{i,\max}\}$, $o_{i,\min}^{'} = \min\{1, \lambda_i p_{i,\min} o_{i,\min}\}$, $w_{\min}^{'} = \sum_j m_j o_{j,\min}^{'}$, and $\psi_{i,\max}^{'} = \max\ \{1, \lambda_i p_{i,\max} o_{i,\max}\}$. We further define

$$W_{ii} = \max\{\frac{o_{i,\max}^{'}}{w_{\min}^{'} P_i^{'}\lambda_i p_{i,\min}} + \psi_{i,\max}^{'}\sum_{j'=1}^{N}\frac{m_{j'}^2 o_{j',\max}^{'}}{{w_{\min}^{'}}^2},$$

$$m_i\frac{o_{i,\max}^{'}}{w_{\min}^{'}} + \psi_{i,\max}^{'}\frac{m_i o_{i,\max}^{'}}{{w_{\min}^{'}}^2 P_i^{'}\lambda_i p_{i,\min}}\},$$

$$W_{ij} = \max\{\psi_{i,\max}^{'}\sum_{j'=1}^{N}\frac{m_{j'}^2 o_{j',\max}^{'}}{{w_{\min}^{'}}^2},$$

$$m_i\frac{o_{i,\max}^{'}}{w_{\min}^{'}} + \psi_{i,\max}^{'}\frac{m_j o_{j,\max}^{'}}{{w_{\min}^{'}}^2 P_j^{'}\lambda_j p_{j,\min}}\}, \ \forall j \neq i.$$

Theorem 10: For $\alpha > 1 - \frac{1}{M}$, where $M = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}^2}$, if $L = 1$, then MMTS converges to the global optimum.

Proof: See Appendix I. ■

Remark: Recall that the optimal transmission probability in the case $L = 1$, i.e., single SIR threshold, and $\alpha = 1$ has been studied in [8] for the case where the network topology is unknown. However, MMTS is optimal for a far wider range of α values.

VII. Numerical Performance Evaluation

In this section, we study the convergence of MMTS and the impact of different system parameters on its performance. We set the number of tiers $N = 10$. The T-R distance in these tiers are $\{15, 20, \cdots, 60\}$ m. The transmission power of each tier is randomly generated between 1 mW and 5 mW in a uniform manner. We choose five SIR thresholds $\{0.2025, 0.7494, 4.4926, 26.1397, 96.1391\}$ from [27], which corresponds to three transmission rates $\{0.1523, 0.6016, 1.9141, 3.9023, 5.5547\}$ bit/s/Hz in [28]. The default pathloss exponent γ is 4. The maximum transmission probability, $p_{n,\max}$, and minimum transmission probability, $p_{n,\min}$, is set as 1 and 10^{-6}, respectively, for $1 \le n \le 10$. We evaluate the performance of MMTS when the number of T-R pairs is 1000, 3000, 5000, 10000, and 30000 per cell and the cell size is 500m, as recommended by [29]. This corresponds to the sum intensity of all tiers, λ, is 1.3×10^{-3}, 3.6×10^{-3}, 6.5×10^{-3}, 1.3×10^{-2}, and 3.9×10^{-2}. We further added an extra smaller intensity setting of 0.65×10^{-3} for a wider simulation setting. We uniformly generate the intensity of each tier. For the α-fair utility, we consider the average throughput per unit area, with the unit bit/s/Hz/m^2.

We choose the initial point in MMTS in the following log-uniform manner. We uniformly draw the initial point, $\mathbf{y}^0 = [y_n^0]_{N\times 1}$, from set $\mathfrak{P}' = \{\mathbf{y} \in \mathbb{R}^N | \log_{10} p_{n,\min} \le y_n \le \log_{10} p_{n,\max}, 1 \le n \le N\}$. Then the initial point in MMTS is $p_n^0 = 10^{y_n^0}$, for $1 \le n \le N$. In our numerical observation, the KKT points of the optimization problem tend to be small. Hence, the selection of log-uniform initial points in log-uniform manner allows us to use smaller initial points with higher probability, which can speed up the rate of convergence. We run MMTS with multiple initial points, from which we pick the best one that leads to the maximum utility at convergence. Since we do not have prior knowledge about the location of the global optimum, by running MMTS with multiple initial points, we can increase the probability of MMTS converging to the global optimum.

We compare the performance of MMTS with 1 and 5 initial points to the following alternatives: 1) the *optimum by exhaustive search* method where we exhaustively search for a global optimizer over the feasible set, 2) the *equal transmission probability* method where we compute a single transmission probability for all transmitters, which is achieved via MMTS over a single tier in which the T-R distance is set to the average distance among all tiers, i.e., 37.5m, an approximation to the actual multi-tier network topology, and the transmission power is the average of the transmission power of all tiers, 3) the method proposed in [7] that is based on a single tier of T-R pairs and a single SIR threshold. We use it to compute the transmission probability after setting the SIR threshold to the average of all SIR thresholds, the transmission power to the average, and T-R distance to 37.5m. We then assign this probability to all T-R pairs. When we compute the utility of this method, we use *multiple SIR thresholds* and *multiple tiers* for fair comparison. Note that [7] further presents a method based on the Shannon-rate upper bound. It gives similar performance, and is omitted for brevity.

A. Convergence of MMTS

We study the impact of number of tiers and α on the convergence of MMTS. For any given number of tiers, we select evenly spaced T-R distances between 15m and 60m. The convergence condition is that the relative difference of the objective value in consecutive iterations is less than 10^{-3}.

In Fig. 2, we show the convergence behavior of MMTS in one realization. We observe that the objective increases in each iteration as expected in the MM framework. In addition, in this realization, MMTS meets the convergence condition in 11, 4, and 53 iterations, when $\alpha = 0$, $\alpha = 1$, and $\alpha = 2$, respectively. Table II further summarizes the average number of iterations when the convergence condition is met. We observe that for $\alpha = 0.5$ and $\alpha = 1$, the number of tiers has little impact on the number of iterations, while for other α values, the number of iterations increases almost linearly in the number of tiers. Furthermore, the simulation results suggest that the value of α has significant impact on the convergence behavior. MMTS converges faster when α is close to 1.

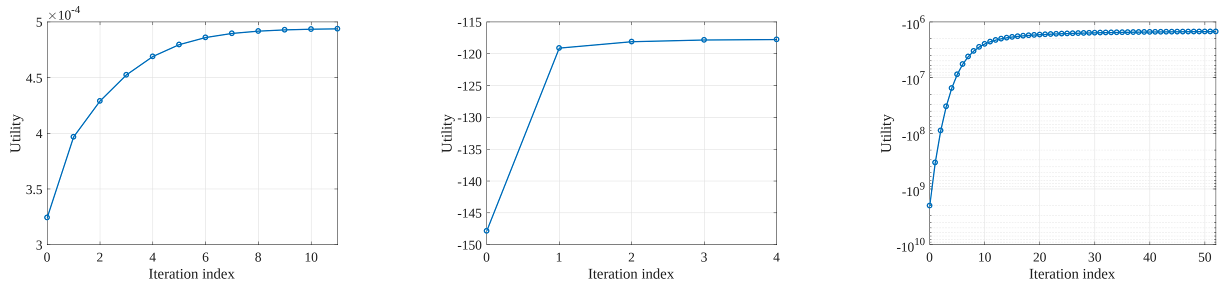
B. Impact of Different System Parameters

We study the impact of intensity, pathloss, and number of SIR thresholds on the utility of the four schemes, when $\alpha = 0$, $\alpha = 1$, and $\alpha = 2$. In Figs. 3-5, we see that under different α values, the utility of MMTS is very close to the optimal utility via exhaustive search. For the case $\alpha = 0$, we observe that running MMTS with multiple initial points can improve the utility, since it can pick a better KKT point compared with running MMTS with 1 initial point. For the cases $\alpha = 1$ and $\alpha = 2$, the utility of MMTS is almost identical to that of the optimal utility via exhaustive search. The reason is that under the simulation settings, the optimization problem is highly likely to have a single KKT point, which is globally optimal by default. Even though the sufficient conditions of Theorems 7, 9, and 10 are not always satisfied, our numerical data indicate that MMTS often converges to the objective value of a good KKT point over a wide range of parameter settings. Furthermore, MMTS substantially outperforms the method in [7] and the *equal transmission probability* method. This suggests the importance of assigning different transmitting probabilities to different types of T-R pairs in maximizing the α-fair utility.

1) Impact of Intensity: In Figs. 3a-3c, we study the utility achieved by MMTS and the other schemes under different node intensity settings. These figures show that, when the intensity is sufficiently high, the utility of MMTS does not

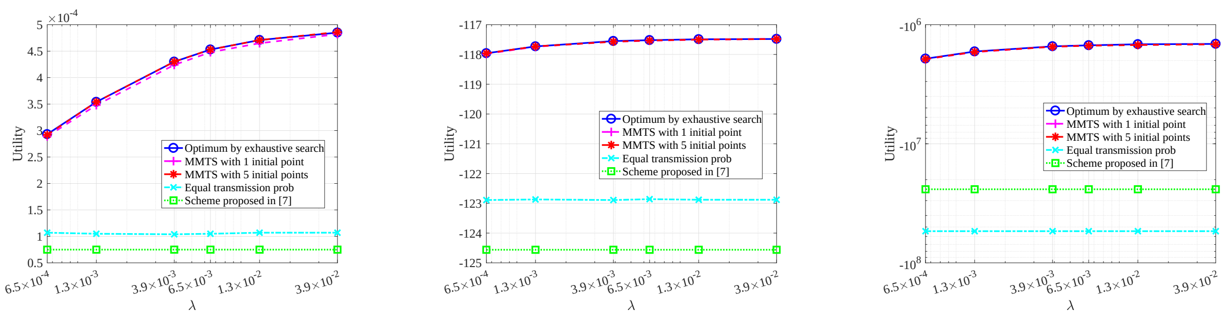
TABLE II: Number of iterations versus number of tiers

# of iterations \ # of tiers (α)	5	10	15	20	25
$\alpha = 0$	10.3	14.2	17.3	19.1	20.8
$\alpha = 0.5$	9.8	8.9	8.7	8.5	8.5
$\alpha = 1$	4.1	4.0	4.0	4.2	4.3
$\alpha = 1.5$	21.3	28.8	36.2	43.4	50.3
$\alpha = 2$	40.7	61.6	82.2	101.0	118.9



(a) $\alpha = 0$ (b) $\alpha = 1$ (c) $\alpha = 2$

Fig. 2: Convergence behavior



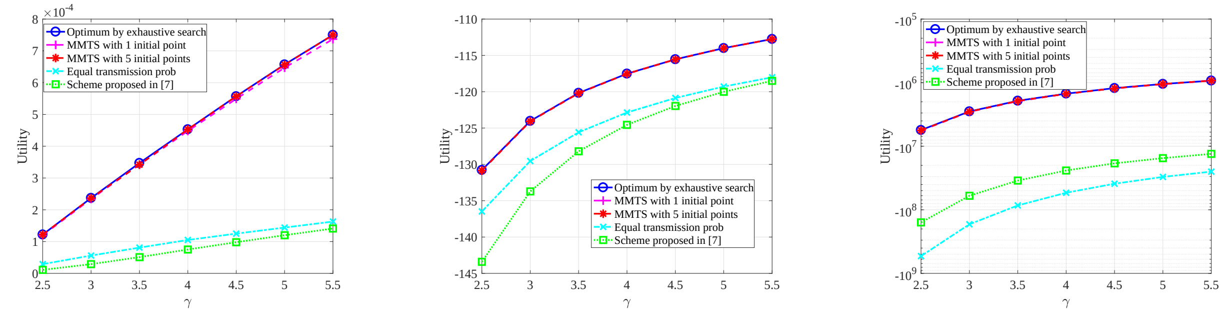
(a) $\alpha = 0$ (b) $\alpha = 1$ (c) $\alpha = 2$

Fig. 3: Utility versus intensity

increase with respect to the intensity, i.e., the system is "saturated". The transmission of a T-R pair introduces interference to other transmitting T-R pairs. When the intensity is sufficiently large, allowing more T-R pairs to transmit does not benefit the utility because of the introduced interference. In addition, we observe that the intensity where the system becomes saturated is larger when $\alpha = 0$ is larger than that when $\alpha = 1$ and $\alpha = 2$. The system is more fair when maximizing the utility with a larger α value. Generally, with larger α, T-R pairs in favorable transmission conditions (e.g., shorter T-R distance or higher transmission power) receive lower throughput, while T-R pairs in unfavorable transmission conditions receive higher throughput. With a larger α value, in order to improve the throughput of the T-R pairs in unfavorable transmission conditions, the interference from other T-R pairs is reduced by decreasing the intensity of the transmitting T-R pairs. Therefore, the intensity where the system becomes saturated becomes smaller with a larger α value.

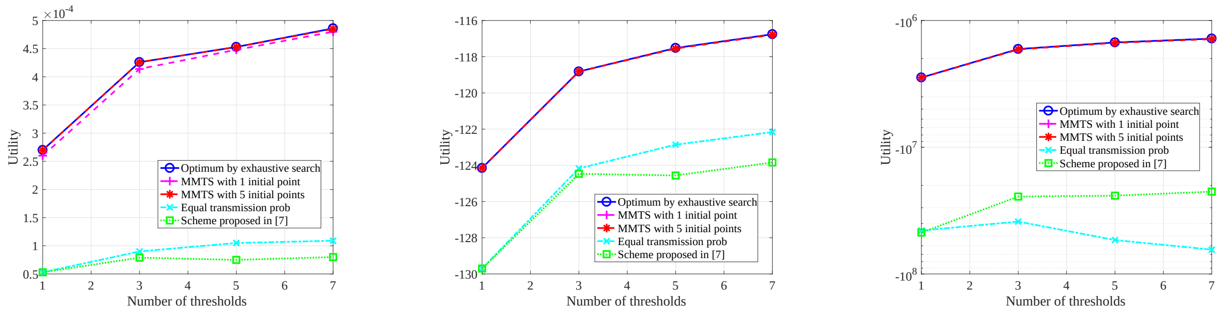
2) Impact of Pathloss Exponent: In Figs. 4a-4c, the utility under different pathloss exponent settings is studied. We observe that the utility of MMTS increases with the increase of pathloss exponent γ. But the behavior of MMTS with respect to γ is different under different α value. The increase of pathloss exponent leads to the weakening of the useful signal, as well as decrease of interference from other T-R pairs. It has different impact on the tiers with different T-R distance and transmission power. From these figures, we see that the utility changes differently with respect to γ under different fairness index α settings. In additon, we observe similar behavior in other alternative schemes.

3) Impact of SIR Thresholds: In Figs. 5a-5c, we study the utility versus the number of SIR thresholds. Evenly spaced SIR thresholds from $\{0.2025, 0.4808, 1.1915, 4.4926, 16.9395, 38.7972, \quad 96.1391\}$ from [27], which corresponds to transmission rates $\{0.1523, 0.377, 0.8770, 1.9141, 3.3223, 4.5234, 5.5547\}$ bit/s/Hz in [28]. We observe that MMTS increases with more thresholds. However, the increase of utility becomes flat when the number of SIR thresholds is sufficiently large. Numerical results such as these can provide design guidelines to system operators on the appropriate number of modulation-coding levels to balance the transmitter complexity and the system performance. Furthermore, in Fig. 5c, the utility of the *equal transmission probability* scheme decreases with more SIR



(a) $\alpha = 0$ (b) $\alpha = 1$ (c) $\alpha = 2$

Fig. 4: Utility versus pathloss exponent



(a) $\alpha = 0$ (b) $\alpha = 1$ (c) $\alpha = 2$

Fig. 5: Utility versus the number of SIR thresholds

thresholds when $\alpha = 2$, and the number of thresholds is larger than 3. In *equal transmission probability* scheme, we approximate the T-R distance of all the T-R pairs with the average when computing the optimal transmission probability. This approximation results in lower likelihood of successful transmission of T-R pairs with longer T-R distance and shorter transmission power when the system has high SIR thresholds. This increases unfairness, and thus decreases the utility. We observe that in Fig. 5a and Fig. 5b, the utility of the method in [7] slightly decreases when the number of SIR thresholds increases from 3 to 5. This is due to the approximation of the multi-tier multi-rate system to the single-tier single-rate system in [7].

VIII. Conclusion

We have studied the optimization of the transmit probabilities in a multi-tier, multi-rate spatial Aloha network with multiple received SIR thresholds with respect to spatial α-fairness. For different ranges of α, the proposed MMTS algorithm utilizes a sequence of iteratively updated lower bound problems, which in turn are decomposed into tier-separable one-dimensional convex problems. The convergence to the objective value of a KKT point is always guaranteed, and several sufficient conditions for global optimality are given. In numerical evaluation, we present the convergence behavior of MMTS and find that MMTS converges faster when α is close to 1. We further study the impact of different system parameters, including the intensity, the pathloss, and the number of SIR thresholds. Our simulation results suggest that MMTS is nearly optimal and has substantial advantage over prior solutions.

Appendix A
Derivation of Laplace Transform of I

The derivation of the Laplace transform of I is as follows:

$$\begin{aligned}
&\mathbb{L}(s) = \mathbb{E}[\exp(-sI)] \\
&= \mathbb{E}_I\left[\exp\left(-s\left(\sum_{k=1,k\neq n}^{N}\sum_{\mathbf{x}\in\Phi_k^{\mathrm{Tx}}}\frac{P_k h_{\mathbf{x0}}}{|\mathbf{x}|^\gamma}\right.\right.\right. \\
&\qquad\left.\left.\left.+\sum_{\mathbf{x}\in\Phi_n^{\mathrm{Tx}}\setminus\{\mathbf{x}_{n1}\}}\frac{P_n h_{\mathbf{x0}}}{|\mathbf{x}|^\gamma}\right)\right)\right] \\
&= \prod_{k=1,k\neq n}^{N}\mathbb{E}_{\Phi_k^{\mathrm{Tx}},\{h_{\mathbf{x0}}\}}\left[\exp\left(-s\sum_{\mathbf{x}\in\Phi_k^{\mathrm{Tx}}}\frac{P_k h_{\mathbf{x0}}}{|\mathbf{x}|^\gamma}\right)\right] \\
&\quad\mathbb{E}_{\Phi_n^{\mathrm{Tx}},\{h_{\mathbf{x0}}\}}\left[\exp\left(-s\sum_{\mathbf{x}\in\Phi_n^{\mathrm{Tx}}\setminus\{\mathbf{x}_{n1}\}}\frac{P_n h_{\mathbf{x0}}}{|\mathbf{x}|^\gamma}\right)\right] \\
&\overset{(a)}{=}\prod_{k=1,k\neq n}^{N}\mathbb{E}_{\Phi_k^{\mathrm{Tx}}}\left[\prod_{\mathbf{x}\in\Phi_k^{\mathrm{Tx}}}\mathbb{E}_{h_{\mathbf{x0}}}\left[\exp\left(-s\frac{P_k h_{\mathbf{x0}}}{|\mathbf{x}|^\gamma}\right)\right]\right] \\
&\quad\mathbb{E}_{\Phi_n^{\mathrm{Tx}}}\left[\prod_{\mathbf{x}\in\Phi_n^{\mathrm{Tx}}\setminus\{\mathbf{x}_{n1}\}}\mathbb{E}_{h_{\mathbf{x0}}}\left[\exp\left(-s\frac{P_n h_{\mathbf{x0}}}{|\mathbf{x}|^\gamma}\right)\right]\right]
\end{aligned}$$

$$\overset{(b)}{=} \prod_{k=1,k\neq n}^{N} \mathbb{E}_{\Phi_k^{\text{Tx}}} \left[\prod_{\mathbf{x}\in\Phi_k^{\text{Tx}}} \frac{1}{1+s\frac{P_k}{|\mathbf{x}|^{\gamma}}} \right]$$

$$\mathbb{E}_{\Phi_n^{\text{Tx}}} \left[\prod_{\mathbf{x}\in\Phi_n^{\text{Tx}}\setminus\{\mathbf{x}_{n1}\}} \frac{1}{1+s\frac{P_n}{|\mathbf{x}|^{\gamma}}} \right]$$

$$\overset{(c)}{=} \prod_{j=1}^{N} \exp\left(p_j\lambda_j \int_{\mathbb{R}^2} \left(1-\frac{1}{1+s\frac{P_j}{|\mathbf{x}|^{\gamma}}}\right) \mathrm{d}\mathbf{x} \right)$$

$$\overset{(d)}{=} \prod_{j=1}^{N} \exp\left(-p_j\lambda_j 2\pi \int_0^{\infty} \left(1-\frac{1}{1+s\frac{P_j}{d^{\gamma}}}\right) d\mathrm{d}d \right)$$

$$\overset{(e)}{=} \prod_{j=1}^{N} \exp\left(-p_j\lambda_j \pi (sP_j)^{\frac{2}{\gamma}} \Gamma(1-\frac{2}{\gamma})\Gamma(1+\frac{2}{\gamma}) \right),$$

where (a) is based on the independence between $\{h_{\mathbf{x0}}\}$ and $\{\Phi_k^{\text{Tx}}\}$, (b) is from the distribution of $\{h_{\mathbf{x0}}\}$, (c) is from the Slivnyak's theorem and probability generating functional (PGFL) of homogeneous PPP, (d) is based on the transformation to polar coordinate, and (e) is from the manipulation of Gamma function.

Appendix B
Proof of Lemma 1

First, we show that function $f(\mathbf{x}) : \mathbb{R}^N \to \mathbb{R}$, given by

$$f(\mathbf{x}) = \frac{\left(\sum_{n=1}^{N} \rho_n \exp(x_n)\right)^{\beta}}{\beta},$$

is convex in $\mathbf{x} = [x_n]_{N\times 1} \in \mathbb{R}^N$, for $\rho_n \geq 0, 1 \leq n \leq N$, and $\beta > 0$. We rewrite $f(\mathbf{x})$ as $f(\mathbf{x}) = \frac{1}{\beta}\exp\left(\beta \log\left(\sum_{i=1}^{N} \rho_n \exp(x_n)\right)\right)$. Then $f(\mathbf{x})$ can be viewed as the composition of $h(y) : \mathbb{R} \to \mathbb{R}$ and $g(\mathbf{x}) : \mathbb{R}^N \to \mathbb{R}$, i.e., $f(\mathbf{x}) = h(g(\mathbf{x}))$, where $h(y) = \frac{1}{\beta}\exp(\beta y)$, and $g(\mathbf{x}) = \log\left(\sum_{n=1}^{N} \rho_n \exp(x_n)\right)$. Since $h(y)$ is convex and nondecreasing in y, and $g(\mathbf{x})$ is convex in $\mathbf{x}$, according to the composition rule [24], function $f(\mathbf{x}) = h(g(\mathbf{x}))$ is convex in $\mathbf{x}$.

Due to the convexity of $f(\mathbf{x})$, for all $\mathbf{x}$ and $\mathbf{x}^{t-1}$ in the domain of f, we have $f(\mathbf{x}) \geq f(\mathbf{x}^{t-1}) + \nabla f(\mathbf{x}^{t-1})(\mathbf{x}-\mathbf{x}^{t-1})$.

Hence

$$\frac{\left(\sum_{n=1}^{N} \rho_n \exp(x_n)\right)^{\beta}}{\beta} \geq \frac{\left(\sum_{n=1}^{N} \rho_n \exp(x_n^{t-1})\right)^{\beta}}{\beta} + \left(\sum_{j=1}^{N} \rho_j \exp(x_j^{t-1})\right)^{\beta-1} \sum_{n=1}^{N} \rho_n \exp(x_n^{t-1})(x_n - x_n^{t-1}), \tag{31}$$

and the equality holds when $\mathbf{x} = \mathbf{x}^{t-1}$.

Since $0 \leq \alpha < 1$, we have

$$\frac{\left(\lambda_n p_n \sum_{l=1}^{L} a_l \exp(-m_{nl}\sum_{j=1}^{N} P_j^{'}\lambda_j p_j)\right)^{1-\alpha}}{1-\alpha}$$

$$\overset{(a)}{=} {\lambda_n}^{1-\alpha} \frac{\left(\sum_{l=1}^{L} a_l \exp(\log p_n - m_{nl}\sum_{j=1}^{N} P_j^{'}\lambda_j p_j)\right)^{1-\alpha}}{(1-\alpha)}$$

$$\overset{(b)}{\geq} -d_n^{t-1}\sum_{j=1}^{N} P_j^{'}\lambda_j(p_j - p_j^{t-1}) + e_n^{t-1}\log\frac{p_n}{p_n^{t-1}} + f_n^{t-1},$$

where (a) results from $p_n = \log(\exp(p_n))$, (b) is from (31), and the equality holds if $\mathbf{p} = \mathbf{p}^{t-1}$.

Appendix C
Proof of Lemma 2

Recall that $e_n^{t-1} \geq 0$,for all $1 \leq n \leq N$. Hence, $e_n^{t-1}\log p_n$ and $d_j^{t-1}P_n^{'}\lambda_n p_n$ are both concave in p_n. Thus, $\widetilde{U}_n(p_n)$ is concave in p_n, which makes the optimization problem $\mathcal{P}_{\text{LB}}^{1,n}$ convex. The Lagrangian is

$$\mathcal{L}(p_n,\mu_1,\mu_2) = -\widetilde{U}_n(p_n) + \mu_1(p_n - p_{n,\max}) + \mu_2(-p_n + p_{n,\min}). \tag{32}$$

KKT conditions are

$$\frac{\partial\mathcal{L}}{\partial p_n}|_{p_n=\tilde{p}_n^*} = -\frac{e_n^{t-1}}{\tilde{p}_n^*} + P_n^{'}\lambda_n\sum_{j=1}^{N} d_j^{t-1} + \mu_1 - \mu_2 = 0, \tag{33}$$

$$\mu_1(\tilde{p}_n^* - p_{n,\min}) = 0, \mu_2(p_{n,\max} - \tilde{p}_n^*) = 0, \tag{34}$$

$$p_{n,\min} \leq \tilde{p}_n^* \leq p_{n,\max}, \mu_1 \geq 0, \mu_2 \geq 0. \tag{35}$$

If $\tilde{p}_n^* \neq p_{n,\min}$ and $\tilde{p}_n^* \neq p_{n,\max}$, then $\tilde{p}_n^* = \frac{e_n^{t-1}}{P_n^{'}\lambda_n\sum_j d_j^{t-1}}$ based on (33). If $\tilde{p}_n^* = p_{n,\min}$, then $\mu_2 = 0$. Substitute $\mu_1 \geq 0, \mu_2 = 0$, and $\tilde{p}_n^* = p_{n,\min}$ to (33), we have $p_{n,\min} \leq \frac{e_n^{t-1}}{P_n^{'}\lambda_n\sum_j d_j^{t-1}}$. Similarly, if $\tilde{p}_n^* = p_{n,\max}$, we have $p_{n,\max} \geq \frac{e_n^{t-1}}{P_n^{'}\lambda_n\sum_j d_j^{t-1}}$. In conclusion, $\tilde{p}_n^* = \left[\frac{e_n^{t-1}}{P_n^{'}\lambda_n\sum_j d_j^{t-1}}\right]_{p_{n,\min}}^{p_{n,\max}}$.

Appendix D
Proof of Lemma 4

Define $x_n = P_n^{'}\lambda_n p_n$, $x_n^{t-1} = P_n^{'}\lambda_n p_n^{t-1}$, $\nu_n^{t-1} = \sum_{l'=1}^{L} a_{l'}\exp(-m_{nl'}\sum_{j=1}^{N} x_j^{t-1})$, and $\theta_{nl}^{t-1} = \exp(-m_{nl}\sum_{j=1}^{N} x_j^{t-1})$, for all n,l. Then we have

$$a_l x_n \exp(-m_{nl}\sum_{j=1}^{N} x_j) = \frac{a_l\theta_{nl}^{t-1}}{\nu_n^{t-1}}\frac{\nu_n^{t-1}}{\theta_{nl}^{t-1}} x_n \exp(-m_{nl}\sum_{j=1}^{N} x_j).$$

Note that $\sum_{l=1}^{L}\frac{a_l\theta_{nl}^{t-1}}{\nu_n^{t-1}} = 1$.

Hence, we have

$$\left(\sum_{l=1}^{L} a_l x_n \exp(-m_{nl}\sum_{j=1}^{N} x_j)\right)^{1-\alpha}$$

$$= \left(\sum_{l=1}^{L}\frac{a_l\theta_{il}^{t-1}}{\nu_i^{t-1}}\frac{\nu_i^{t-1}}{\theta_{il}^{t-1}} x_i \exp(-m_{il}\sum_{j=1}^{N} x_j)\right)^{1-\alpha}$$

$$\overset{(a)}{\leq} \sum_{l=1}^{L}\frac{a_l\theta_{nl}^{t-1}}{\nu_n^{t-1}}\left(\frac{\nu_n^{t-1}}{\theta_{nl}^{t-1}} x_n \exp(-m_{nl}\sum_{j=1}^{N} x_j)\right)^{1-\alpha}$$

$$= \sum_{l=1}^{L} a_l\left(\frac{\theta_{nl}^{t-1}}{\nu_n^{t-1}}\right)^{\alpha} x_n^{1-\alpha}\prod_{j=1}^{N}\exp(-m_{nl}(1-\alpha)x_j),$$

where (a) is from Jensen's inequality and the fact that function $f(x) = x^\beta$ is convex in x when $\beta < 0$ and $x > 0$. Note that equality holds when $\mathbf{x} = \mathbf{x}^{t-1}$, where $\mathbf{x} = [x_n]_{N\times 1}$, and $\mathbf{x}^{t-1} = [x_n^{t-1}]_{N\times 1}$.

Next, note that for all $z_n > 0, z_n^{t-1} > 0$, we have $\prod_{n=1}^N z_n \le \sum_{n=1}^N \frac{\prod_{j=1,j\neq n} z_j^{t-1}}{N z_n^{t-1\,N-1}} z_n{}^N$, and the equality holds if $z_n = z_n^{t-1}$, for all n. This can be easily proved by the inequality of arithmetic and geometric means. Hence, we have

$$x_n^{1-\alpha} \prod_{j=1}^N \exp(-m_{nl}(1-\alpha)x_j) \le \frac{\zeta_{nl}^{t-1} {x_n}^{\alpha'}}{N+1} + \sum_{j=1}^N \frac{\phi_{njl}^{t-1} \exp(-\alpha' m_{il} x_j)}{N+1}. \tag{36}$$

where

$$\zeta_{nl}^{t-1} = \frac{\prod_{k=1}^N \exp(-(1-\alpha)m_{nl}x_k^{t-1})}{x_n^{t-1\,(1-\alpha)N}},$$

$$\phi_{njl}^{t-1} = \frac{{x_n^{t-1}}^{1-\alpha} \prod_{k=1,k\neq j}^N \exp(-(1-\alpha)m_{nl}x_k^{t-1})}{\exp(-(1-\alpha)N m_{nl} x_j^{t-1})}.$$

Recall that $1-\alpha < 0$. From (36), we have

$$\frac{\left(\sum_{l=1}^L a_l x_n \exp(-m_{nl}\sum_{j=1}^N x_j)\right)^{1-\alpha}}{1-\alpha} \ge \sum_{l=1}^L a_l \left(\frac{\theta_{nl}^{t-1}}{\nu_n^{t-1}}\right)^{\alpha} \left(\frac{\zeta_{nl}^{t-1}{x_n}^{\alpha'}}{\alpha'} + \sum_{j=1}^N \frac{\phi_{njl}^{t-1}\exp(-\alpha' m_{il} x_j)}{\alpha'}\right), \tag{37}$$

where the equality holds if $\mathbf{x} = \mathbf{x}^{t-1}$.

Substituting $x_n = P_n^{'}\lambda_i p_i$, and $x_n^{t-1} = P_n^{'}\lambda_n p_n^{t-1}$ into (37) completes the proof.

Appendix E
Proof of Theorem 6

The boundedness of the feasible set in the optimization problem guarantees that the points generated by MMTS have at least one limit point [30]. We need to verify the following two conditions of our proposed lower bounds to show that MMTS converges to the objective value of a KKT point [31].

Condition 1: $\forall \mathbf{p}, \mathbf{p}^{t-1} \in \mathfrak{P}$; for $0 \le \alpha < 1$, $U(\mathbf{p}) \ge \tilde{U}(\mathbf{p}, \mathbf{p}^{t-1})$; for $\alpha = 1$, $U(\mathbf{p}) \ge \hat{U}(\mathbf{p}, \mathbf{p}^{t-1})$; and for $\alpha > 1$, $U(\mathbf{p}) \ge \bar{U}(\mathbf{p}, \mathbf{p}^{t-1})$. Furthermore, in all cases above, the equalities hold when $\mathbf{p} = \mathbf{p}^{t-1}$.

This condition has been verified in section V.

Condition 2: $\forall \mathbf{p}, \mathbf{p}^{t-1} \in \mathfrak{P}$, for $0 \le \alpha < 1$ and $1 \le k \le N$, $\frac{\partial U(\mathbf{p})}{\partial p_k}|_{\mathbf{p}=\mathbf{p}^{t-1}} = \frac{\partial \tilde{U}(\mathbf{p},\mathbf{p}^{t-1})}{\partial p_k}|_{\mathbf{p}=\mathbf{p}^{t-1}}$; for $\alpha = 1$ and $1 \le k \le N$, $\frac{\partial U(\mathbf{p})}{\partial p_k}|_{\mathbf{p}=\mathbf{p}^{t-1}} = \frac{\partial \hat{U}(\mathbf{p},\mathbf{p}^{t-1})}{\partial p_k}|_{\mathbf{p}=\mathbf{p}^{t-1}}$; and for $\alpha > 1$ and $1 \le k \le N$, $\frac{\partial U(\mathbf{p})}{\partial p_k}|_{\mathbf{p}=\mathbf{p}^{t-1}} = \frac{\partial \bar{U}(\mathbf{p},\mathbf{p}^{t-1})}{\partial p_k}|_{\mathbf{p}=\mathbf{p}^{t-1}}$.

To show that this condition is satisfied, we first note that

$$\frac{\partial U(\mathbf{p})}{\partial p_k} = {\lambda_k}^{1-\alpha} {p_k}^{-\alpha} \exp\left(-m_{il}\sum_{j=1}^N p_j \lambda_j P_j^{'}\right) - \sum_{i=1}^N \frac{(\lambda_i p_i)^{1-\alpha} P_k^{'}\lambda_k \sum_{l=1}^L a_l m_{il} \exp\left(-m_{il}\sum_{j=1}^N p_j\lambda_j P_j^{'}\right)}{\left(\sum_{l=1}^L a_l \exp\left(-m_{il}\sum_{j=1}^N p_j \lambda_j P_j^{'}\right)\right)^{\alpha}}.$$

For $0 \le \alpha < 1$,

$$\frac{\partial \tilde{U}(\mathbf{p},\mathbf{p}^{t-1})}{\partial p_k}|_{\mathbf{p}=\mathbf{p}^{t-1}} = \frac{e_k^{t-1}}{p_k^{t-1}} - \sum_{i=1}^N d_i^{t-1} P_k^{'}\lambda_k. \tag{38}$$

Plugging (11) into (38), we can verify that $\frac{\partial U(\mathbf{p})}{\partial p_k}|_{\mathbf{p}=\mathbf{p}^{t-1}} = \frac{\partial \tilde{U}(\mathbf{p},\mathbf{p}^{t-1})}{\partial p_k}|_{\mathbf{p}=\mathbf{p}^{t-1}}$, for $1 \le k \le N$.

Similarly, we can verify that $\frac{\partial U(\mathbf{p})}{\partial p_k}|_{\mathbf{p}=\mathbf{p}^{t-1}} = \frac{\partial \hat{U}(\mathbf{p},\mathbf{p}^{t-1})}{\partial p_k}|_{\mathbf{p}=\mathbf{p}^{t-1}}$, for $\alpha = 1$ and $1 \le k \le N$, and $\frac{\partial U(\mathbf{p})}{\partial p_k}|_{\mathbf{p}=\mathbf{p}^{t-1}} = \frac{\partial \bar{U}(\mathbf{p},\mathbf{p}^{t-1})}{\partial p_k}|_{\mathbf{p}=\mathbf{p}^{t-1}}$, for $\alpha > 1$ and $1 \le k \le N$.

Appendix F
Proof of Theorem 7

First, we present a useful lemma.

Lemma 11: [32] Consider an iterative algorithm,

$$\mathbf{x}(t+1) = T(\mathbf{x}(t)), t = 0, 1, \cdots,$$

where mapping $T : \mathcal{X} \to \mathcal{X}$, and $\mathcal{X}$ is a closed subset of $\mathbb{R}^N$. If T satisfies

$$|T(\mathbf{x}) - T(\mathbf{y})| \le \sigma|\mathbf{x} - \mathbf{y}|, \forall \mathbf{x}, \mathbf{y} \in \mathcal{X}$$

where $|\cdot|$ is some norm and σ is a constant in $[0, 1)$, then the mapping T has a unique fixed point.

In this paper, we consider l_2-norm. For notational convenience, we let $x_i = P_i^{'}\lambda_i p_i$, $x_i^t = P_i^{'}\lambda_i p_i^t$, $x_{i,\max} = P_i^{'}\lambda_i p_{i,\max}$, $x_{i,\min} = P_i^{'}\lambda_i p_{i,\min}$, $1 \le i \le N$, and $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^N | x_{i,\min} \le x_i \le x_{i,\max}, 1 \le i \le N\}$.

We further define a vector-valued function $\mathbf{f}(\mathbf{x})$, where

$$f_i(\mathbf{x}) = \frac{\left(\frac{\sum_{l=1}^L a_l x_i \exp(-m_{il}\sum_{k=1}^N x_k)}{P_i^{'}}\right)^{1-\alpha}}{\sum_j \frac{{x_j}^{1-\alpha}\left(\sum_{l=1}^L a_l m_{jl}\exp(-m_{jl}\sum_{k=1}^N x_k)\right)}{{P_j^{'}}^{1-\alpha}\left(\sum_{l=1}^L a_l \exp(-m_{jl}\sum_{k=1}^N x_k)\right)^{\alpha}}}, \ \forall i. \tag{39}$$

The Jacobian matrix of $\mathbf{f}(\mathbf{x})$ is defined as $\mathbf{J}(\mathbf{x}) = [\frac{\partial f_i}{\partial x_j}]_{N\times N}$.

Furthermore, we use the following mapping

$${x_i}^{t+1} = \left[f_i(\mathbf{x}^t)\right]_{x_{i,\min}}^{x_{i,\max}}.$$

Let $o_i = \sum_{l=1}^L a_l \exp(-m_{il}\sum_{k=1}^N x_k)$, $\xi_i = \sum_{l=1}^L a_l m_{il} \exp(-m_{il}\sum_{k=1}^N x_k)$, and $\delta_i = \sum_{l=1}^L a_l {m_{il}}^2 \exp(-m_{il}\sum_{k=1}^N x_k)$.

The numerator and denominator of $f_i(\mathbf{x})$ is given by $\psi_i(\mathbf{x}) = \frac{{x_i}^{1-\alpha}{o_i}^{1-\alpha}}{{P_i^{'}}^{1-\alpha}}$, and $\omega(\mathbf{x}) = \sum_{j'} \frac{{x_{j'}}^{1-\alpha}\xi_{j'}}{{P_{j'}^{'}}^{1-\alpha}{o_{j'}}^{\alpha}}$ respectively. After simple derivation, the partial derivatives of $\psi_i(\mathbf{x})$ and $\omega(\mathbf{x})$ are give by

$$\frac{\partial \psi_i}{\partial x_i} = \frac{(1-\alpha)(o_i - x_i\xi_i)}{{P_i^{'}}^{1-\alpha}{x_i}^{\alpha}{o_i}^{\alpha}}, \frac{\partial \psi_i}{\partial x_j} = \frac{(\alpha-1){x_i}^{1-\alpha}\xi_i}{{P_i^{'}}^{1-\alpha}{o_i}^{\alpha}}, \ \ \forall j \neq i,$$

and

$$\frac{\partial \omega}{\partial x_j} = \sum_{j'} \frac{{x_{j'}}^{1-\alpha}}{{P_{j'}^{'}}^{1-\alpha}{o_{j'}}^{\alpha}}\left(\frac{\alpha{\xi_{j'}}^2}{o_{j'}} - \delta_{j'}\right) + \frac{(1-\alpha){x_j}^{-\alpha}\xi_j}{{P_j^{'}}^{1-\alpha}{o_j}^{\alpha}}, \ \forall j.$$

Then we have

$$\frac{\partial f_i}{\partial x_i} = \frac{\frac{\partial \psi_i}{\partial x_i}\omega - \frac{\partial \omega}{\partial x_i}\psi_i}{\omega^2}$$

$$= \frac{(1-\alpha)o_i}{wP_i'^{1-\alpha}x_i^{\alpha}o_i^{\alpha}} + \psi_i \sum_{j'} \frac{x_{j'}^{1-\alpha}}{w^2 P_{j'}'^{1-\alpha} o_{j'}^{\alpha}}(\delta_{j'} - \frac{\alpha \xi_{j'}^2}{o_{j'}}) - \left(\frac{(1-\alpha)x_i\xi_i}{wP_i'^{1-\alpha}x_i^{\alpha}o_i^{\alpha}} + \psi_i \frac{(1-\alpha)x_i^{-\alpha}\xi_i}{w^2 P_i'^{1-\alpha} o_i^{\alpha}} \right),$$

and

$$\frac{\partial f_i}{\partial x_j} = \frac{\frac{\partial \psi_i}{\partial x_j}\omega - \frac{\partial \omega}{\partial x_j}\psi_i}{\omega^2} = \psi_i \sum_{j'} \frac{x_{j'}^{1-\alpha}}{w^2 P_{j'}'^{1-\alpha} o_{j'}^{\alpha}}(\delta_{j'} - \frac{\alpha \xi_{j'}^2}{o_{j'}}) - \left(\frac{(1-\alpha)x_i\xi_i}{wP_i'^{1-\alpha}x_i^{\alpha}o_i^{\alpha}} + \psi_i \frac{(1-\alpha)x_j^{-\alpha}\xi_j}{w^2 P_j'^{1-\alpha} o_j^{\alpha}} \right), \forall j \neq i.$$

Based on the Cauchy-Schwarz inequality, it is easy to show

$$\delta_j o_j \geq \xi_j^2, \forall j.$$

Then since $\alpha \leq 1$, we have $\delta_j \geq \frac{\alpha \xi_j^2}{o_j}$. Then it is obvious that

$$\left|\frac{\partial f_i}{\partial x_i}\right| \leq (1-\alpha)V_{ii}, \left|\frac{\partial f_i}{\partial x_j}\right| \leq (1-\alpha)V_{ij}$$

Based on the assumption in Theorem 7, $|\mathbf{J}(\mathbf{x})| \leq \sqrt{(1-\alpha)^2 \sum_{i=1}^N \sum_{j=1}^N V_{ij}^2} < 1$. Hence, based on Mean Value Theorem, for all $\mathbf{x}^1, \mathbf{x}^2 \in \mathcal{X}$, we have

$$|\mathbf{f}(\mathbf{x}^1) - \mathbf{f}(\mathbf{x}^2)| \leq |\mathbf{J}(\widetilde{\mathbf{x}})||\mathbf{x}^1 - \mathbf{x}^2| \leq \sqrt{(1-\alpha)^2 \sum_{i=1}^N \sum_{j=1}^N V_{ij}^2} |\mathbf{x}^1 - \mathbf{x}^2|, \tag{40}$$

where $\widetilde{\mathbf{x}}$ is a convex combination of $\mathbf{x}^1$ and $\mathbf{x}^2$. Based on Lemma 11, the proposed algorithm converges to the objective value of the unique fixed point, which must be the global optimum.

Appendix G
Proof of Corollary 8

Recall that $a_l = \mathsf{r}_l - \mathsf{r}_{l-1}$ and $\mathsf{r}_0 = 0$. Therefore, $\sum_{l=1}^L a_l = \mathsf{r}_L$. For notional convenience, we define $a = \mathsf{r}_L$. It is easy to prove that $o_{i,max} \leq a\varrho_{1,\max}$, $o_{i,\min} \geq a\varrho_{1,\min}$, $\xi_{i,\max} \leq a\varrho_{2,\max}$, $\xi_{i,\min} \geq a\varrho_{2,\min}$, and $\delta_{i,\max} \leq a\varrho_{4,\max}$ for all $1 \leq i \leq N$.

Furthermore,

$$\omega_{\min} = \sum_j \frac{(\lambda_j p_{j,\min})^{1-\alpha}\xi_{j,\min}}{o_{j,\max}^{\alpha}} \geq \sum_j \frac{\min\{1, \lambda_j p_{j,\min}\} a\varrho_{2,\min}}{(a\varrho_{1,\max})^{\alpha}} \overset{(a)}{\geq} \sum_j \frac{\min\{1, \lambda_j p_{j,\min}\}\varrho_{2,\min}}{\max\{1, \varrho_{1,\max}\}} a^{1-\alpha} \geq \varrho_{3,\min} a^{1-\alpha},$$

where (a) is is from the fact that $\varrho_{il,\max} < 1$ for all $1 \leq i \leq N$ and $1 \leq l \leq L$.

In addition, we have

$$\psi_{i,\max} = (\lambda_i p_{i,\max} o_{i,\max})^{1-\alpha} \leq (\lambda_i p_{i,\max} \rho_{1,\max})^{1-\alpha} a^{1-\alpha} \leq \max\{1, \lambda_i p_{i,\max} \rho_{1,\max}\} a^{1-\alpha}.$$

Therefore,

$$\frac{o_{i,\max}^{1-\alpha}}{w_{\min} P_i'(\lambda_i p_{i,\min})^{\alpha}} \leq \frac{a^{1-\alpha}(\varrho_{1,\max})^{1-\alpha}}{\varrho_{3,\min} a^{1-\alpha} P_i'(\lambda_i p_{i,\min})^{\alpha}} \leq \frac{1}{\varrho_{3,\min} P_i' \min\{1, \lambda_i p_{i,\min}\}}.$$

Similarly, we have

$$\frac{(\lambda_i p_{i,\max})^{1-\alpha}}{w_{\min}^2 o_{i,\min}^{\alpha}} \delta_{i,\max} \leq \frac{\max\{1, \lambda_i p_{i,\max}\}\varrho_{4,\max}}{\varrho_{1,\min}\varrho_{3,\min}^2} a^{1-\alpha},$$
$$\frac{(\lambda_i p_{i,\max})^{1-\alpha}\xi_{i,max}}{w_{\min} o_{i,\min}^{\alpha}} \leq \frac{\max\{1, \lambda_i p_{i,\max}\}\varrho_{2,\max}}{\varrho_{1,\min}\varrho_{3,\min}},$$
$$\frac{(\lambda_i p_{i,\min})^{-\alpha}\xi_{i,\max}}{w_{\min}^2 P_i' o_{i,\min}^{\alpha}} \leq \frac{\varrho_{2,\max} a^{1-\alpha}}{\varrho_{1,\min}\varrho_{3,\min}^2 P_i' \min\{1, \lambda_i p_{i,\min}\}}.$$

After simple derivation, we have

$$(1-\alpha)\sum_{i=1}^N \sum_{j=1}^N V_{ij} \leq (1-\alpha)\tilde{\beta} + (\check{\beta} + (1-\alpha)\dot{\beta})a^{2(1-\alpha)}.$$

Recall that $a = \mathsf{r}_L$ and $1 - \frac{1}{\tilde{\beta}} < \alpha < 1$. If $\mathsf{r}_L < \left(\frac{1-(1-\alpha)\tilde{\beta}}{(\check{\beta}+(1-\alpha)\dot{\beta})}\right)^{\frac{1}{2(1-\alpha)}}$, we have

$$(1-\alpha)\tilde{\beta} + (\check{\beta} + (1-\alpha)\dot{\beta})a^{2(1-\alpha)} < 1,$$

and

$$(1-\alpha)^2 \sum_{i=1}^N \sum_{j=1}^N V_{ij}^2 \leq \left((1-\alpha)\sum_{i=1}^N \sum_{j=1}^N V_{ij}\right)^2 < 1.$$

Based on Theorem 7, MMTS converges to the global optimum.

Appendix H
Proof of Theorem 9

For $\alpha = 1$, we have

$$\frac{\partial f_i}{\partial x_j} = \frac{1}{w^2}\sum_{j'}(\frac{\delta_{j'}}{o_{j'}} - \frac{\xi_{j'}^2}{o_{j'}^2}). \tag{41}$$

Furthermore, $\frac{\delta_{j'}}{o_{j'}} \leq m_{j'L}^2$ and $\frac{\xi_{j'}^2}{o_{j'}^2} \geq m_{j'1}^2$ for all j'. In addition, $\omega = \sum_{j'} \frac{\xi_{j'}}{o_{j'}} \geq \sum_{j'} m_{j'1}$. Therefore

$$\frac{\partial f_i}{\partial x_j} = \frac{1}{w^2}\sum_{j'}(\frac{\delta_{j'}}{o_{j'}} - \frac{\xi_{j'}^2}{o_{j'}^2}) \leq \frac{1}{\left(\sum_{j'} m_{j'1}\right)^2}\sum_{j'}\left(m_{j'L}^2 - m_{j'1}^2\right).$$

Similar to the proof of Theorem 7, the sufficient condition for global optimality is

$$\frac{1}{\left(\sum_j m_{j1}\right)^2}\sum_j\left(m_{jL}^2 - m_{j1}^2\right) < \frac{1}{N^2}. \tag{42}$$

We relax the sufficient condition (42) as

$$\sum_j \left({m_{jL}}^2 - {m_{j1}}^2\right) < \frac{1}{N^2}\sum_j {m_{j1}}^2, \tag{43}$$

based on the fact that $\sum_j {m_{j1}}^2 \le \left(\sum_j m_{j1}\right)^2$.

Plugging $m_{il} = \frac{{R_i}^2}{P_i'}\pi {T_l}^{\frac{2}{\gamma}}\Gamma(1-\frac{2}{\gamma})\Gamma(1+\frac{2}{\gamma})$ into (43), we have

$$T_L < \left(\frac{1}{N^2}+1\right)^{\frac{\gamma}{4}} T_1. \tag{44}$$

Appendix I
Proof of Theorem 10

For notational convenience, we let $m_i = m_{i1}$ for all i.

Case 1: $\alpha \in (1-\frac{1}{M}, 1)$.

For $L=1$, we have $\delta_i = {m_i}^2 o_i$ and $\xi_i = m_i o_i$. Similar to the proof of Theorem 7, we have

$$\begin{aligned}\left|\frac{\partial f_i}{\partial x_i}\right| =&(1-\alpha)\left|\frac{{o_i}^{1-\alpha}}{w{P_i'}^{1-\alpha}{x_i}^{\alpha}} + \psi_i \sum_{j'} \frac{m_{j'}^2 {x_{j'}}^{1-\alpha}{o_{j'}}^{1-\alpha}}{w^2 {P_{j'}'}^{1-\alpha}}\right.\\ &\left. - \left(\frac{x_i\xi_i}{w{P_i'}^{1-\alpha}{x_i}^{\alpha}{o_i}^{\alpha}} + \psi_i \frac{{x_i}^{-\alpha}\xi_i}{w^2{P_i'}^{1-\alpha}{o_i}^{\alpha}}\right)\right|\\ \le&(1-\alpha)W_{ii},\end{aligned} \tag{45}$$

and

$$\begin{aligned}\left|\frac{\partial f_i}{\partial x_j}\right| =&(1-\alpha)\left|\psi_i \sum_{j'} \frac{m_{j'}^2 {x_{j'}}^{1-\alpha}{o_{j'}}^{1-\alpha}}{w^2 {P_{j'}'}^{1-\alpha}}\right.\\ &\left. - \left(\frac{x_i\xi_i}{w{P_i'}^{1-\alpha}{x_i}^{\alpha}{o_i}^{\alpha}} + \psi_i \frac{{x_j}^{-\alpha}\xi_j}{w^2{P_j'}^{1-\alpha}{o_j}^{\alpha}}\right)\right|\\ \le&(1-\alpha)W_{ij}, \forall j\neq i.\end{aligned} \tag{46}$$

Since $\alpha > 1-\frac{1}{M}$, $|\mathbf{J}| \le (1-\alpha)M < 1$. Following the same approach in Theorem 7, the proposed algorithm converges to the global optimum.

Based on the following inequality

$$\sum_i m_i \frac{o_{i,\max}'}{w_{\min}'} = \frac{\sum_i m_i o_{i,\max}'}{\sum_j m_j o_{j,\min}'} \ge 1,$$

we can further show that $M > N$ after simple manipulation of W_{ij}.

Case 2: $\alpha \ge 1$.

For $\alpha = 1$, since $L=1$, the objective degrades to

$$U(\mathbf{p}) = \sum_{i=1}^N \log(\lambda_i p_i) + \log(a_1 \exp(-m_{i1}\sum_{j=1}^N P_j'\lambda_j p_j).$$

Obviously, $U(\mathbf{p})$ is concave in $\mathbf{p}$.

For $\alpha > 1$, since $L=1$, the objective degrades to

$$U(\mathbf{p}) = \sum_{i=1}^N \frac{\left(\lambda_i p_i a_1 \exp\left(-m_{i1}\sum_{j=1}^N p_j\lambda_j P_j'\right)\right)^{1-\alpha}}{1-\alpha}.$$

Let

$$g_i'(\mathbf{x}) = \frac{(x_i \exp(-m_{i1}\sum_{j=1}^N x_j))^{1-\alpha}}{(P_i')^{1-\alpha}(1-\alpha)},$$

where $\mathbf{x} \in \mathbb{R}_n^+$.

Then we have

$$\begin{aligned} g_i'(\mathbf{x}) =& \frac{(\exp(\log(x_i) - m_{il}\sum_{j=1}^N x_j))^{1-\alpha}}{(P_i')^{1-\alpha}(1-\alpha)}\\ =& \frac{\exp((1-\alpha)\log(x_i) - (1-\alpha)m_{il}\sum_{j=1}^N x_j))}{(P_i')^{1-\alpha}(1-\alpha)}\\ =& \frac{\exp(h_i'(\mathbf{x}))}{(P_i')^{1-\alpha}(1-\alpha)},\end{aligned} \tag{47}$$

where $h_i'(\mathbf{x}) = (1-\alpha)\log(x_i) - (1-\alpha)m_{il}\sum_{j=1}^N x_j$. For $\alpha > 1$, $h_i'(\mathbf{x})$ is convex in $\mathbf{x}$. Furthermore, $\exp(x)$ is convex in x and non-decreasing with respect to x. Hence, $\exp(h_i'(\mathbf{x}))$ is convex in $\mathbf{x}$ according to the composition rule. Therefore $g_i'(\mathbf{x})$ is concave in $\mathbf{x}$. We also have $U(\mathbf{p}) = a_1^{1-\alpha}\sum_{i=1}^N g_i'(\mathbf{p})$, where $g_i'(\mathbf{p}) = g_i'(\mathbf{x})|_{x_i = P_i'\lambda_i p_i}$. Therefore, we conclude that $U(\mathbf{p})$ is concave in $\mathbf{p}$.

Hence, for $\alpha \ge 1$ and $L=1$, problem $\mathcal{P}$ is a convex optimization problem. Furthermore, it is easy to verify that Slater's condition is satisfied. Thus, KKT conditions are sufficient for global optimality. Combining this with Theorem 6, we see that MMTS converges to the global optimum.

References

[1] J. Liu, N. Kato, J. Ma, and N. Kadowaki, "Device-to-device communication in LTE-Advanced networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 1923–1940, Fourthquarter 2015.

[2] 3GPP, "Study on LTE device to device proximity services; radio aspects," TR 36.843, March 2014.

[3] J. W. Lee, M. Chiang, and A. R. Calderbank, "Utility-optimal random-access control," *IEEE Trans. on Wireless Communications*, vol. 6, no. 7, pp. 2741–2751, July 2007.

[4] A. H. Mohsenian-Rad, J. Huang, M. Chiang, and V. W. S. Wong, "Utility-optimal random access: Reduced complexity, fast convergence, and robust performance," *IEEE Trans. on Wireless Communications*, vol. 8, no. 2, pp. 898–911, Feb. 2009.

[5] A. H. Mohsenian-Rad, V. W. S. Wong, and R. Schober, "Optimal SINR-based random access," in *Proc. IEEE International Conference on Computer Communications (INFOCOM)*, 2010.

[6] F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler, "An Aloha protocol for multihop mobile wireless networks," *IEEE Trans. on Information Theory*, vol. 52, no. 2, pp. 421–436, Feb. 2006.

[7] ——, "Stochastic analysis of spatial and opportunistic Aloha," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1105–1119, September 2009.

[8] F. Baccelli, B. Blaszczyszyn, and C. Singh, "Analysis of a proportionally fair and locally adaptive spatial Aloha in Poisson networks," in *Proc. IEEE International Conference on Computer Communications (INFOCOM)*, 2014.

[9] Z. Tong and M. Haenggi, "A throughput-optimum adaptive Aloha MAC scheme for full-duplex wireless networks," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, 2015.

[10] N. Lee, X. Lin, J. G. Andrews, and R. W. Heath, "Power control for D2D underlaid cellular networks: Modeling, algorithms, and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 1, pp. 1–13, Jan. 2015.

[11] A. Memmi, Z. Rezki, and M. S. Alouini, "Power control for D2D underlay cellular networks with channel uncertainty," *IEEE Trans. on Wireless Communications*, vol. 16, no. 2, pp. 1330–1343, Feb. 2017.

[12] M. H. Cheung, A. H. Mohsenian-Rad, V. W. S. Wong, and R. Schober, "Utility-optimal random access for wireless multimedia networks," *IEEE Wireless Communications Letters*, vol. 1, no. 4, pp. 340–343, August 2012.

[13] ——, "Random access for elastic and inelastic traffic in WLANs," *IEEE Trans. on Wireless Communications*, vol. 9, no. 6, pp. 1861–1866, June 2010.

[14] K. Kar, S. Sarkar, and L. Tassiulas, "Achieving proportional fairness using local information in Aloha networks," *IEEE Trans. on Automatic Control*, vol. 49, no. 10, pp. 1858–1863, Oct. 2004.

[15] M. Haenggi, "The meta distribution of the SIR in Poisson bipolar and cellular networks," *IEEE Trans. on Wireless Communications*, vol. 15, no. 4, pp. 2577–2589, April 2016.

[16] ——, "Outage, local throughput, and capacity of random wireless networks," *IEEE Trans. on Wireless Communications*, vol. 8, no. 8, pp. 4350–4359, August 2009.

[17] ——, "The local delay in Poisson networks," *IEEE Trans. on Information Theory*, vol. 59, no. 3, pp. 1788–1802, March 2013.

[18] Z. Gong and M. Haenggi, "The local delay in mobile Poisson networks," *IEEE Trans. on Wireless Communications*, vol. 12, no. 9, pp. 4766–4777, September 2013.

[19] F. Baccelli and B. Blaszczyszyn, "A new phase transitions for local delays in MANETs," in *Proc. IEEE International Conference on Computer Communications (INFOCOM)*, March 2010.

[20] M. Haenggi, "Local delay in static and highly mobile Poisson networks with Aloha," in *Proc. IEEE International Conference on Communications (ICC)*, May 2010.

[21] ——, "Local delay in Poisson networks with and without interference," in *Proc. Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sep. 2010.

[22] Y. Xu, T. Le-Ngoc, and S. Panigrahi, "Global concave minimization for optimal spectrum balancing in multi-user DSL networks," *IEEE Trans. on Signal Processing*, vol. 56, no. 7, pp. 2875–2885, July 2008.

[23] J. Papandriopoulos and J. S. Evans, "Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks," in *Proc. IEEE International Conference on Communications (ICC)*, 2006.

[24] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[25] J. Lyu, Y. H. Chew, and W. Wong, "Efficient and scalable distributed autonomous spatial Aloha networks via local leader election," *IEEE Trans. on Vehicular Technology*, vol. 65, no. 12, pp. 9954–9967, Dec. 2016.

[26] M. H. Cheung, V. W. S. Wong, and R. Schober, "SINR-based random access for cognitive radio: Distributed algorithm and coalitional game," *IEEE Trans. on Wireless Communications*, vol. 10, no. 11, pp. 3887–3897, Nov. 2011.

[27] J. C. Ikuno, M. Wrulich, and M. Rupp, "System level simulation of LTE networks," in *Proc. IEEE Vehicular Technology Conference (VTC)*, May 2010.

[28] 3GPP, "E-UTRA; physical layer procedures," TS 36.213, October 2009.

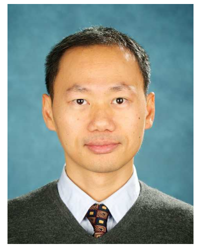
[29] ——, "Study on RAN Improvements for Machine-type Communications," TR 37.868, March 2011.

[30] D. P. Bertsekas, *Nonlinear programming*. Athena scientific Belmont, 1999.

[31] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681–683, August 1978.

[32] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*. Prentice hall, 1989.

Yujie Xu (S'18) received B.E. and M.E. degrees from the Beijing University of Posts and Telecommunications, Beijing, China, in 2011 and 2014, respectively. He is currently a Ph.D. student at the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada. His research interests include resource allocation in IoT networks and application of learning techniques to wireless networks.



Ben Liang (S'94-M'01-SM'06-F'18) received honors-simultaneous B.Sc. (valedictorian) and M.Sc. degrees in Electrical Engineering from Polytechnic University in Brooklyn, New York, in 1997 and the Ph.D. degree in Electrical Engineering with a minor in Computer Science from Cornell University in Ithaca, New York, in 2001. In the 2001-2002 academic year, he was a visiting lecturer and post-doctoral research associate at Cornell University. He joined the Department of Electrical and Computer Engineering at the University of Toronto in 2002, where he is now a Professor. His current research interests are in networked systems and mobile communications. He has served on the editorial boards of the IEEE Transactions on Mobile Computing since 2017 and the IEEE Transactions on Communications since 2014, and he was an editor for the IEEE Transactions on Wireless Communications from 2008 to 2013 and an associate editor for Wiley Security and Communication Networks from 2007 to 2016. He regularly serves on the organizational and technical committees of a number of conferences. He is a Fellow of IEEE and a member of ACM and Tau Beta Pi.

Gary Boudreau (M'84-SM'11) received a B.A.Sc. in Electrical Engineering from the University of Ottawa in 1983, an M.A.Sc. in Electrical Engineering from Queens University in 1984 and a Ph.D. in electrical engineering from Carleton University in 1989. From 1984 to 1989 he was employed as a communications systems engineer with Canadian Astronautics Limited and from 1990 to 1993 he worked as a satellite systems engineer for MPR Teltech Ltd. For the period spanning 1993 to 2009 he was employed by Nortel Networks in a variety of wireless systems and management roles within the CDMA and LTE base station product groups. In 2010 he joined Ericsson Canada where he is currently employed in the 5G systems architecture group. His interests include digital and wireless communications as well as digital signal processing.

Hossein Seyedmehdi received a Masters degree from the National University of Singapore in 2008 and a PhD degree from the University of Toronto in 2014 both in Electrical and Computer Engineering. He was affiliated with Ericsson Canada until 2017 where he worked on the 5th generation (5G) of wireless technologies including the commercialization of massive MIMO. He has numerous papers in the area of wireless systems in addition to more than 10 patents. His research interests include Information Theory for Wireless Communications, Signal Processing for MIMO channels, Algorithms for Radio Access Networks, and Future Radio Technologies.