Wireless Max–Min Utility Fairness With General Monotonic Constraints by Perron–Frobenius Theory

Liang Zheng, *Member, IEEE*, Y.-W. Peter Hong, *Senior Member, IEEE*, Chee Wei Tan, *Senior Member, IEEE*, Cheng-Lin Hsieh, and Chia-Han Lee, *Member, IEEE*

***Abstract*— This paper presents a systematic approach for solving wireless max–min utility fairness optimization problems in multiuser wireless networks with general monotonic constraints. These problems are often challenging to solve due to their nonconvexity. By establishing a connection between this class of optimization problems and the class of conditional eigenvalue problems that can be addressed by a generalized nonlinear Perron–Frobenius theory, we show how these problems can be solved optimally using an iterative algorithm that converges geometrically fast. The mathematical development in this paper unifies previous work and allows us to handle a broader class of competitive utility functions with general nonlinear monotonic constraints. Several representative applications illustrate the effectiveness of the proposed framework, including the max–min quality-of-service subject to robust interference temperature constraints in cognitive radio networks, the min–max weighted mean-square error subject to signal-to-interference-and-noise ratio constraints in multiuser downlink systems, the max–min throughput subject to nonlinear power constraints in energy-efficient wireless networks, the max–min sigmoid utility in multimedia wireless networks, and the min–max outage probability subject to outage constraints in heterogeneous wireless networks. Numerical results are presented to demonstrate the fast-convergence behavior of the algorithms to the optimal fixed-point solution characterized by our generalized nonlinear Perron–Frobenius theoretic framework.**

***Index Terms*— Optimization, max-min utility fairness, nonlinear Perron-Frobenius theory, wireless resource allocation.**

Manuscript received June 29, 2015; revised April 22, 2016; accepted August 8, 2016. Date of publication October 5, 2016; date of current version November 18, 2016. This work was supported in part by the Ministry of Science and Technology, Taiwan, under Grant 105-2218-E-007-010 and Grant 104-2918-I-007-006, in part by the Research Grants Council of Hong Kong, under Grant RGC 125212 and Grant 11212114, in part Qualcomm Inc., and in part by the Science, Technology and Innovation Commission of Shenzhen Municipality, under Project JCYJ20160229165220746. This paper was presented in part at the 2014 IEEE International Conference on Computer Communications [1]. *(Corresponding author: Y.-W. Peter Hong.)*

L. Zheng is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: liangz@princeton.edu).

Y.-W. P. Hong is with the Institute of Communications Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan (e-mail: ywhong@ee.nthu.edu.tw).

C. W. Tan is with the Department of Computer Science, City University of Hong Kong, Hong Kong (e-mail: cheewtan@cityu.edu.hk).

C.-L. Hsieh was with the Institute of Communications Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan. He is now with ASUSTek Computer Inc., Taipei 11262, Taiwan. (e-mail: ian_hsieh@asus.com).

C.-H. Lee is with the Department of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan (e-mail: chiahan@nctu.edu.tw).

Communicated by T. Javidi, Associate Editor for Communication Networks.

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/TIT.2016.2615183

I. Introduction

The demand for broadband mobile data services has grown rapidly in recent years. Furthermore, new wireless services such as unlicensed spectrum usage in cognitive radio networks and spectrum sharing in heterogeneous wireless networks create multiuser interference that has a significant impact on wireless network performance. Joint interference and resource management is thus critical to ensure that sharing is done fairly in a collaborative and efficient manner. However, maintaining a balanced and fair operation becomes increasingly difficult as multiuser interference increases with the denser ad-hoc deployment of wireless networks.

Fairness in wireless networks can be modeled by nonlinear utility functions of wireless link metrics, such as signal-to-interference-and-noise ratio (SINR), mean-square error (MSE) and outage probability [2]–[8]. The seminal work in [2] is the earliest to propose balancing SINR fairness in satellite communications, and this was later adopted in wireless cellular networks [9]. These link metrics are useful because the SINR measures the received signal quality (which impacts the data rate), MSE characterizes the error of the estimated transmit symbol at the receiver, and the outage probability measures the likelihood that the received SINR falls below a given threshold. These wireless link metrics are often affected by channel conditions, interference, and quality-of-service requirements, e.g., interference temperature constraints or outage probability constraints in cognitive radio networks. The total network fairness utility is then maximized over the joint solution space of wireless link variables such as power. The main challenges of solving these wireless utility maximization problems come from the highly coupled and nonlinear dependency between link metrics and optimization variables, making these nonconvex optimization problems notoriously difficult to solve. Designing scalable and low-complexity algorithms to solve them optimally is even harder.

In [4], [5], and [10], techniques based on geometric programming (GP) were proposed to solve a certain class of nonconvex wireless utility maximization problems that can be approximately transformed into convex ones. In [6], Gibbs sampling techniques were used to solve nonconvex wireless utility maximization problems, but the optimality of the solutions were not guaranteed. In [7] and [8], deterministic wireless utility maximization problems that depend on the users' transmission rates and powers were examined. More recently, in [11]–[16], a specialized (i.e., concave) form of nonlinear

Perron-Frobenius theory [17]–[19] was applied to tackle max-min utility fairness problems in wireless networks. We refer the readers to [19] for an overview of this nonlinear Perron-Frobenius theoretic approach. These works showed that a certain class of max-min utility fairness problems can be equivalently converted into conditional eigenvalue problems with concave self-mappings and can be solved using fixed-point algorithms with a norm normalization in each iteration. However, these works were only able to incorporate affine power constraints but cannot handle other types of nonlinear constraints typically encountered in wireless networks, e.g., nonlinear power constraints, stochastic interference temperature constraints and outage probability constraints.

In this work, we broaden the class of solvable wireless max-min utility optimization problems to those with general monotonic constraints (that could not be handled by existing works [11]–[16]). This is done by establishing the mathematical connection between the problems mentioned above and the conditional eigenvalue problems that can be solved by a generalized form of nonlinear Perron-Frobenius theory presented in [20]. By doing so, a broader class of wireless max-min utility optimization problems with general monotonic constraints can be solved optimally using an algorithm that converges geometrically fast. Compared to [11]–[16], the connection between application and theory is less apparent in our case due to the inclusion of general monotonic constraints and, thus, requires mathematical development that is more technically involved. In particular, the key lies in the careful choice of a monotone scale function that projects the solution obtained in each iteration to its largest feasible value in the same direction (c.f., Section III-B). We provide a list of assumptions on the utility functions and constraints in order for our optimization framework to hold, and examine five representative case studies to demonstrate its wide applicability. The case studies include problems in cognitive radio networks, multiuser downlink systems, energy-efficient wireless networks, multimedia-based wireless networks and heterogeneous cellular networks. In particular, these case studies incorporate nonconvex monotonic constraints such as nonlinear power constraints, interference temperature constraints, and outage probability constraints. Numerical simulations are then provided for all case studies to demonstrate the effectiveness of the proposed technique.

The remainder of this article is organized as follows. In Section II, we provide a brief review of classical Perron-Frobenius theory, its concave version in [17]–[19], and the generalized nonlinear version that is to be adopted in this work, i.e., [20]. In Section III, we provide the mathematical development required to establish the connection between wireless max-min utility optimization problems with general monotonic constraints and conditional eigenvalue problems by leveraging the generalized nonlinear Perron-Frobenius theory. The five case studies and their corresponding experiments are presented in Sections IV and V, respectively. Finally, we conclude the paper in Section VI. All of the proofs can be found in the Appendix.

Notations: Let $\mathcal{R}^m$ be the m-dimensional Euclidean space. For $\mathbf{x} = [x_1, \ldots, x_m]^\top$ and $\mathbf{y} = [y_1, \ldots, y_m]^\top \in \mathcal{R}^m$, we have $\mathbf{x} < \mathbf{y}$ if $x_i < y_i$, $\forall i$; $\mathbf{x} \leq \mathbf{y}$ if $x_i \leq y_i$, $\forall i$, but $\mathbf{x} \neq \mathbf{y}$; and $\mathbf{x} \leqq \mathbf{y}$ if $x_i \leq y_i$, $\forall i$. Moreover, let $\mathcal{R}^m_+$ be the set of vectors $\mathbf{x} \in \mathcal{R}^m$ with $\mathbf{x} \geqq \mathbf{0}$.

II. Brief Review of Classical and Nonlinear Perron Frobenius Theory

In this section, we provide a brief review of the classical (linear) Perron-Frobenius Theory, which has been instrumental in [2], and its concave extension with applications in [11]–[16]. This provides the basis for an extension to the more generalized nonlinear case, which is to be addressed in this paper. The generalized theory unifies previous works, leading to a systematic way of characterizing the optimal solution and of designing optimal algorithms to solve a broader class of nonconvex wireless network optimization problems [19].

The classical (linear) Perron-Frobenius theory asserts that an irreducible nonnegative matrix as a linear mapping has a unique spectral radius with a unique unit-norm eigenvector that has strictly positive entries. A basic form is stated below.

Theorem 1 [21]: Let $\mathbf{T} : \mathcal{R}^L_+ \to \mathcal{R}^L_+$ be a linear mapping defined by $\mathbf{T}(\mathbf{x}) = \mathbf{A}\mathbf{x}$, for any $\mathbf{x} \in \mathcal{R}^L_+$, where $\mathbf{A}$ is an $L \times L$ irreducible nonnegative matrix. Then,

(a) the conditional eigenvalue problem $\mathbf{T}(\mathbf{x}) = \lambda \mathbf{x}$ has a unique solution with $\mathbf{x}^* > \mathbf{0}$, $\|\mathbf{x}^*\| = 1$ and $\lambda^* > 0$;
(b) the solution is $\mathbf{x}^\star = \lim_{n\to\infty} \tilde{\mathbf{T}}^n(\mathbf{x})$, where $\tilde{\mathbf{T}}(\mathbf{x}) = \mathbf{T}(\mathbf{x})/\|\mathbf{T}(\mathbf{x})\|$, for all $\mathbf{x} \geq \mathbf{0}$.

This theorem characterizes both the eigenvalue property of irreducible nonnegative matrices and the fixed-point of linear mappings. It has important applications in areas such as probability theory (for analyzing Markov chains), dynamical systems, economics, and Internet search engines (Google Pagerank algorithm) [21], [22], etc.

Moreover, due to the need to address nonlinear mappings in many applications, e.g., in control theory [23], computer science [24], mathematical biology [25], the above theorem has also been extended to its concave version in [17] and [18], where $\mathbf{T} : \mathcal{R}^L_+ \to \mathcal{R}^L_+$ satisfies $\mathbf{T}(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \geq \alpha\mathbf{T}(\mathbf{x}) + (1-\alpha)\mathbf{T}(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{R}^L_+$ and all $\alpha \in [0, 1]$ and $\mathbf{T}(\mathbf{x}) > \mathbf{0}$ for all $\mathbf{x} \geq \mathbf{0}$. In [11]–[16], the authors showed that a certain class of max-min utility optimization problems with linear power constraints can be equivalently converted into conditional eigenvalue problems similar to that in Theorem 1(a), but with $\mathbf{T}$ being concave, and can be solved by an iterative algorithm implied by Theorem 1(b) that converges geometrically fast, i.e., there exists $\delta \in (0, 1)$ such that $\|\tilde{\mathbf{T}}^n(\mathbf{x}) - \mathbf{x}^*\| = \mathcal{O}(\delta^n)$. In particular, the normalization by $\|\mathbf{T}(\mathbf{x})\|$ in Theorem 1(b) allows one to consider norm or affine constraints on the optimization variable $\mathbf{x}$ [11]–[16].

In this work, we consider a further generalization of the nonlinear Perron-Frobenius theory where the normalization is generalized by a monotone scale function β as stated below.

Theorem 2 [20]: Suppose that $\beta : \mathcal{R}^L_+ \to \mathcal{R}_+$ is not identically 0, positively homogeneous (i.e., $\beta(\lambda\mathbf{x}) = \lambda\beta(\mathbf{x})$ for $\mathbf{x} \geqq \mathbf{0}$ and $\lambda \geqq 0$), and monotonic (i.e., $\mathbf{0} \leqq \mathbf{x} \leqq \mathbf{y}$ implies $\beta(\mathbf{x}) \leqq \beta(\mathbf{y})$), and that $\mathbf{T} : \mathcal{R}^L_+ \to \mathcal{R}^L_+$ satisfies the following conditions: (i) there exists $a > 0$, $b > 0$, and a vector $\mathbf{e} > \mathbf{0}$ such that $a\mathbf{e} \leqq \mathbf{T}(\mathbf{x}) \leqq b\mathbf{e}$, for all $\mathbf{x} \in \mathcal{R}^L_+$ with $\beta(\mathbf{x}) = 1$;

(ii) for any $\mathbf{x}, \mathbf{y} \in \mathcal{R}_+^L$ with $\beta(\mathbf{x}) = \beta(\mathbf{y}) = 1$ and $0 \leqq \lambda \leqq 1$: If $\lambda\mathbf{x} \leqq \mathbf{y}$, then $\lambda\mathbf{T}(\mathbf{x}) \leqq \mathbf{T}(\mathbf{y})$; and, if $\lambda\mathbf{x} \leq \mathbf{y}$ with $\lambda < 1$, then $\lambda\mathbf{T}(\mathbf{x}) < \mathbf{T}(\mathbf{y})$. Then, the following properties hold:

(a) $\lambda\mathbf{x} = \mathbf{T}(\mathbf{x})$ has a unique solution $\mathbf{x}^* \in \mathcal{R}_+^L$ with $\beta(\mathbf{x}^*) = 1$ and $\lambda^* > 0$.

(b) $\mathbf{x}^* = \lim_{n\to\infty} \tilde{\mathbf{T}}^n(\mathbf{x})$, where $\tilde{\mathbf{T}}(\mathbf{x}) = \mathbf{T}(\mathbf{x})/\beta(\mathbf{T}(\mathbf{x}))$, for any $\mathbf{x} \geqq \mathbf{0}$ with $\beta(\mathbf{T}(\mathbf{x})) > 0$.

Interestingly, the class of functions $\mathbf{T}$ (i.e., functions that satisfy conditions (i) and (ii)) and monotone scale functions $\beta(\mathbf{x})$ considered in Theorem 2 include as special cases the concave mappings and norm normalizations considered in the concave Perron-Frobenius theory in [17] and [18]. In particular, the monotone function $\beta(\mathbf{x})$ considered in Theorem 2 is key to incorporating general monotonic constraints, such as stochastic interference temperature constraints, outage probability constraints, and other nonlinear power constraints, that could not be handled before. Unlike previous works [11]–[16] that only address linear power constraints, the mapping from the monotonic constraints in the considered optimization problem to the monotone scale function $\beta(\mathbf{x})$ in Theorem 2 is nontrivial. This is done in the following by determining a suitable choice of $\beta(\mathbf{x})$ in terms of the monotone constraints and by characterizing its optimality conditions.

III. General Max-Min Utility Optimization With Monotonic Constraints

Let us consider a general wireless network with L users and let $p_1, \ldots, p_L$ be the transmit powers of the L users. In particular, p_i is the transmit power of the i-th user. We assume that the power vector $\mathbf{p} = [p_1, \ldots, p_L]^\top$ can be adjusted to optimize the overall network utility. Based on the generalization of nonlinear Perron-Frobenius theory [20], we develop in the following an optimization framework to solve a class of max-min utility optimization problems under general monotonic constraints. We make some reasonable assumptions below that the utility functions and constraints of the optimization problem must satisfy in order for the proposed framework to apply, but these assumptions do not require either convexity or linearity of the optimization problem. In fact, the proposed techniques are particularly useful when convexity and linearity do not hold and other existing techniques cannot be applied (c.f. Section IV).

A. Problem Formulation

Let $u_i : \mathcal{R}_+^L \to \mathcal{R}_+$, for $i \in \{1, \ldots, L\}$, be a continuous function of $\mathbf{p}$ that models the utility of user i and let $g_k : \mathcal{R}_+^L \to \mathcal{R}_+$, for $k \in \{1, \ldots, K\}$, be a continuous function of $\mathbf{p}$ that is used to specify the k-th system constraint. The class of problems considered in this work can be formulated by:

$$\text{maximize} \quad \min_{i=1,\ldots,L} \; u_i(\mathbf{p}) \tag{1a}$$

$$\text{subject to } \mathbf{g}(\mathbf{p}) \leqq \bar{\mathbf{g}} \tag{1b}$$

$$\text{variables} : \mathbf{p}, \tag{1c}$$

where $\mathbf{g}(\mathbf{p}) = [g_1(\mathbf{p}), \ldots, g_K(\mathbf{p})]^\top$ and $\bar{\mathbf{g}} = [\bar{g}_1, \ldots, \bar{g}_K]^\top$ are vectors of constraint functions and values, respectively.

Specifically, the class of utility functions in (1a) must satisfy the following assumptions.

Assumption 1 (Competitive Utility Functions):

- *Positivity*: For all i, $u_i(\mathbf{p}) > 0$ if $\mathbf{p} > \mathbf{0}$ and, in addition, $u_i(\mathbf{p}) = 0$ if and only if $p_i = 0$.
- *Competitiveness*: For all i, u_i is strictly increasing with respect to p_i and is strictly decreasing with respect to p_j, for $j \neq i$, when $p_i > 0$.
- *Directional Monotonicity*: For $\lambda > 1$ and $p_i > 0$, $u_i(\lambda\mathbf{p}) > u_i(\mathbf{p})$, for all i.

The competitiveness assumption models the interaction between users in a wireless network and the directional monotonicity captures the increase in utility as the total power consumption increases.

Moreover, the class of constraints considered in (1b), called *monotonic constraints*, must satisfy the following assumptions.

Assumption 2 (Monotonic Constraints):

- *Strict Monotonicity*: For all k, $g_k(\mathbf{p}_1) > g_k(\mathbf{p}_2)$ if $\mathbf{p}_1 > \mathbf{p}_2$, and $g_k(\mathbf{p}_1) \geqq g_k(\mathbf{p}_2)$ if $\mathbf{p}_1 \geqq \mathbf{p}_2$.
- *Feasibility*: The set $\{\mathbf{p} > \mathbf{0} : \mathbf{g}(\mathbf{p}) \leqq \bar{\mathbf{g}}\}$ is non-empty.
- *Validity*: For any $\mathbf{p} > \mathbf{0}$, there exists $\lambda > 0$ such that $g_k(\lambda\mathbf{p}) \geqq \bar{g}_k$, for some k.

The strict monotonicity captures the increase in cost or resource consumption as $\mathbf{p}$ increases, the feasibility ensures that there exists a positive power vector in the feasible set, and the validity ensures that the set of constraints is meaningful. If the validity condition does not hold, the corresponding constraint can be simply removed without loss of generality. Due to the monotonicity of the functions $\{g_k\}_{k=1}^K$, we shall refer to (1b) as the set of monotonic constraints.

Notice that Assumptions 1 and 2 are necessary to equivalently convert the optimization problem into the conditional eigenvalue problem to be described in Section III-B. Before deriving the general solution for (1), let us give some practical examples of such objectives and constraints in the following. In particular, the class of utility functions satisfying Assumption 1 can include many standard performance metrics in wireless networks, e.g.,

- the SINR under frequency-flat fading, i.e.,

$$u_i(\mathbf{p}) = \mathsf{SINR}_i(\mathbf{p}) \triangleq \frac{G_{ii}p_i}{\sum_{j=1, j\neq i}^{L} G_{ij}p_j + \eta_i}, \tag{2}$$

where G_{ij} is the channel fading gain from transmitter j to receiver i, and η_i is the receiver noise variance;

- the link capacity, i.e.,

$$u_i(\mathbf{p}) = \log\left(1 + \mathsf{SINR}_i(\mathbf{p})\right), \tag{3}$$

where the SINR is given by (2);

- the reliability function (which is the complement of the outage probability [5], [11]), i.e.,

$$\begin{aligned} u_i(\mathbf{p}) &= \Pr\left(\mathsf{SINR}_i^R(\mathbf{p}) \geq \gamma_i\right) \\ &\triangleq e^{\frac{-\gamma_i\eta_i}{G_{ii}p_i}} \prod_{j=1, j\neq i}^{L} \left(1 + \frac{\gamma_i G_{ij} p_j}{G_{ii} p_i}\right)^{-1}, \end{aligned} \tag{4}$$

where γ_i is the SINR threshold of the i-th user and $\mathsf{SINR}_i^R(\mathbf{p})$ is the SINR received at the i-th receiver under Rayleigh fading. In particular, we have

$$\mathsf{SINR}_i^R(\mathbf{p}) \triangleq \frac{G_{ii}R_{ii}p_i}{\sum_{j=1,j\neq i}^{L} G_{ij}R_{ij}p_j + \eta_i}, \tag{5}$$

where $\{R_{ij}\}$ are independent and identically distributed (i.i.d.) exponential random variables with unit mean. Note that the SINR under Rayleigh fading in (5) is a random variable whose statistics depend on $\mathbf{p}$.

The class of monotonic constraints satisfying Assumption 2 also includes a large class of constraints that are representative in wireless networks, such as

- linear power constraints, i.e.,

$$\mathbf{g}(\mathbf{p}) = \mathbf{A}\mathbf{p} \leqq \bar{\mathbf{p}}, \tag{6}$$

where $\bar{\mathbf{p}} \in \mathbb{R}_+^{K\times 1}$ is a positive upper bound vector for power constraints and $\mathbf{A} \in \mathbb{R}_{\geq 0}^{K\times L}$ is a nonnegative weight matrix (e.g., $K = L$ and $\mathbf{A} = \mathbf{I}$ in the case of individual power constraints, and $K = 1$ and $\mathbf{A} = \mathbf{1}^\top$, i.e., $\mathbf{A}$ is an all-one row vector, in the case of sum power constraints);
- stochastic interference temperature constraints [26],

$$\mathbf{g}(\mathbf{p}) \leqq \boldsymbol{\varepsilon} \tag{7}$$

where $g_k(\mathbf{p}) = \Pr(\mathbf{w}_k^T\mathbf{p} \geq t_k)$, for $k = 1,\dots,K$, $\boldsymbol{\varepsilon} = [\varepsilon_1,\dots,\varepsilon_K]^T$, and $\mathbf{w}_k$ is the vector of interference channel gains.
- nonlinear power constraints [27], i.e.,

$$\mathbf{g}(\mathbf{p}) = \mathbf{J}(\mathbf{p}) \leqq \bar{\mathbf{p}}, \tag{8}$$

where $\mathbf{J} : \mathcal{R}_+^L \to \mathcal{R}_+^K$ is a nonlinear monotonic function of the power vector and can be used to model nonlinearities in the circuit, e.g., the power amplifier [28].

In many cases, u_i and g_k in (1) are complicated nonconvex functions in $\mathbf{p}$, making the problem difficult to solve using standard convex optimization approaches [29]. In some special cases, GP (or related approximation algorithms) [10] can be used to obtain efficient solutions, but it cannot be used to address the general case optimally. Leveraging results from the generalized nonlinear Perron-Frobenius theory, we derive next an efficient iterative algorithm to optimally solve the class of problems described by (1) under Assumptions 1 and 2.

B. General Solution

By introducing an auxiliary variable τ, the problem in (1) can be reformulated as:

$$\begin{aligned} &\text{maximize } \tau && \text{(9a)}\\ &\text{subject to } u_i(\mathbf{p}) \geq \tau, \quad i = 1,\dots,L, && \text{(9b)}\\ &\qquad\qquad g_k(\mathbf{p}) \leq \bar{g}_k, \quad k = 1,\dots,K, && \text{(9c)}\\ &\text{variables} : \mathbf{p},\ \tau. && \text{(9d)} \end{aligned}$$

We shall refer to the constraints in (9b) as the objective constraints and those in (9c) as the system constraints.

Lemma 1: For $\{u_i\}_{i=1}^L$, $\{g_k\}_{k=1}^K$, and $\{\bar{g}_k\}_{k=1}^K$ that satisfy Assumptions 1 and 2, the optimal solution $(\tau^*, \mathbf{p}^*)$ is positive, i.e., $\tau^* > 0$ and $\mathbf{p}^* > \mathbf{0}$, and, at optimality, all objective constraints are tight and at least one system constraint is active. That is, $u_i(\mathbf{p}^*) = \tau^*$, for all i, and $g_k(\mathbf{p}^*) = \bar{g}_k$, for some k.

Lemma 1 states the tightness of both objective constraints in (9b) and system constraints in (9c), which can be proved by the competitiveness and directional monotonicity of utility functions in Assumption 1. The proof is given in Appendix A. By Lemma 1 and the positivity of the utility functions (i.e., $u_i(\mathbf{p}^*) > 0$, for all i), it follows that

$$\frac{1}{\tau^*}p_i^* = \frac{1}{u_i(\mathbf{p}^*)}p_i^*,\ i = 1,\dots,L, \tag{10}$$

for which, we define the following function.

Definition 1: The function $T_i : \mathcal{R}_+^L \to \mathcal{R}_+$ of each user's power p_i is defined as

$$T_i(\mathbf{p}) \triangleq \frac{1}{u_i(\mathbf{p})}p_i, \quad i = 1,\dots,L. \tag{11}$$

This means that the optimal power vector $\mathbf{p}^*$ is a solution to the fixed point equation $\frac{1}{\tau^*}\mathbf{p}^* = [T_1(\mathbf{p}^*),\dots,T_L(\mathbf{p}^*)]^\top \triangleq \mathbf{T}(\mathbf{p}^*)$ with T_i defined in (11).

Definition 2: The function $\beta : \mathcal{R}_+^L \to \mathcal{R}_+$ of $\mathbf{p}$ is defined as

$$\beta(\mathbf{p}) \triangleq \min\{\beta' \geqq 0 : g_k(\mathbf{p}/\beta') \leqq \bar{g}_k, \forall k\}. \tag{12}$$

We term $\beta(\mathbf{p})$ the *scale* of $\mathbf{p}$.

The monotonic constraints stated in Assumption 2 leads to the following lemma for the scale $\beta(\mathbf{p})$ defined in Definition 2. The proof is given in Appendix B.

Lemma 2: The scale $\beta : \mathcal{R}_+^L \to \mathcal{R}_+$ defined in Definition 2 satisfies the following properties:

1) β is not identically zero, and $\beta(\mathbf{p}) > 0$, for all $\mathbf{p} > \mathbf{0}$;
2) $\beta(\lambda\mathbf{p}) = \lambda\beta(\mathbf{p})$ for $\mathbf{p} \geqq \mathbf{0}$ and $\lambda \geqq 0$ (i.e., positively homogeneous);
3) $\mathbf{0} \leqq \mathbf{p} \leqq \mathbf{q}$ implies $\beta(\mathbf{p}) \leqq \beta(\mathbf{q})$ (i.e., monotonic).

Notice that, when $\mathbf{p} > 0$, the scale $\beta(\mathbf{p}) > 0$ (c.f. Lemma 2) is chosen such that, after normalization, the power vector $\tilde{\mathbf{p}} = \mathbf{p}/\beta(\mathbf{p})$ yields the largest feasible solution in the direction of $\mathbf{p}$. In this case, at least one of the system constraints must be satisfied with equality, i.e., $g_k(\mathbf{p}/\beta(\mathbf{p})) = \bar{g}_k$ for some k.

Definition 3: The set $\mathcal{U}$ for $\mathbf{p}$ is defined as

$$\mathcal{U} \triangleq \left\{\mathbf{p} \geqq \mathbf{0} : \beta(\mathbf{p}) = 1\right\}. \tag{13}$$

Lemma 1 implies that the set of feasible solutions of $\mathbf{p}$ must include, as a subset, the set of $\mathbf{p}$ that has scale equal to 1, i.e., $\left\{\mathbf{p} \geqq \mathbf{0} : g_k(\mathbf{p}) \leqq \bar{g}_k, \forall k\right\} \supseteq \mathcal{U}$. We formalize this implication in the following lemma.

Lemma 3: The optimal solution of (1) is included in $\mathcal{U}$.

The proof is given in Appendix C. By Lemma 3, it follows that the solution of the optimization problem in (9) (and, thus, (1)) is a solution of the conditional eigenvalue problem [20], where the objective is to find $(\tau^*, \mathbf{p}^*)$ such that

$$\frac{1}{\tau^*}\mathbf{p}^* = \mathbf{T}(\mathbf{p}^*),\ \tau^* \in \mathcal{R},\ \mathbf{p}^* \in \mathcal{U}, \tag{14}$$

where $\mathbf{T}(\mathbf{p}) = [T_1(\mathbf{p}),\dots,T_L(\mathbf{p})]^\top$ and $T_i(\mathbf{p})$ is defined in Definition 1. Note that the equivalence of (9) and (14) relies on

Algorithm 1 Max-Min Utility Optimization Under Monotonic Constraints

1) Initialize power vector $\mathbf{p}(0) > \mathbf{0}$.
2) Update power vector $\mathbf{p}(t+1)$:

$$p_i(t+1) = T_i(\mathbf{p}(t)), \quad i = 1, \ldots, L. \tag{15}$$

3) Scale power vector $\mathbf{p}(t+1)$:

$$p_i(t+1) \leftarrow \frac{p_i(t+1)}{\beta(\mathbf{p}(t+1))}, \quad i = 1, \ldots, L, \tag{16}$$

where $\beta(\mathbf{p}(t+1))$ is calculated by using Algorithm 2 or 3 (c.f., Section III-C).
4) Repeat Steps 2 and 3 until convergence.

Assumptions 1 and 2, which holds for many practical scenarios in wireless networking, as to be shown in Section IV. Interestingly, the conditional eigenvalue problem in (14) is exactly the problem tackled by the generalized nonlinear Perron-Frobenius Theory in Theorem 2(a) with $\lambda^* = \frac{1}{\tau^*}$. Hence, given that $\mathbf{T}$ satisfies conditions (i) and (ii) in Theorem 2 and by Lemma 2, the solution can be found by Theorem 2(b), which leads to the iterative procedure in Algorithm 1.

In particular, Algorithm 1 consists of two main steps: the iterative computation of $\mathbf{p}(t+1) = \mathbf{T}(\mathbf{p}(t))$ (i.e., Step 2) and the scaling by $\beta(\mathbf{p}(t+1))$ (i.e., Step 3). The scaling by $\beta(\mathbf{p}(t+1))$ projects the solution $\mathbf{p}(t+1)$ back into the feasible set $\mathcal{U}$ (i.e., the largest feasible solution in its direction). The value of $\beta(\mathbf{p}(t+1))$ can be efficiently computed by Algorithm 2 or 3 as to be described in Subsection III-C. By combining the above two steps, we effectively compute the function $\tilde{\mathbf{T}}(\mathbf{p}(t+1)) = \mathbf{T}(\mathbf{p}(t+1))/\beta(\mathbf{T}(\mathbf{p}(t+1)))$ which can be shown to be a contraction mapping under conditions (i) and (ii) [20]. The existence and the uniqueness of the solution of the conditional eigenvalue problem in (14) and the convergence of Algorithm 1 can be guaranteed by Theorem 2, which is restated below for completeness, using notations tailored to the problem at hand.

Theorem 3: Suppose that $\mathbf{T} : \mathcal{R}_+^L \rightarrow \mathcal{R}_+^L$, where $\mathbf{T}(\mathbf{p}) = [T_1(\mathbf{p}), \ldots, T_L(\mathbf{p})]^\top$ and $T_i(\mathbf{p})$ is defined in Definition 1, satisfies the following conditions: (i) there exist numbers $a > 0$, $b > 0$, and a vector $\mathbf{e} > \mathbf{0}$ such that $a\mathbf{e} \leqq \mathbf{T}(\mathbf{p}) \leqq b\mathbf{e}$, for all $\mathbf{p} \in \mathcal{U}$; (ii) for any $\mathbf{p}$, $\mathbf{q} \in \mathcal{U}$ and $0 \leqq \lambda \leqq 1$: If $\lambda\mathbf{p} \leqq \mathbf{q}$, then $\lambda\mathbf{T}(\mathbf{p}) \leqq \mathbf{T}(\mathbf{q})$; and, if $\lambda\mathbf{p} \leq \mathbf{q}$ with $\lambda < 1$, then $\lambda\mathbf{T}(\mathbf{p}) < \mathbf{T}(\mathbf{q})$. Then, the following properties hold:

(a) The conditional eigenvalue problem in (14) has a unique solution $\mathbf{p}^* \in \mathcal{U}$ and $\tau^* > 0$.
(b) The power vector $\mathbf{p}(t)$ in Algorithm 1 converges to $\mathbf{p}^*$ (i.e., the solution of (14) and, thus, (1)) for any initial point $\mathbf{p}(0) \geqq \mathbf{0}$ with $\beta(\mathbf{T}(\mathbf{p}(0))) > 0$.

The above procedure allows one to solve the max-min utility optimization problem optimally since the convergence to a unique solution must be global. The following corollary follows directly from the results in [20] and gives an immediate consequence of the conditions in Theorem 3.

Corollary 1: For $\beta(\mathbf{p})$ that is zero only when $\mathbf{p} = \mathbf{0}$, the properties in Theorem 3 hold for $\mathbf{T}$ that is positive and

Algorithm 2 Computation of β via Bisection Search

Initialization:
i) Set $i \leftarrow 0$, $L \leftarrow 0$, and $U \leftarrow 2^i$.
ii) If there exists k such that $g_k(\mathbf{p}/U) > \bar{g}_k$, then increment $i \leftarrow i+1$ and set $U \leftarrow 2^i$.
iii) Repeat (ii) until $g_k(\mathbf{p}/U) \leq \bar{g}_k$ for all k.

Bisection Search
1) Set $\beta \leftarrow (U+L)/2$. If $g_k(\mathbf{p}/\beta) > \bar{g}_k$ for some k, then set $L \leftarrow \beta$. Otherwise, set $U \leftarrow \beta$.
2) Repeat until $|U - L| < \epsilon$.

concave, i.e., $\mathbf{T}(\mathbf{p}) > \mathbf{0}$ for $\mathbf{p} \geq \mathbf{0}$ and

$$\mathbf{T}(\lambda\mathbf{p} + (1-\lambda)\mathbf{q}) \geqq \lambda\mathbf{T}(\mathbf{p}) + (1-\lambda)\mathbf{T}(\mathbf{q}), \tag{17}$$

for all $\mathbf{p} \geqq \mathbf{0}$, $\mathbf{q} \geqq \mathbf{0}$, and $0 \leqq \lambda \leqq 1$.

Even though conditions (i) and (ii) in Theorem 3 are more general, the conditions in Corollary 1 are easier to verify and are, in fact, sufficient for most applications. The requirement that "$\beta(\mathbf{p})$ is zero only when $\mathbf{p} = \mathbf{0}$" is satisfied when the power of all users are finitely constrained, i.e., no user's transmit power can be infinity. Notice that $\mathbf{T}$ is the self-mapping in (14) (not the utility functions in the original optimization problem). Hence, its concavity does not imply the concavity of the utility functions or the convexity of the original optimization problem in (1) or (9).

In the next subsection, we provide two efficient algorithms that can be used as an inner loop of Algorithm 1 to compute the scale $\beta(\mathbf{p}(t))$ at Step 2 in each iteration t.

C. Efficient Computation of $\beta(\mathbf{p})$

Recall that for given $\mathbf{p} > \mathbf{0}$, the scale $\beta(\mathbf{p})$ is chosen such that the normalized power vector $\tilde{\mathbf{p}} \triangleq \mathbf{p}/\beta(\mathbf{p})$ guarantees that at least one system constraint in (1) is tight. In general, this scale can be found via a bisection search, where the upper bound of $\beta(\mathbf{p})$ is first established in the initialization phase and the actual value of $\beta(\mathbf{p})$ is then obtained by decreasing the interval between the upper and lower bounds geometrically. This procedure is summarized in Algorithm 2. The advantage of the bisection search is that its worst-case convergence time can be computed explicitly for a given ϵ and initial values of L and U [29].

Even though Algorithm 2 is sufficient for computing $\beta(\mathbf{p})$, a more efficient algorithm can also be proposed by identifying the relation between β and the standard interference function defined in [30]. Moreover, it is interesting to show that, for $\mathbf{p} > \mathbf{0}$, finding $\beta(\mathbf{p})$ by (12) is equivalent to choosing the minimum value of $\beta' > 0$ such that

$$I_k(\beta') \triangleq \beta' g_k(\mathbf{p}/\beta')/\bar{g}_k \leqq \beta', \quad k = 1, \ldots, K. \tag{18}$$

Based on this observation, an efficient iterative algorithm can be proposed for the computation of $\beta(\mathbf{p})$ by generalizing the results from [30]. Following [30], let us define the following class of standard functions.

Definition 4 (Standard Interference Function [30]): A function $I : \mathcal{R}_+ \rightarrow \mathcal{R}_+$ is *standard* if the following

Algorithm 3 Computation of β via Fixed-Point Iteration

1) Set initial value $\beta(0) > 0$.
2) Set $\beta(t+1) \leftarrow \max_k I_k(\beta(t))$.
3) Repeat Step 2 until convergence.

conditions are satisfied for all $\beta > 0$[1]:

1) Monotonicity: If $\beta^{(a)} \geqq \beta^{(b)}$, then $I(\beta^{(a)}) \geqq I(\beta^{(b)})$.
2) Scalability: For $\lambda > 1$, $\lambda I(\beta) > I(\lambda\beta)$.

For $\{I_k\}_{k=1}^K$ that are standard, Algorithm 3 can instead be used for the computation of $\beta(\mathbf{p})$. With a slight abuse of notation, we denote by $\beta(t)$ the value of β in the t-th iteration of the algorithm.

The following theorem shows the convergence and the validity of the above algorithm.

Theorem 4: For $\mathbf{p} > 0$ and for $\{g_k\}_{k=1}^K$ such that $I_k(\beta) \triangleq \beta g_k(\mathbf{p}/\beta)$ is standard, $\forall k$, the following properties hold:

1) Algorithm 3 converges to a unique fixed point β^*.
2) The fixed point β^* is equal to $\beta(\mathbf{p})$ defined in (12).

The above theorem is proved following similar procedures as in [30] and is omitted for brevity. However, it is worthwhile to note that, in [30], the properties in Definition 4 are required to hold for $\beta = 0$ as well. However, by initializing Algorithm 3 with $\beta(0) > 0$, the value of $\beta(t)$ will always be positive for all t and, thus, the convergence proof in [30] holds for our case as well. In the following, we present a sufficient condition on g_k for efficient evaluation of whether I_k is standard.

Corollary 2: The properties in Theorem 4 hold for $\mathbf{p} > \mathbf{0}$ and for $\{g_k\}_{k=1}^K$ that are concave and satisfy Assumption 2.

The proof relies on showing that I_k is standard if g_k is concave and monotone. Details can be found in Appendix D. Note that Algorithm 2 is sufficient to compute $\beta(\mathbf{p}(t+1))$ at Step 3 of Algorithm 1, but Algorithm 3 provides a more efficient alternative when Theorem 4 or Corollary 2 hold.

IV. Case Studies for the Optimization Framework

In this section, we utilize the mathematical tools developed in the previous section to solve representative problems in five different wireless networking applications, namely, the max-min quality-of-service subject to robust interference temperature constraints in cognitive radio networks, the min-max weighted MSE subject to SINR constraints in multiuser downlink systems, the max-min throughput subject to nonlinear power constraints in energy-efficient wireless networks, the max-min sigmoid utility subject to delay and weighted power constraints in multimedia wireless networks, and the min-max outage probability subject to outage probability constraints in heterogeneous networks. The optimal solutions can be computed using Algorithm 1.

A. Case Study I: Cognitive Radio Networks With Deterministic and Stochastic Interference Temperature Constraints

Let us consider a cognitive radio (CR) network with M primary receivers (PRs) and L secondary transmitter-receiver pairs. Here, we consider a representative example where the power vector $\mathbf{p} = [p_1, \ldots, p_L]^T$ associated with the L secondary transmitters (STs) is adjusted to maximize the minimum SINR among secondary receivers (SRs) subject to interference temperature constraints at PRs and individual power constraints at STs. This problem was previously investigated in [15] and [31] by considering only deterministic interference temperature constraints, which requires perfect channel state information (CSI) at all users. However, in practice, perfect CSI is difficult to obtain between STs and PRs due to channel estimation error or limited feedback, and often only their statistics can be acquired. In this case, interference temperature constraints can only be satisfied in a probabilistic sense, resulting in stochastic interference temperature constraints that often cannot be expressed in closed-form. These constraints were addressed in [26] by using approximate closed-form expressions and by relaxing the optimization problem into a sequence of GP problems. With the tool developed in the previous section, we show that the above problem can be solved exactly (without approximation or convex relaxation).

Specifically, let us consider the case where the instantaneous channels between all ST-SR pairs and those between STs and a set of M' ($\leqq M$) PRs, labelled $1, \ldots, M'$, are perfectly known, but the channels between STs and the other PRs, labelled $M'+1, \ldots, M$, are unknown.[2] The SINR at SR i is given by

$$\mathrm{SINR}_i(\mathbf{p}) = \frac{G_{ii} R_{ii} p_i}{\sum_{j=1, j\neq i}^{L} G_{ij} R_{ij} p_j + \eta_i}, \tag{19}$$

where G_{ij} and R_{ij}, respectively, are the long-term path gain and short-term fading gain between ST j and SR i, and η_i is the effective noise variance at SR i (which includes the interference from primary users). The interference temperature at PR m is given by

$$q_m(\mathbf{p}) = \sum_{i=1}^{L} G_{mi}^{\mathrm{PU}} R_{mi}^{\mathrm{PU}} p_i, \tag{20}$$

where G_{mi}^{PU} and R_{mi}^{PU} are also the long-term and short-term channel gains between ST i and PR m. Interference temperature can be constrained deterministically with perfect CSI, but can only be constrained stochastically with imperfect CSI. The representative problem that we consider can thus be formulated as follows:

$$\text{maximize} \quad \min_{i=1,\ldots,L} \; \mathrm{SINR}_i(\mathbf{p}) \tag{21a}$$

$$\text{subject to} \; q_m(\mathbf{p}) \leq t_m, \quad m = 1, \ldots, M', \tag{21b}$$

$$\Pr(q_m(\mathbf{p}) \geq t_m) \leq \varepsilon_m, \quad m = M'+1, \ldots, M, \tag{21c}$$

[1]Notice that, even though $\beta \geq 0$ was required in [30], the results hold equally for just $\beta > 0$. Moreover, notice that positivity (i.e., $I(\beta) > 0$) was required explicitly in [30], but can actually be implied by monotonicity and scalability since the latter two yield $tI(\beta) > I(t\beta) \geq I(\beta)$, for $t > 1$.

[2]The case where some ST to PR channels are known and some are not is considered only to demonstrate that our technique is applicable in both cases. In practice, the former case may occur when the primary and secondary systems are cooperative whereas the latter case occurs when they are not.

$$p_i \leq \bar{p}_i, \quad i = 1, \ldots, L, \tag{21d}$$

$$\text{variables}: \mathbf{p}, \tag{21e}$$

where (21b) represents the deterministic interference temperature constraints, (21c) represents the stochastic interference temperature constraints, and (21d) represents the individual power constraints. It is easy to verify that all constraints satisfy Assumption 2.

To justify the use of Algorithm 1, it is necessary to verify that the utility functions

$$u_i(\mathbf{p}) = \mathsf{SINR}_i(\mathbf{p}) = \frac{G_{ii} R_{ii} p_i}{\sum_{j=1, j\neq i}^{L} G_{ij} R_{ij} p_j + \eta_i}, \; i = 1, \ldots, L \tag{22}$$

satisfy Assumption 1 and the conditions in Theorem 3 (or Corollary 1). The former is straightforward to show and, thus, we focus only on the latter. To do so, let us first define $\mathbf{F}$ as the $L \times L$ nonnegative matrix with entries

$$F_{ij} = \begin{cases} 0, & \text{for } i = j, \\ \frac{G_{ij} R_{ij}}{G_{ii} R_{ii}}, & \text{for } i \neq j, \end{cases} \tag{23}$$

and define the vector $\mathbf{v}$ as

$$\mathbf{v} = \left(\frac{\eta_1}{G_{11} R_{11}}, \ldots, \frac{\eta_L}{G_{LL} R_{LL}} \right)^T. \tag{24}$$

Then, we can express the objective functions as

$$u_i(\mathbf{p}) = \mathsf{SINR}_i(\mathbf{p}) = \frac{p_i}{(\mathbf{F}\mathbf{p} + \mathbf{v})_i}, \; i = 1, \ldots, L. \tag{25}$$

Notice that $\mathbf{T}(\mathbf{p}) = [T_1(\mathbf{p}), \ldots, T_L(\mathbf{p})]^T = \mathbf{F}\mathbf{p} + \mathbf{v}$ is affine and, thus, is concave. Moreover, since $\mathbf{F}$ is a nonnegative matrix and $\mathbf{v}$ has elements that are strictly positive, $\mathbf{T}$ is a positive mapping, i.e., $\mathbf{T}(\mathbf{p}) > \mathbf{0}$, for all $\mathbf{p} \geq \mathbf{0}$. Hence, the conditions in Corollary 1 are satisfied and, thus, optimal power control can be performed iteratively with guaranteed convergence by Algorithm 1.

Moreover, the value of $\beta(\mathbf{p}(t))$ in each iteration t can be computed using the bisection search described in Algorithm 2. Since the stochastic interference temperature constraints in general cannot be obtained in closed form, most works that deal with such probabilistic constraints resort to approximation methods, as in [26]. The resulting solution is thus suboptimal. However, the need for closed-form expressions of the stochastic interference temperature constraints can be avoided in our scheme. This is because, in Algorithm 1, the constraints need only to be considered in Step 3 to determine the scaling of $\mathbf{p}(t+1)$, i.e., $\beta(\mathbf{p}(t+1))$, and by adopting the bisection search (i.e., Algorithm 2) to compute $\beta(\mathbf{p}(t+1))$, we only need to determine in each iteration of Algorithm 2 whether or not the stochastic constraints are satisfied for given $\mathbf{p}(t+1)$ and $\beta = (U + L)/2$. The latter can be done by replacing the probabilities with their sample average approximations. By eliminating the subinterval in which the solution does not lie, we obtain a smaller search interval for the next iteration of the bisection search. The accuracy of the solution can be made arbitrarily small by increasing the sample size, which is not the case for the approximation considered in [26].

B. Case Study II: Reliable Multiuser Downlink System With SINR Requirement Constraints

In this subsection, we consider a single cell multiuser downlink system, where the base station (the transmitter) is equipped with an antenna array and each user (the receiver) has a single receive antenna, receiving data from the corresponding transmitter simultaneously on a shared spectrum. Assume that L users operate over a common frequency-flat channel. Channel information is available at both the transmitter and the receivers. The accuracy and computational complexity of decoding depends on the MSE [13], [32], thereby the optimization of MSE is required for reliable transmission. In addition, as SINR requirements are imposed on all users, we seek to minimize the maximum weighted MSE between the transmitted and estimated symbols subject to SINR constraints.

Let $\mathbf{p} = [p_1, \ldots, p_L]^\top$ be the transmit power vector. The SINR of user i is given by (25) with the nonnegative matrix $\mathbf{F}$ and the vector $\mathbf{v}$ defined by (23) and (24) respectively. Suppose that the linear minimum mean-square error (LMMSE) estimator is used to estimate the received symbol at each user. In this case, we can express the MSE of the i-th user as

$$\mathsf{MSE}_i(\mathbf{p}) = \frac{1}{1 + \mathsf{SINR}_i(\mathbf{p})}. \tag{26}$$

Our objective is to minimize the maximum weighted MSE, i.e., $\min_{\mathbf{p}} \max_{i=1,\ldots,L} w_i \mathsf{MSE}_i(\mathbf{p})$, where $\mathbf{w}$ is a positive vector with the entry w_i assigned to the i-th link to reflect its priority. A larger w_i denotes a higher priority. Notice that solving $\min_{\mathbf{p}} \max_{i=1,\ldots,L} w_i \mathsf{MSE}_i(\mathbf{p})$ is equivalent to solving $\max_{\mathbf{p}} \min_{i=1,\ldots,L} u_i(\mathbf{p})$, where

$$u_i(\mathbf{p}) = \frac{1 + \mathsf{SINR}_i(\mathbf{p})}{w_i} = \frac{(\mathbf{F}\mathbf{p} + \mathbf{v})_i + p_i}{w_i (\mathbf{F}\mathbf{p} + \mathbf{v})_i}. \tag{27}$$

Hence, the problem can be formulated as

$$\text{maximize} \; \min_{i=1,\ldots,L} \; u_i(\mathbf{p}) \tag{28a}$$

$$\text{subject to } \mathsf{SINR}_i(\mathbf{p}) \geq \gamma_i, \quad i = 1, \ldots, L \tag{28b}$$

$$\text{variables}: \mathbf{p}, \tag{28c}$$

where γ_i is the SINR threshold of user i, i.e., the received SINR at user i must be higher than γ_i. Let $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_L]^\top$ be the SINR threshold vector. In this case, the SINR constraints can be shown to satisfy Assumption 2 by rewriting the constraints in the equivalent matrix form given by $(\mathbf{I} - \mathrm{diag}(\boldsymbol{\gamma})\mathbf{F})\mathbf{p} \geq \mathrm{diag}(\boldsymbol{\gamma})\mathbf{v}$. Notably, $g_i(\mathbf{p}) = (\mathrm{diag}(\boldsymbol{\gamma})\mathbf{F} - \mathbf{I})\mathbf{p}$ is affine, and thus, is concave. Therefore, Corollary 2 holds and Algorithm 3 can be adopted to compute $\beta(\mathbf{p})$.

To show that $\{u_i\}_{i=1}^L$ satisfies Assumption 1, we first observe that $u_i(\mathbf{p})$ is positive and monotonically increases with respect to p_i and that $\partial u_i(\mathbf{p})/\partial p_j = -w_i^{-1} p_i F_{ij} (\mathbf{F}\mathbf{p} + \mathbf{v})^{-2} < 0$, for $j \neq i$. Therefore, the competitiveness holds. Moreover, for $\lambda > 1$, we have $u_i(\lambda\mathbf{p}) = \frac{1}{w_i}\big(1 + \frac{p_i}{(\mathbf{F}\mathbf{p} + \frac{1}{\lambda}\mathbf{v})_i}\big) \geqq u_i(\mathbf{p})$, $\forall i$, which implies directional monotonicity.

Next, we verify that $\{u_i\}_{i=1}^L$ satisfies the conditions in Theorem 3 (or equivalently, Corollary 1). Let us define $\mathbf{T}(\mathbf{p}) = [T_1(\mathbf{p}), \ldots, T_L(\mathbf{p})]^\top$, where

$$T_i(\mathbf{p}) = \frac{w_i p_i}{1 + \mathsf{SINR}_i(\mathbf{p})} = w_i (\mathbf{F}\mathbf{p} + \mathbf{v})_i \frac{\mathsf{SINR}_i(\mathbf{p})}{1 + \mathsf{SINR}_i(\mathbf{p})}. \quad (29)$$

It is easy to see that $\mathbf{T}$ is positive. Moreover, to show that $\mathbf{T}$ is concave in $\mathbf{p}$, let us invoke the following lemma from [12].

Lemma 4 [12]: Let $\mathbf{A} \in \mathcal{R}^{m \times n}$, $\mathbf{b} \in \mathcal{R}^m$, $\mathbf{c} \in \mathcal{R}^n$, and $d \in \mathcal{R}$. If $h : \mathcal{R}^m \to \mathcal{R}$ is concave, then

$$g(\mathbf{x}) = (\mathbf{c}^\top \mathbf{x} + d) h((\mathbf{A}\mathbf{x} + \mathbf{b})/(\mathbf{c}^\top \mathbf{x} + d)) \quad (30)$$

is concave on $\{\mathbf{x} | \mathbf{c}^\top \mathbf{x} + d > 0, (\mathbf{A}\mathbf{x} + \mathbf{b})/(\mathbf{c}^\top \mathbf{x} + d) \in \mathbf{dom}(h)\}$, where $\mathbf{dom}(h)$ denotes the domain of function h.

By considering $h(s) = \frac{s}{1+s}$, which is concave, and by letting $m = 1$, $\mathbf{A} = \mathbf{e}_i^\top$, $\mathbf{b} = 0$, $\mathbf{c}^\top$ be the i-th row of $\mathbf{F}$, $\mathbf{x} = \mathbf{p}$, and $d = \eta_i / G_{ii}$, it follows from Lemma 4 that T_i (and, thus, $\mathbf{T}$) is concave. Hence, it follows from Corollary 1 that the optimal solution to this problem can be obtained iteratively with guaranteed convergence using Algorithm 1.

C. Case Study III: Energy-Efficient Wireless Networks With Nonlinear Power Constraints

Most works in the literature on wireless communications make the idealized assumption that the transmit power is linear to the actual power consumption in each user's device. However, this is not the case in practice due to circuit nonlinearities. For example, under the Rapp model [28], the input and output power relation of a nonlinear power amplifier is written as

$$p_i = \alpha p_{i,\text{in}} \left[1 + \left(\frac{\alpha p_{i,\text{in}}}{b_{\text{sat}}} \right)^q \right]^{-1/q}, \quad (31)$$

where p_i is the output power, $p_{i,\text{in}}$ is the input power, α is the linear gain of the power amplifier, and b_{sat} is the saturation power as $p_{i,\text{in}} \to \infty$. Such nonlinearities may have a significant impact on the system performance [27]. In addition, the cost of emitting a certain amount of power may also be nonlinear due to, e.g., incremental pricing on the energy usage.

Let us consider a general wireless network with L transmit-receive pairs and aim to maximize (through adjustment of $\mathbf{p}$) the minimum throughput among the L transmission links subject to nonlinear constraints on power and cost. Using previous developments of the nonlinear Perron-Frobenius theory, a similar max-min throughput optimization problem was examined in [12] under linear power constraints. However, the tools developed in [12] are not sufficient to address the nonlinear power and cost constraints considered here.

Specifically, by assuming that the input-output power relation can be modeled by the strictly monotonic function h_i, i.e., $p_i = h_i(p_{i,\text{in}})$ (e.g., (31)), the power constraints at the transmitters can be generally expressed as

$$\mathbf{A}\tilde{\mathbf{h}}(\mathbf{p}) \le \bar{\mathbf{p}}, \quad (32)$$

where $\tilde{\mathbf{h}}(\mathbf{p}) = [h_1^{-1}(p_1), \ldots, h_L^{-1}(p_L)]^T$ is the vector of the actual powers that are consumed, and $\bar{\mathbf{p}}$ is a $K \times 1$ positive vector of constraint values. Here, $\mathbf{A}$ is a $K \times L$ nonnegative weight matrix, where no column or row is identically zero.

Moreover, the power consumed by each user, say user i, may also be subject to a (nonlinear) strictly monotonic cost function $c_i : \mathcal{R}_+ \to \mathcal{R}_+$. The transmission may thus also be subject to a linear constraint on the cost, i.e.,

$$\mathbf{B}\mathbf{c}(\mathbf{p}) \le \bar{\mathbf{c}}, \quad (33)$$

where $\mathbf{c}(\mathbf{p}) = [c_1(p_1), \ldots, c_L(p_L)]^T$, $\bar{\mathbf{c}}$ is an $S \times 1$ positive vector of constraint values, and $\mathbf{B}$ is an $S \times L$ nonnegative matrix with no row or column that is identically zero.

Let us consider the representative example where our goal is to maximize the minimum throughput among L users subject to the nonlinear power and cost constraints given above. By the Shannon capacity formula, we model the throughput of user i as $u_i(\mathbf{p}) = \log(1 + \mathsf{SINR}_i(\mathbf{p}))$, where $\mathsf{SINR}_i(\mathbf{p})$ is defined as in (2). Similar to (25), we can also write u_i as

$$u_i(\mathbf{p}) = \log\left(1 + \frac{p_i}{(\mathbf{F}\mathbf{p} + \mathbf{v})_i}\right). \quad (34)$$

The problem can thus be formulated as

$$\begin{aligned} &\text{maximize} \ \min_{i=1,\ldots,L} \ u_i(\mathbf{p}) && (35\text{a}) \\ &\text{subject to } \mathbf{A}\tilde{\mathbf{h}}(\mathbf{p}) \le \bar{\mathbf{p}} && (35\text{b}) \\ &\qquad\qquad \mathbf{B}\mathbf{c}(\mathbf{p}) \le \bar{\mathbf{c}} && (35\text{c}) \\ &\text{variables : } \mathbf{p}. && (35\text{d}) \end{aligned}$$

By properties of $\mathbf{A}$, $\mathbf{B}$, $\tilde{\mathbf{h}}$ and $\mathbf{c}$, strict monotonicity of the constraints is trivially satisfied; and, through appropriate choices of $\bar{\mathbf{p}}$ and $\bar{\mathbf{c}}$, feasibility and validity can be guaranteed as well. Therefore, Assumption 2 is satisfied. Furthermore, since $u_i(\mathbf{p})$ is a monotonic function of $\mathsf{SINR}_i(\mathbf{p})$, it inherits the competitiveness and directional monotonicity from $\mathsf{SINR}_i(\mathbf{p})$. Also, by definition of u_i, we have $u_i(\mathbf{p}) = 0$ when $p_i = 0$. Hence, Assumption 1 is satisfied as well.

Next, it is necessary to verify that $\{u_i\}_{i=1}^L$ satisfy the conditions in Theorem 3 or Corollary 1. Let us define

$$T_i(\mathbf{p}) = \frac{p_i}{\log(1 + \mathsf{SINR}_i(\mathbf{p}))}. \quad (36)$$

By taking the derivative of $T_i(\mathbf{p})$ with respect to p_j, we have

$$\frac{\partial T_i(\mathbf{p})}{\partial p_j} = \begin{cases} \frac{1}{(\log(x))^2} \left[\log(x) - (1 - \frac{1}{x})\right], & \text{for } i = j, \\ \frac{1}{(\log(x))^2} \frac{1}{x} (x-1)^2 \frac{G_{ij} R_{ij}}{G_{ii} R_{ii}}, & \text{otherwise}, \end{cases} \quad (37)$$

where $x = 1 + \mathsf{SINR}_i(\mathbf{p}) > 1$ for all $p_i > 0$. We can see that $\partial T_i(\mathbf{p})/\partial p_j > 0$, for all $p_i > 0$, when $j \neq i$. To show that $\partial T_i(\mathbf{p})/\partial p_i > 0$, let us consider the function $J(x) \triangleq \log(x) - (1 - \frac{1}{x})$. Notice that $J(1) = 0$ and that, for $x > 1$, its first-order derivative is $\partial J(x)/\partial x = \frac{1}{x^2}(x - 1) > 0$. This indicates that $J(x)$ is positive for all $x > 1$ and, thus, $\partial T_i(\mathbf{p})/\partial p_i > 0$ for $p_i > 0$. Therefore, all entries of $\nabla T_i(\mathbf{p})$ are positive for $p_i > 0$ and, thus, $T_i(\mathbf{p})$ increases monotonically in $\mathbf{p}$. Since $\mathbf{T}(\mathbf{0}) = \mathbf{0}$, it follows that $\mathbf{T}$ is positive for $\mathbf{p} \ge \mathbf{0}$.

Similar to the proof of concavity in Section IV-B, by taking $h(s) = \frac{s}{\log(1+s)}$, $m = 1$, $\mathbf{A} = \mathbf{e}_i^T$, $\mathbf{b} = 0$, $\mathbf{c}^T$ as the i-th row of $\mathbf{F}$, $\mathbf{x} = \mathbf{p}$, and $d = \eta_i / G_{ii}$, it follows from Lemma 4 that T_i (and, thus, $\mathbf{T}$) is concave. Hence, it follows from Corollary 1 that this problem can be solved iteratively with guaranteed convergence by Algorithm 1.

D. Case Study IV: Delay-Constrained Multimedia Wireless Networks

With the rapid development of 4G and 5G wireless networks, a wide range of real-time multimedia applications is inevitable and generate mass traffic of video streaming. In this subsection, we solve the physical-layer optimization problem, which involves power allocation as well as rate and SINR computations, to offer opportunities for application needs. In particular, in multimedia wireless networks, the utility of real-time services is often modeled by the sigmoid function of data rates, as described in [33] and [34]. This provides a good measure of user experience, especially in video streaming applications. The utility function of the i-th user in this case can be expressed as

$$u_i(\mathbf{p}) = \frac{1}{1 + e^{-\log(1+\mathsf{SINR}_i(\mathbf{p}))}}, \tag{38}$$

where $\mathbf{p} = [p_1, \ldots, p_L]^\top$ is the transmit power vector and the SINR of user i is given by (25). In delay-sensitive applications, it may be necessary to constrain the average queuing delay in order to avoid packet losses. By assuming an M/M/1 queue, i.e., Poisson-distributed packet arrivals, user i's average queuing delay can be modeled as [33]

$$D_i(\mathbf{p}) = \frac{1}{\log(1 + \mathsf{SINR}_i(\mathbf{p})) - \lambda_i}, \tag{39}$$

where λ_i is the arrival rate for user i.

To guarantee fairness among users, we consider the maximization of the minimum sigmoid utility among users subject to delay and weighted power constraints:

$$\begin{aligned} &\text{maximize} \ \min_{i=1,\ldots,L} \ u_i(\mathbf{p}) && (40a)\\ &\text{subject to } D_i(\mathbf{p}) \le \bar{D}_i, \quad i = 1, \ldots, L, && (40b)\\ &\qquad \mathbf{a}^\top \mathbf{p} \le \bar{p}, && (40c)\\ &\text{variables} : \mathbf{p}, && (40d) \end{aligned}$$

where $\bar{D}_i$ is the maximum tolerable delay for the i-th user, $\bar{p}$ and $\mathbf{a} \in \mathcal{R}_+^L$ are respectively the upper bound and the positive weight vector for the weighted power constraint.

By matrix transformation, the delay constraint can be rewritten as an affine function in $\mathbf{p}$: $(\mathbf{I} - \mathrm{diag}(e^{\frac{1}{\mathbf{D}}+\boldsymbol{\lambda}} - \mathbf{1})\mathbf{F})\mathbf{p} \ge \mathrm{diag}(e^{\frac{1}{\mathbf{D}}+\boldsymbol{\lambda}} - \mathbf{1})\mathbf{v}$. Since the delay constraint and the weighted power constraint are in fact both affine, Assumption 2 is satisfied. Furthermore, similar to (34), the monotonicity of $u_i(\mathbf{p})$ with $\mathsf{SINR}_i(\mathbf{p})$ inherits the competitiveness and directional monotonicity from $\mathsf{SINR}_i(\mathbf{p})$. Also, positivity can be directly observed from (38). Hence, Assumption 1 is also satisfied.

Next, let us verify that $\{u_i\}_{i=1}^L$ satisfies the conditions in Theorem 3 (or equivalently, Corollary 1). We define $\mathbf{T}(\mathbf{p}) = [T_1(\mathbf{p}), \ldots, T_L(\mathbf{p})]^\top$ such that

$$T_i(\mathbf{p}) = p_i \left(1 + e^{-\log\left(1+\mathsf{SINR}_i(\mathbf{p})\right)}\right). \tag{41}$$

It is easy to see that $\mathbf{T}$ is positive. Again, by Lemma 4, we take $h(s) = s(1 + e^{-\log(1+s)})$, $m = 1$, $\mathbf{A} = \mathbf{e}_i^T$, $\mathbf{b} = 0$, $\mathbf{c}^T$ as the i-th row of $\mathbf{F}$, $\mathbf{x} = \mathbf{p}$, and $d = \eta_i / G_{ii}$. Since $h''(s) = -2(1+s)^{-2} e^{-\log(1+s)} \le 0$ indicates the concavity of h, $\mathbf{T}$ is also concave. Therefore, it follows from Corollary 1 that we can solve the problem optimally by Algorithm 1 with a geometric convergence rate.

E. Case Study V: Heterogeneous Wireless Cellular Networks With Outage Constraints

In this subsection, we consider the downlink of a heterogeneous cellular network with $L + M$ base stations transmitting to their corresponding users under Rayleigh fading. Among these base stations, we assume that L base stations can adjust transmission power to optimize the overall network utility whereas M base stations maintain a nonadjustable and fixed transmission power over time. Our goal is to minimize the worst outage probability among the L base stations whose powers are adjustable subject to constraints on the outage probability of the M base stations whose powers are nonadjustable (e.g., high priority macrocell base stations).

Let $\mathbf{p} = [p_1, \ldots, p_L]^\top$ be the transmit power vector of the L base stations with adjustable power and let $p_{L+1}, \ldots, p_{L+M}$ be the transmit powers of base stations with fixed nonadjustable powers. The transmissions are subject to co-channel interference and, thus, the SINR of the user served by base station i is given by (5). For simplicity, the noise variance is assumed to be the same for all users, i.e., $\eta_i = \eta_n$, $\forall i$. Then, the outage probability of the user served by base station i can be written in closed-from as [5], [11]:

$$\begin{aligned} O_i(\mathbf{p}) &= \Pr(\mathsf{SINR}_i(\mathbf{p}) < \gamma_i) && (42)\\ &= 1 - e^{\frac{-\gamma_i \eta_n}{G_{ii} p_i}} \prod_{j=1, j\neq i}^{L+M} \left(1 + \frac{\gamma_i G_{ij} p_j}{G_{ii} p_i}\right)^{-1}, && (43) \end{aligned}$$

where γ_i is the SINR threshold of the user served by base station i.

Notice that finding $\mathbf{p}$ that minimizes the worst outage probability, i.e., $\min_{\mathbf{p}} \max_i O_i(\mathbf{p})$, is equivalent to finding $\mathbf{p}$ for the max-min problem $\max_{\mathbf{p}} \min_i u_i(\mathbf{p})$ with

$$\begin{aligned} u_i(\mathbf{p}) &= [-\log(1 - O_i(\mathbf{p}))]^{-1} && (44)\\ &= \left[\frac{\gamma_i \eta_n}{G_{ii} p_i} + \sum_{j=1, j\neq i}^{L+M} \log\left(1 + \frac{\gamma_i G_{ij} p_j}{G_{ii} p_i}\right)\right]^{-1}. && (45) \end{aligned}$$

Hence, the problem can be formulated as

$$\begin{aligned} &\text{maximize} \ \min_{i=1,\ldots,L} \ u_i(\mathbf{p}) && (46a)\\ &\text{subject to } O_m(\mathbf{p}) \le \bar{O}_m, \quad m = L+1, \ldots, L+M, && (46b)\\ &\qquad p_i \le \bar{p}_i, \quad i = 1, \ldots, L, && (46c)\\ &\text{variables} : \mathbf{p}. && (46d) \end{aligned}$$

Again, it is straightforward to show that the above constraints satisfy Assumption 2. In particular, for the constraints in (46b), we can see that the outage probability $O_m(\mathbf{p})$ increases as the entries of $\mathbf{p}$ increase since user m's power is assumed to be fixed.

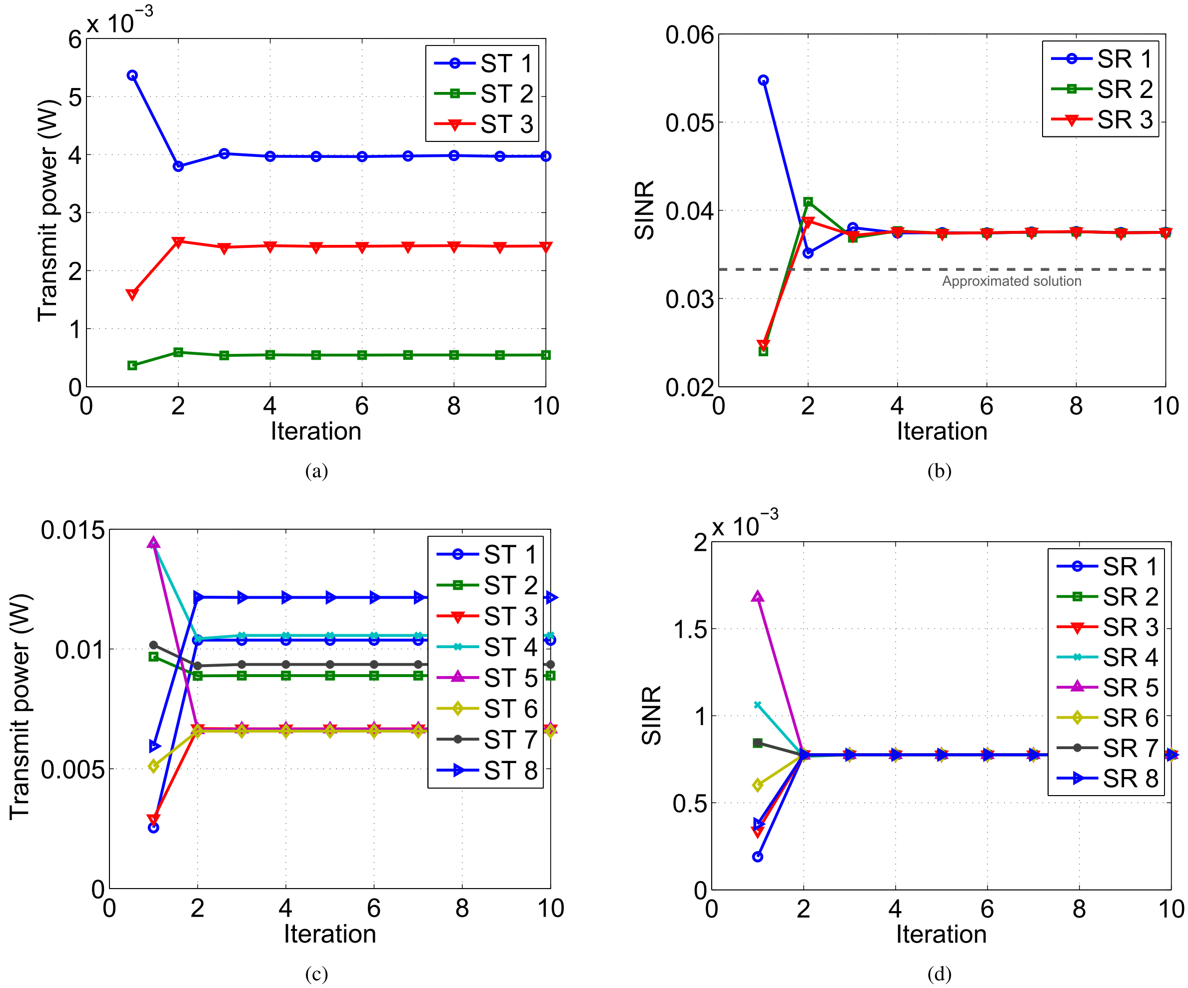


Fig. 1. In Experiment I, the transmit powers of STs and the SINR of the SRs both converge rapidly, with the latter reaching a common value for all SRs. (a) Evolution of the transmit powers at 3 STs. (b) Evolution of the SINRs at 3 SRs. (c) Evolution of the transmit powers at 8 out of the 100 STs. (d) Evolution of the SINRs at 8 out of the 100 SRs.

Next, we show that the utility functions $\{u_i\}_{i=1}^{L}$ considered here satisfy Assumption 1 and the conditions in Theorem 3 (or, more practically, those in Corollary 1). First, notice that positivity and competitiveness can be seen directly from the definition in (44). In particular, we can see that $u_i(\mathbf{p})$ is equal to 0 when $p_i = 0$ and that it increases monotonically with p_i but decreases monotonically with p_j, for all $j \neq i$, when $p_i > 0$. Moreover, for $\lambda \geqq 1$, we have

$$u_i(\lambda \mathbf{p}) = \left[\frac{\gamma_i \eta_n}{G_{ii} \lambda p_i} + \sum_{j=1, j\neq i}^{L+M} \log\left(1 + \frac{\gamma_i G_{ij} p_j}{G_{ii} p_i}\right)\right]^{-1} \geqq u_i(\mathbf{p}),$$

for all i, and, thus, directional monotonicity follows.

Next, let $\mathbf{T}(\mathbf{p}) = [T_1(\mathbf{p}), \ldots, T_L(\mathbf{p})]^\top$, where

$$T_i(\mathbf{p}) = \frac{p_i}{u_i(\mathbf{p})} = \frac{\gamma_i \eta_n}{G_{ii}} + \sum_{j=1, j\neq i}^{L} p_i \log\left(1 + \frac{\gamma_i G_{ij} p_j}{G_{ii} p_i}\right). \tag{47}$$

It is straightforward to see that $\mathbf{T}$ is positive. To show that $\mathbf{T}$ is concave, let us recall the fact that $t\log(1 + x/t)$ is strictly concave in (x, t) for strictly positive t since it is the perspective function of the strictly concave function $\log(1+x)$ [29]. Hence, $T_i(\mathbf{p})$ can be viewed as the sum of strictly concave perspective functions and, therefore, $\mathbf{T}$ is strictly concave in $\mathbf{p}$. Hence, by Corollary 1, the optimal solution can be obtained efficiently with guaranteed convergence using Algorithm 1.

It is interesting to remark that, in computing $\beta(\mathbf{p})$, the constraints in (46b) can be reformulated to allow for the adoption of Algorithm 3. That is, we can let

$$g_m(\mathbf{p}) \triangleq -\log\left[1 - O_m(\mathbf{p})\right] \tag{48}$$

$$= \frac{\gamma_m \eta_n}{G_{mm} p_m} + \sum_{j=1, j\neq m}^{L+M} \log\left(1 + \frac{\gamma_m G_{jm} p_j}{G_{mm} p_m}\right) \tag{49}$$

$$\leq -\log\left[1 - \bar{O}_m\right] \triangleq \bar{g}_m, \tag{50}$$

where p_m and p_j, for $j > L$, are constant. In this case, g_m is a concave and monotone function of $\mathbf{p}$. Hence, Corollary 2 holds and Algorithm 3 is applicable.

V. Numerical Examples

In this section, we provide numerical examples to demonstrate the effectiveness of Algorithm 1 for solving the max-min utility optimization problems described in Section IV. Moreover, Experiments I and II are also performed for networks with 100 users to demonstrate the scalability of our proposed

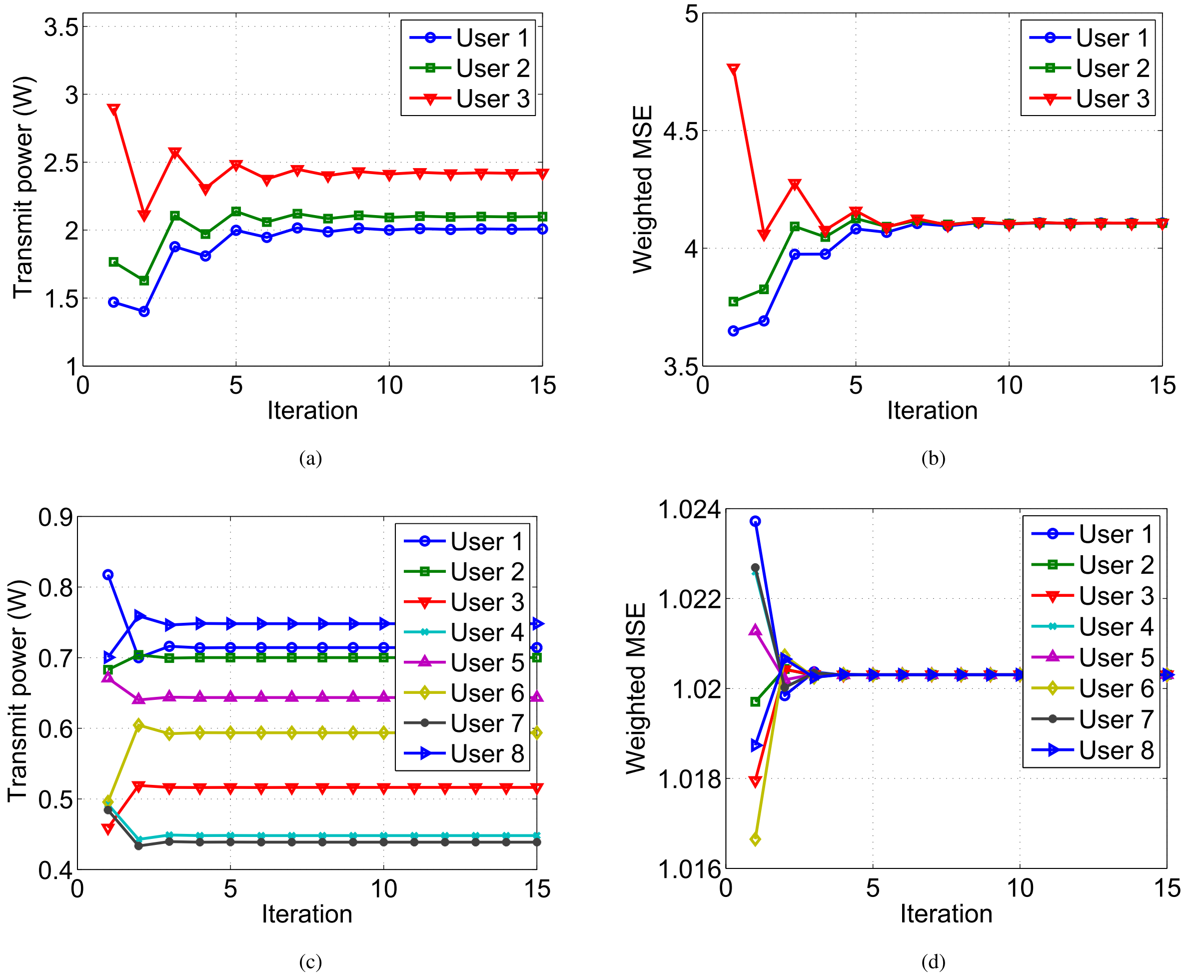


Fig. 2. In Experiment II, the transmit powers and the weighted MSE converge rapidly (within 9 iterations) using Algorithm 1. (a) Evolution of the transmit powers intended for the 3 users. (b) Evolution of the weighted MSE at 3 users. (c) Evolution of the transmit powers intended for 8 out of the 100 users. (d) Evolution of the weighted MSE at 8 out of the 100 users.

algorithms in a large network. Experiments III, IV and V can also scale, but are not shown for brevity. In large-scale experiments, all parameters are generated randomly. The initial transmit powers are generated randomly in all experiments.

A. Experiment I (Cognitive Radio Networks With Deterministic and Stochastic Interference Temperature Constraints)

Let us consider a cognitive radio network with $M = 2$ PRs and $L = 3$ ST-SR pairs. Following [26], the long-term path gain is given by $G_{ij} = d_{ij}^{-\nu} s_{ij}$, where d_{ij} and s_{ij} are the distance and the shadowing effect between ST j and SR i, respectively, and $\nu = 3.5$ is the path-loss exponent. s_{ij} is defined as $10^{X/10}$ with X being a Gaussian random variable with mean 0 and standard deviation 10 (in dB). Moreover, the short-term fading gain R_{ij} is assumed to be Nakagami-m with $m = 10$. Assume that the STs have instantaneous knowledge of their channels to PR 1, which are set as $[G_{11}^{\mathrm{PU}}, G_{21}^{\mathrm{PU}}, G_{31}^{\mathrm{PU}}]^T = [303^{-3.5} \cdot 0.02, 284^{-3.5} \cdot 143.45, 218^{-3.5} \cdot 3.04]^T$ and $[R_{11}^{\mathrm{PU}}, R_{21}^{\mathrm{PU}}, R_{31}^{\mathrm{PU}}]^T = [0.87, 1.14, 0.99]^T$, but the STs have only the distance information between themselves and PR 2, i.e., $[d_{11}^{\mathrm{PU}}, d_{21}^{\mathrm{PU}}, d_{31}^{\mathrm{PU}}]^T = [219, 209, 191]^T$, as well as the channel statistics. The instantaneous channels between STs and SRs are assumed to be known and are given by

$$\mathbf{G} = \begin{bmatrix} 71^{-3.5} \times 0.60 & 61^{-3.5} \times 0.74 & 253^{-3.5} \times 3.13 \\ 83^{-3.5} \times 44.98 & 65^{-3.5} \times 2.37 & 199^{-3.5} \times 0.22 \\ 222^{-3.5} \times 0.32 & 204^{-3.5} \times 0.39 & 84^{-3.5} \times 1.96 \end{bmatrix}$$

and

$$\mathbf{R} = \begin{bmatrix} 1.13 & 1.25 & 1.18 \\ 1.15 & 0.90 & 1.11 \\ 0.99 & 0.97 & 0.93 \end{bmatrix}.$$

The interference thresholds for both PRs are set to -80 dBW (i.e., $t_1 = t_2 = 10^{-8}$), and the interference temperature outage probability threshold for PR 2 is set to 0.01 (i.e., $\varepsilon_2 = 0.01$) [26]. The maximum transmit power for all STs is 5 W (i.e., $\bar{p}_1 = \bar{p}_2 = \bar{p}_3 = 5$).

In Fig. 1, we can see that both the ST transmit powers and the SINR at the SRs converge rapidly within 5 iterations using Algorithm 1. The interference temperature outage probability is evaluated in each iteration of the bisection search for $\beta(\mathbf{p})$ using sample average approximations (averaged over 10^7 realizations). The terminating parameter in Algorithm 2 is set as $\epsilon = 5 \cdot 10^{-5}$. The "approximated solution" indicated in

Fig. 1(b) represents the converged SINR value obtained with the approximate interference temperature outage probability expression adopted in [26], which yields only a suboptimal solution. In Figs. 1(c) and 1(d), we consider the case with $M = 60$ PRs and $L = 100$ ST-SR pairs. The instantaneous channels to all PRs are assumed to be known in this case. Here, we show the evolution of the transmit powers of 8 out of the 100 STs and the SINR at their corresponding SRs, demonstrating the scalability of Algorithm 1.

B. Experiment II (Reliable Multiuser Downlink System With SINR Requirement Constraints)

Let us consider a single-cell multiuser downlink system with a single base station serving $L = 3$ users, i.e., users 1, 2 and 3. The downlink transmit powers are p_1, p_2 and p_3. In the channel gain matrix $\mathbf{G} = [G_{ij}]_{i,j=1}^{L} > \mathbf{0}_{L\times L}$, the diagonal entries G_{ii}, for all i, represents the effective channel gain from the transmitter to user i and G_{ij}, for $j \neq i$, represents the effective cross-channel interference coefficient between the signal intended for user j and the reception at user i (i.e., the inner product between the channel vector at user i and the beamforming vector used for user j's signal). We let

$$\mathbf{G} = \begin{bmatrix} 0.65 & 0.22 & 0.14 \\ 0.15 & 0.78 & 0.26 \\ 0.19 & 0.13 & 0.72 \end{bmatrix}, \tag{51}$$

and consider the weights $\mathbf{w} = [0.42\ 0.45\ 0.5]^\top$ for the weighted MSE objective. The noise power of each user is 1 W, and the SINR threshold vector is $\boldsymbol{\gamma} = [0.52\ 0.77\ 0.65]^\top$.

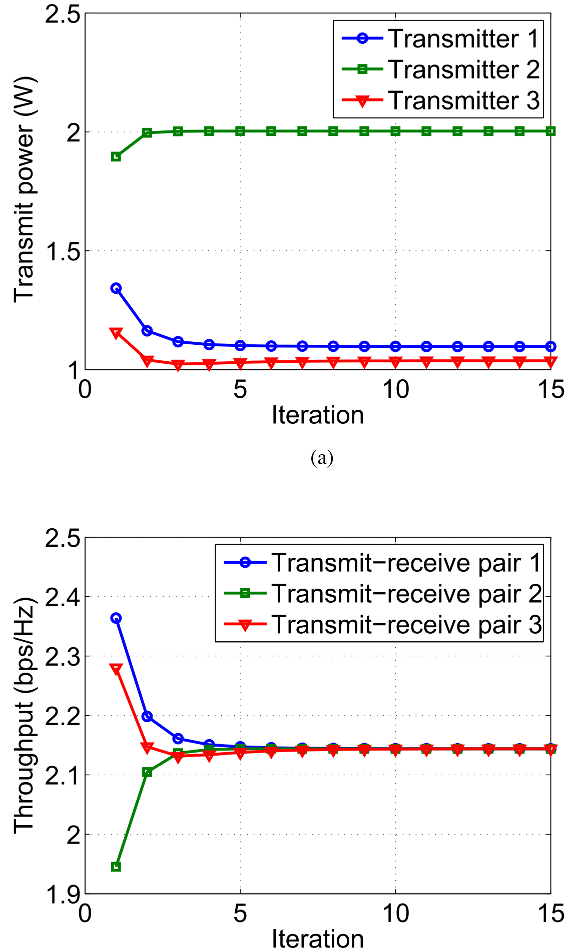
Figs. 2(a) and 2(b) plot the evolution of the base station transmit powers and the resulting weighted MSE at the 3 users. In this numerical example, $\beta(\mathbf{p}(t))$ is evaluated using Algorithm 3. We can observe from these figures that both the transmit powers and the MSEs converge very fast within 9 iterations to the optimal solution (verifying Theorem 3). Moreover, for a large network with $L = 100$ users, we show in Figs. 2(c) and 2(d) the evolution of the base station transmit powers for 8 of the 100 users as well as their weighted MSE. These figures again demonstrate the scalability of Algorithm 1.

C. Experiment III (Energy-Efficient Wireless Networks With Nonlinear Power Constraints)

Let there be $L = 3$ transmit-receive pairs and consider the power amplifier (PA SM1720-50) whose input-output power relation can be fitted into the Rapp model with parameters $b_{\text{sat}} = 10$ W, $\alpha = 316.2278$ and $q = 4$ [27]. Moreover, let us set the maximum input power as $\bar{p} = 0.01$ W, the noise variance as 1 pW, and the nonnegative weight matrix as

$$\mathbf{A} = \begin{bmatrix} 0.9134 & 0.2785 & 0.9649 \\ 0.6324 & 0.5469 & 0.1576 \\ 0.0975 & 0.9575 & 0.9706 \end{bmatrix}. \tag{52}$$

Let us consider the following nonlinear pricing: the cost is \$1/Wm when the power consumption is under 70 Watt-minute (Wm), \$5/Wm between 70 Wm and 110 Wm, and \$10/Wm when over 110 Wm. The transmit powers are



(a)

(b)

Fig. 3. In Experiment III, the transmit powers and the throughput values converge rapidly (within 10 iterations) using Algorithm 1. (a) Evolution of the transmit powers in a wireless network with 3 transmit-receive pairs. (b) Evolution of the throughput of 3 transmit-receive pairs in a wireless network.

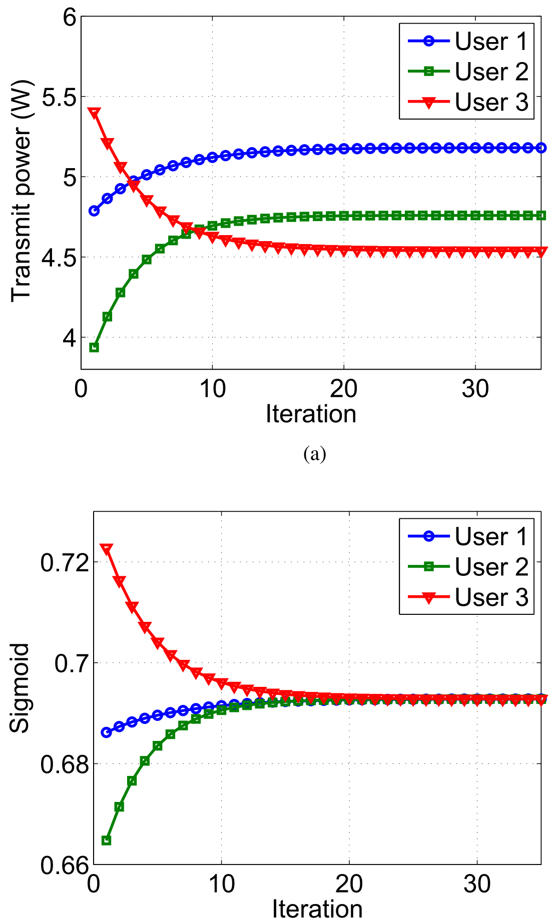
updated every minute. The total cost of the three transmit-receive pairs should be less than \$500 per minute, i.e., $\mathbf{B} = \mathbf{1}^T$ and $\bar{c} = 500$. The following channel gain matrix $\mathbf{G}$ is used:

$$\mathbf{G} = \begin{bmatrix} 80^{-4} & 103^{-4} & 208^{-4} \\ 227^{-4} & 98^{-4} & 94^{-4} \\ 305^{-4} & 207^{-4} & 23^{-4} \end{bmatrix}. \tag{53}$$

In Figs. 3(a) and 3(b), we again show that both the transmit powers and the throughput converge rapidly within 10 iterations using Algorithm 1. The value of $\beta(\mathbf{p}(t))$ is again computed using Algorithm 2 with $\epsilon = 5 \cdot 10^{-5}$.

D. Experiment IV (Delay-Constrained Multimedia Wireless Networks)

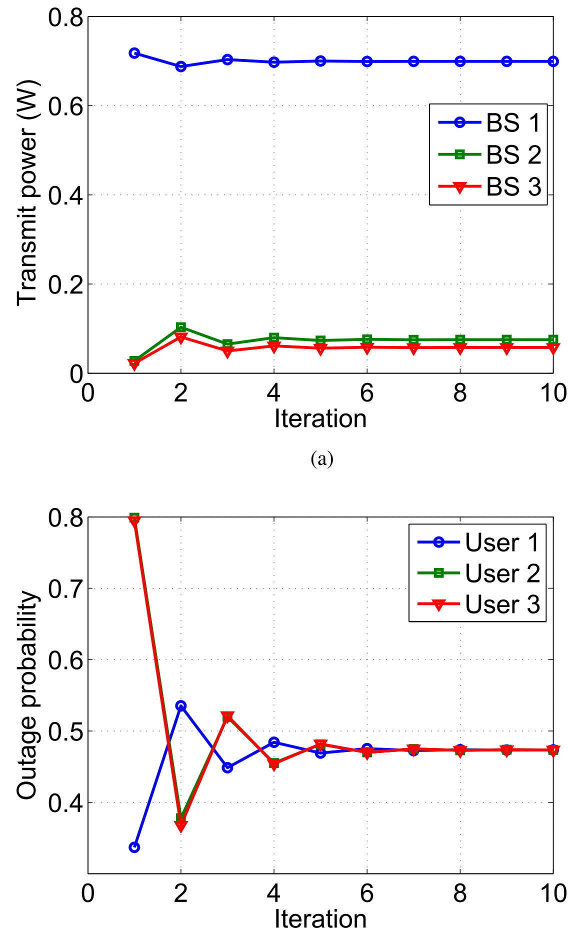
Let us consider a multimedia-based wireless network where the service provider serves $L = 3$ users. The transmit powers for their multimedia traffic are p_1, p_2 and p_3, respectively. We use the same channel gain matrix $\mathbf{G}$ as in Experiment II. We consider the maximum tolerable delay of $\bar{D}_i = 5$ seconds and the arrival rate $\lambda_i = 0.5$ for all users as well as



(a)

(b)

Fig. 4. In Experiment IV, the transmit powers and the sigmoid utility values both converge rapidly (within 40 iterations) using Algorithm 1. (a) Evolution of the transmit powers of 3 base stations. (b) Evolution of the outage probabilities of 3 users.



(a)

(b)

Fig. 5. In Experiment IV, the transmit powers and the outage probabilities both converge rapidly, with the latter reaching a common value for all users. (a) Evolution of the transmit powers of 3 base stations (BS). (b) Evolution of the outage probabilities of 3 users. (b) Evolution of the outage probabilities of 3 users.

$\mathbf{a} = [0.68, 0.56, 0.84]^\top$ and $\bar{p} = 10$ W for the weighted power. The noise power of each user is 1 W.

In Figs. 4(a) and 4(b), we plot the evolution of the transmit powers and the sigmoid functions. The transmit powers and the sigmoid utility functions are shown to converge rapidly within 40 iterations using Algorithm 1.

E. Experiment V (Heterogeneous Wireless Cellular Networks With Outage Constraints)

We consider three base stations (namely, base station 1, base station 2, and base station 3) whose transmit powers are adjustable and one base station (i.e., base station 4) whose transmit power is nonadjustable. Assume that one base station serves only one user and, thus, users 1, 2, 3, and 4 are served by base stations 1, 2, 3, and 4, respectively. For power-adjustable base stations, the maximum transmit power is 1 W (i.e., $\bar{p}_1 = \bar{p}_2 = \bar{p}_3 = 1$) and the SINR threshold is $\gamma_1 = \gamma_2 = \gamma_3 = 1$. For base station 4, whose transmit power is nonadjustable, the transmit power is fixed as 2 W (i.e., $p_4 = 2$). The SINR threshold of user 4 is set as $\gamma_4 = 10$ and the outage constraint is $\bar{O}_4 = 0.1$. The noise variance is given by $\eta_n = 1$ pW for all users. By considering only the path loss, the channel gain gain matrix is given by

$$\mathbf{G} = \begin{bmatrix} 93^{-4} & 124^{-4} & 152^{-4} & 127^{-4} \\ 67^{-4} & 55^{-4} & 215^{-4} & 191^{-4} \\ 119^{-4} & 213^{-4} & 58^{-4} & 188^{-4} \\ 169^{-4} & 164^{-4} & 179^{-4} & 53^{-4} \end{bmatrix}. \tag{54}$$

Fig. 5 plots the evolution of the base stations' transmit powers and the outage probabilities for all users in the heterogeneous network. The fast convergence (within 6 iterations) is observed for both the transmit powers and the outage probabilities. Here, $\beta(\mathbf{p}(t))$ is evaluated using Algorithm 3.

VI. CONCLUSION

In this paper, we proposed an optimization framework based on the generalization of nonlinear Perron-Frobenius theory to solve max-min utility fairness optimization problems with nonlinear monotonic constraints. The proposed techniques were shown to be widely applicable to many max-min utility fairness resource allocation problems with realistic monotonic constraints, such as nonlinear power constraints, interference constraints, and stochastic outage constraints, etc. In particular, algorithms with low complexity (no configuration needed) were designed to solve these nonconvex problems in a optimal and scalable manner. Five representative case studies in cognitive radio networks, multiuser downlink systems, energy-efficient wireless networks, multimedia-based wireless networks and heterogeneous networks were presented to illustrate the applicability of our framework. Extensive numerical examples verified our theoretical claims and demonstrated the optimality and fast-convergent property of our algorithms.

Appendix

A. Proof of Lemma 1

Suppose that $(\tau^*, \mathbf{p}^*)$ is an optimal solution. To show that $(\tau^*, \mathbf{p}^*)$ is positive, we first suppose that there exists i such that $p_i^* = 0$. Then, $u_i(\mathbf{p}^*) = 0$ and $\tau^* = \min_i u_i(\mathbf{p}^*) = 0$. However, by the feasibility assumption (c.f. Assumption 2), there must exist a vector $\mathbf{p}' > \mathbf{0}$ that is feasible and, by the positivity assumption (c.f. Assumption 1), we can find $\tau \triangleq \min_i u_i(\mathbf{p}') > 0 = \tau^*$, which contradicts the assumption that $(\tau^*, \mathbf{p}^*)$ is optimal. Hence, $(\tau^*, \mathbf{p}^*)$ is positive.

Next, we show that all of the utility constraints are tight at optimality. Suppose that there exists i such that $u_i(\mathbf{p}^*) > \tau^*$. Then, by the fact that $\mathbf{p}^* > \mathbf{0}$, as shown above, we can choose $\hat{\mathbf{p}}$ such that $0 < \hat{p}_i < p_i^*$ and $\hat{p}_j = p_j^*$, for all $j \neq i$, and such that $\tau^* < u_i(\hat{\mathbf{p}}) < u_i(\mathbf{p}^*)$, due to the competitiveness of the utility functions (c.f. Assumption 1). However, this yields $u_j(\hat{\mathbf{p}}) > u_j(\mathbf{p}^*)$, for all $j \neq i$. In this case, we can choose τ such that $\tau \triangleq \min_i u_i(\hat{\mathbf{p}}) > \min_i u_i(\mathbf{p}^*) = \tau^*$, which contradicts the assumption that τ^* is optimal. Hence, all of the objective constraints must be tight at optimality.

Finally, we show that at least one system constraint is tight at optimality. Suppose that $g_k(\mathbf{p}^*) < \bar{g}_k$, for all k. Since $\mathbf{p}^* > \mathbf{0}$, there exists $\lambda > 1$ and $\mathbf{p}' \triangleq \lambda\mathbf{p}^*$ such that $g_k(\mathbf{p}^*) < g_k(\mathbf{p}') \leq \bar{g}_k$, for all k. Then, by the fact that $\mathbf{p}^* > \mathbf{0}$ and the directional monotonicity of the utility functions (c.f. Assumption 1), it follows that $u_i(\mathbf{p}') > u_i(\mathbf{p}^*)$, for all i. This contradicts the assumption that $\mathbf{p}^*$ is optimal. Hence, at least one of the system constraints must be tight at optimality.

B. Proof of Lemma 2

The properties of β are shown as follows.

1) Suppose that statement 1) in Lemma 2 is not true. Then, there exists $\mathbf{p}' > \mathbf{0}$ such that $\beta(\mathbf{p}') = 0$. This implies that $g_k(\mathbf{p}'/\beta') < \bar{g}_k$ for all k and for all $\beta' > 0$. This contradicts the validity assumption in Assumption 2. Hence, it follows that $\beta(\mathbf{p}) > 0$, for all $\mathbf{p} > \mathbf{0}$.

2) The property follows trivially for $\lambda = 0$ since $\beta(0\mathbf{p}) = \beta(\mathbf{0}) = 0 = 0\beta(\mathbf{p})$. To prove the property for $\lambda > 0$, let us recall, from the definition of β, that $g_k(\mathbf{p}/\beta(\mathbf{p})) \leqq \bar{g}_k$, for all k. Thus, for $\lambda > 0$, it follows that $g_k(\lambda\mathbf{p}/\lambda\beta(\mathbf{p})) \leqq \bar{g}_k$, for all k. This implies that $\lambda\beta(\mathbf{p}) \geq \beta(\lambda\mathbf{p}) \triangleq \min\{\beta' : g_k(\lambda\mathbf{p}/\beta') \leqq \bar{g}_k, \forall k\}$. Suppose that $\lambda\beta(\mathbf{p}) > \beta(\lambda\mathbf{p})$. Then, we can choose $\hat{\beta} \triangleq \beta(\lambda\mathbf{p})/\lambda < \beta(\mathbf{p})$ such that $g_k(\mathbf{p}/\hat{\beta}) = g_k(\lambda\mathbf{p}/\beta(\lambda\mathbf{p})) \leqq \bar{g}_k$, for all k, which contradicts the definition that $\beta(\mathbf{p}) \triangleq \min\{\beta' : g_k(\mathbf{p}/\beta') \leqq \bar{g}_k, \forall k\}$. Hence, $\lambda\beta(\mathbf{p}) = \beta(\lambda\mathbf{p})$.

3) The property is trivially satisfied for $\beta(\mathbf{p}) = 0$. Now, let us consider the case where $\beta(\mathbf{p}) > 0$. By the monotonicity of g_k, we know that $g_k(\mathbf{p}) \leqq g_k(\mathbf{q})$, for any $\mathbf{p} \leqq \mathbf{q}$. Therefore, for $\beta' > 0$, it follows that $g_k(\mathbf{p}/\beta') \leqq g_k(\mathbf{q}/\beta')$, for any $\mathbf{p} \leqq \mathbf{q}$. Hence, $\{\beta' : g_k(\mathbf{p}/\beta') \leqq \bar{g}_k, \forall k\} \supseteq \{\beta' : g_k(\mathbf{q}/\beta') \leqq \bar{g}_k, \forall k\}$ and, thus, $\beta(\mathbf{p}) \triangleq \min\{\beta' : g_k(\mathbf{p}/\beta') \leqq \bar{g}_k, \forall k\} \leq \{\beta' : g_k(\mathbf{q}/\beta') \leqq \bar{g}_k, \forall k\} \triangleq \beta(\mathbf{q})$.

C. Proof of Lemma 3

Let $\mathbf{p}^*$ be an optimal solution of (1). Suppose that $\beta(\mathbf{p}^*) > 1$ and recall, from Lemma 1, that $\mathbf{p}^*$ must be nonzero in all components. It then follows by the strict monotonicity and validity of the constraints that $g_k(\mathbf{p}^*) > g_k(\mathbf{p}^*/\beta(\mathbf{p}^*)) = \bar{g}_k$ for some k. Therefore, $\mathbf{p}^*$ is infeasible and contradicts the assumption that $\mathbf{p}^*$ is optimal. Hence, $\beta(\mathbf{p}^*) \leqq 1$.

Moreover, by Lemma 1, we know that $g_k(\mathbf{p}^*) \leqq \bar{g}_k$, for all k, and there exists k' such that $g_{k'}(\mathbf{p}^*) = \bar{g}_{k'}$. Suppose that $\beta(\mathbf{p}^*) < 1$, then the power vector $\tilde{\mathbf{p}} = \mathbf{p}^*/\beta(\mathbf{p}^*)$ is also feasible. By the fact that $\mathbf{p}^* > \mathbf{0}$ (c.f. Lemma 1) and by the directional monotonicity of the utility functions, it follows that $u_i(\tilde{\mathbf{p}}) > u_i(\mathbf{p}^*)$, for all i. This implies that $\min_i u_i(\tilde{\mathbf{p}}) > \min_i u_i(\mathbf{p}^*)$, which contradicts the assumption that $\mathbf{p}^*$ is optimal. Hence, $\beta(\mathbf{p}^*) = 1$ and, thus, $\mathbf{p}^* \in \mathcal{U}$.

D. Proof of Corollary 2

This proof relies on showing that $I_k \triangleq \beta g_k(\mathbf{p}/\beta)$ is standard by Definition 4 for all k. In particular, to show the monotonicity of I_k, let us first take the derivative of $I_k(\beta)$ with respect to β, which yields

$$I_k'(\beta) = g_k(\mathbf{p}/\beta) - \frac{1}{\beta}\mathbf{p}^T \nabla_{\mathbf{p}} g_k(\mathbf{p}/\beta), \tag{55}$$

where $\nabla_{\mathbf{p}}$ represents the gradient operator with respect to $\mathbf{p}$. By the change of variable $x = 1/\beta$, we have

$$f(x) \triangleq I_k'(1/x) = g_k(x\mathbf{p}) - x\mathbf{p}^T \nabla_{\mathbf{p}} g_k(x\mathbf{p}). \tag{56}$$

Notice that $f(0) = g_k(0) \geqq 0$ and $f'(x) = -x\mathbf{p}^T \nabla^2_{\mathbf{p}} g_k(x\mathbf{p})\mathbf{p} \geqq 0$, $\forall x > 0$, due to the concavity of g_k, where $\nabla^2_{\mathbf{p}}$ represents the Hessian operator with respect to $\mathbf{p}$. Therefore, $f(x) = I_k'(1/x) \geqq 0$, $\forall x \geqq 0$ and, thus, $I'(\beta) \geqq 0$, $\forall \beta > 0$. Hence, we can conclude that I_k is nondecreasing for $\beta > 0$, i.e., monotonicity holds for I_k. Moreover, for $\lambda > 1$ and $\mathbf{p} > \mathbf{0}$, we have $\lambda I_k(\beta) = \lambda\beta g_k(\mathbf{p}/\beta) > \lambda\beta g_k(\mathbf{p}/(\lambda\beta)) = I_k(\lambda\beta)$ since $\mathbf{p}/\beta > \mathbf{p}/(\lambda\beta)$. Hence, scalability holds for I_k. Since I_k satisfies monotonicity and scalability, it is standard by Definition 4 and, thus, properties of Theorem 4 hold for this case as well.

References

[1] Y. W. Hong, C. W. Tan, L. Zheng, C.-L. Hsieh, and C.-H. Lee, "A unified framework for wireless max-min utility optimization with general monotonic constraints," in *Proc. IEEE INFOCOM*, Apr. 2014, pp. 2076–2084.

[2] J. M. Aein, "Power balancing in systems employing frequency reuse," *COMSAT Tech. Rev.*, vol. 3, no. 2, pp. 277–299, 1973.

[3] X. Lin, N. Shroff, and R. Srikant, "A tutorial on cross-layer optimization in wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1452–1463, Aug. 2006.

[4] M. Chiang, "Balancing transport and physical layers in wireless multi-hop networks: Jointly optimal congestion control and power control," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 1, pp. 104–116, Jan. 2005.
[5] S. Kandukuri and S. Boyd, "Optimal power control in interference-limited fading wireless channels with outage-probability specifications," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 46–55, Jan. 2002.
[6] S. C. Borst, M. G. Markakis, and I. Saniee, "Nonconcave utility maximization in locally coupled systems, with applications to wireless and wireline networks," *IEEE/ACM Trans. Netw.*, vol. 22, no. 2, pp. 674–687, Apr. 2014.
[7] P. Hande, S. Rangan, M. Chiang, and X. Wu, "Distributed uplink power control for optimal SIR assignment in cellular data networks," *IEEE/ACM Trans. Netw.*, vol. 16, no. 6, pp. 1420–1433, Dec. 2008.
[8] J.-W. Lee, R. R. Mazumdar, and N. B. Shroff, "Joint opportunistic power scheduling and end-to-end rate control for wireless ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 801–809, Mar. 2007.
[9] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 41, no. 1, pp. 57–62, Feb. 1992.
[10] M. Chiang, C. W. Tan, D. P. Palomar, D. O. Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.
[11] C. W. Tan, "Optimal power control in rayleigh-fading heterogeneous wireless networks," *IEEE/ACM Trans. Netw.*, vol. 24, no. 2, pp. 940–953, Apr. 2016.
[12] D. W. H. Cai, C. W. Tan, and S. H. Low, "Optimal max-min fairness rate control in wireless networks: Perron-frobenius characterization and algorithms," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 648–656.
[13] C. W. Tan, M. Chiang, and R. Srikant, "Maximizing sum rate and minimizing MSE on multiuser downlink: Optimality, fast algorithms and equivalence via max-min SINR," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6127–6143, Dec. 2011.
[14] C. W. Tan, M. Chiang, and R. Srikant, "Fast algorithms and performance bounds for sum rate maximization in wireless networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 3, pp. 706–719, Jun. 2013.
[15] L. Zheng and C. W. Tan, "Cognitive radio network duality and algorithms for utility maximization," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 500–513, Mar. 2013.
[16] L. Zheng and C. W. Tan, "Maximizing sum rates in cognitive radio networks: Convex relaxation and global optimization algorithms," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 3, pp. 667–680, Mar. 2014.
[17] U. Krause, "Concave Perron–Frobenius theory and applications," *Nonlinear Anal.*, vol. 47, no. 3, pp. 1457–1466, 2001.
[18] B. Lemmens and R. Nussbaum, *Nonlinear Perron–Frobenius Theory*. Cambridge, U.K.: Cambridge Univ, 2012.
[19] C. W. Tan, "Wireless network optimization by Perron–Frobenius theory," *Found. Trends Netw.*, vol. 9, nos. 2–3, pp. 107–218, 2015.
[20] U. Krause, "Perron's stability theorem for nonlinear mappings," *J. Math. Econ.*, vol. 15, no. 3, pp. 275–282, 1986.
[21] A. Berman and R. Plemmons, *Nonnegative Matrices in The Mathematical Sciences*. London, U.K.: Academic, 1979.
[22] H. Minc, *Nonnegative Matrices*. New York, NY, USA: Wiley, 1988.
[23] M. Akian and S. Gaubert, "Spectral theorem for convex monotone homogeneous maps, and ergodic control," *Nonlinear Anal. Theory Methods Appl.*, vol. 52, no. 2, pp. 637–679, 2003.
[24] V. D. Blondel, L. Ninove, and P. V. Dooren, "An affine eigenvalue problem on the nonnegative orthant," *Linear Algebra Appl.*, vol. 404, pp. 69–84, Jul. 2005.
[25] P. Takáč, "Asymptotic behavior of discrete-time semigroups of sublinear, strongly increasing mappings with applications to biology," *Nonlinear Anal. Theory Methods Appl.*, vol. 14, no. 1, pp. 35–42, 1990.
[26] E. Dall, "Anese, S.-J. Kim, G. B. Giannakis, and S. Pupolin, "Power control for cognitive radio networks under channel uncertainty," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3541–3551, Oct. 2011.
[27] J. Joung, C. K. Ho, and S. Sun, "Spectral efficiency and energy efficiency of OFDM systems: Impact of power amplifiers and countermeasures," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 208–220, Feb. 2014.
[28] C. Rapp, "Effects of HPA-nonlinearity on a 4-DPSK/OFDM-signal for a digital sound broadcasting signal," in *Proc. 2nd Eur. Conf. Satellite Commun.*, Oct. 1991, pp. 179–184.
[29] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[30] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.
[31] K. R. Krishnan and H. Luss, "Power selection for maximizing SINR in femtocells for specified SINR in macrocell," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2011, pp. 563–568.
[32] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum–utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.
[33] H. P. Shiang and M. V. D. Schaar, "Queuing-based dynamic channel selection for heterogeneous multimedia applications over cognitive radio networks," *IEEE Trans. Multimedia*, vol. 10, no. 5, pp. 896–909, Aug. 2008.
[34] Y. Wu *et al.*, "A learning-based QoE-driven spectrum handoff scheme for multimedia transmissions over cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 11, pp. 2134–2148, Nov. 2014.

Liang Zheng (S'13–M'15) is a postdoctoral research associate at Department of Electrical Engineering of Princeton University. She received the bachelor's degree in software engineering from Sichuan University, Chengdu, China, in 2011, and the Ph.D. in computer science from the City University of Hong Kong in 2015. Previously, she was a visiting student at Princeton University in 2014. Her research interests are in big data analytics, network economics, cloud computing, communication networks, nonlinear optimization and its applications. She was a recipient of the IBM Students for a Smarter Planet Award in 2014. She was selected as one of the 2016 Rising Stars in EECS at the Carnegie Mellon University Rising Stars Workshop.

Y.-W. Peter Hong (S'01–M'05–SM'13) received his B.S. degree from National Taiwan University, Taipei, Taiwan, in 1999, and his Ph.D. degree from Cornell University, Ithaca, NY, in 2005, both in electrical engineering. He joined the Institute of Communications Engineering and the Department of Electrical Engineering at National Tsing Hua University, Hsinchu, Taiwan, in Fall 2005, where he is now a Full Professor. His research interests include physical layer secrecy, cooperative communications, and signal processing for wireless networks. Dr. Hong received the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2010, the Y. Z. Hsu Scientific Paper Award in 2011, the National Science Council (NSC) Wu Ta-You Memorial Award in 2011, the Chinese Institute of Electrical Engineering (CIEE) Outstanding Young Electrical Engineer Award in 2012, and the Best Paper Award from the Asia- Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) in 2013. Dr. Hong served as Associate Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING (2010-2014) and IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY (2011-2016), and is currently an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS.

Chee Wei Tan (M'08–SM'12) received the M.A. and Ph.D. degrees in electrical engineering from Princeton University in 2006 and 2008, respectively. He is an Associate Professor with the City University of Hong Kong. He was a Postdoctoral Scholar at the California Institute of Technology (Caltech). He was a Visiting Faculty with Qualcomm R&D in 2011. His research interests are in networks, statistical inference in data analytics, and optimization theory and its applications. Dr. Tan was the recipient of the 2008 Princeton University Gordon Wu Prize for Excellence. He was the Chair of the IEEE Information Theory Society Hong Kong Chapter in 2014 and 2015, during which he received the 2015 IT Society Chapter of the Year Award. He was twice selected to participate at the U.S. National Academy of Engineering China- America Frontiers of Engineering Symposium in 2013 and 2015. He currently serves as an Editor for both the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE/ACM TRANSACTIONS ON NETWORKING.

Cheng-Lin Hsieh received his B.S. degree in electrical engineering, in 2010, and his M.S. degree in communications engineering, in 2013, both from National Tsing Hua University, Hsinchu, Taiwan. He joined ASUSTek Computer Inc., Taipei, Taiwan, in 2014, where he is now a senior engineer.

Chia-Han Lee (M'10) received the B.S. degree from National Taiwan University in 1999, the M.S. degree from the University of Michigan, Ann Arbor, in 2003, and the Ph.D. degree from Princeton University in 2008, all in electrical engineering. From 1999 to 2001, he served in the ROC Army as a Missile Operation Officer. From 2008 to 2009, he was a Postdoctoral Research Associate with the University of Notre Dame, USA. From 2010 to 2016, he was with Academia Sinica as an Assistant Research Fellow and then an Associate Research Fellow. Since 2016, he has been with National Chiao Tung University as an Associate Professor. His research interests include wireless communications and networks. He is an Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and IEEE COMMUNICATIONS LETTERS.