Fast Algorithms and Performance Bounds for Sum Rate Maximization in Wireless Networks

Chee Wei Tan,1 Mung Chiang2 and R. Srikant3
1 City University of Hong Kong  2 Princeton University  3 University of Illinois at Urbana-Champaign

Abstract—Sum rate maximization by power control is a nonconvex optimization problem and achieves a rate region that is in general nonconvex. We derive approximation ratios to the sum rate optimal value by studying the solutions to two related problems, sum rate maximization using an SIR approximation and max-min weighted SIR optimization. We also show that these two problems can be solved efficiently, using much faster algorithms than the existing ones in the literature. Furthermore, using a new parameterization of the sum rate maximization problem, we obtain a characterization of the power controlled rate region and its convexity property in various asymptotic regimes. As an application of these results, we analyze the connection-level stability of utility maximization term priority) and assigned by the network to the various asymptotic regimes. As an application of these results, we analyze the connection-level stability of utility maximization (obtained by varying \(\mathbf{w}\) and solving for \(\mathbf{p}^*\) accordingly) is in general a nonconvex set. Moreover, (1) may even be hard to approximate \([8]\). This paper aims at answering the following questions, interesting in their own right as well as for their importance in understanding cross-layer optimization involving transmit powers:

- Can some related efficiently-solvable problems provide provable approximation ratios to (1)? We show how to solve (1) by (i) approximating the function that describes rate as a function of SIR and (ii) by solving the max-min SIR problem in Section III (Theorem 1, 3, 4). We derive algorithms that are much faster than existing algorithms for these two extensively studied problems, and then quantify their approximation ratios with respect to \(\mathbf{p}^*\) of (1) in Section IV (Theorem 6,7).
- Can we completely characterize the resulting rate region obtained by solving (1)? We provide the answer in a closed-form expression in Section V (Theorem 8), which quantifies the intuition that power-controlled rate region is convex for sufficiently weak interference channels or sufficiently small maximum powers.

Overall, the contributions of the paper are as follows:

1) We start with the weighted sum rate maximization problem. Then, as in past work based on Geometric Programming (GP) \([1]\), we use convex approximation by approximating \(\log(1 + \text{SIR})\) by \(\log\text{SIR}\), followed by a change of variables. At this point, we make a departure and show that there exists a faster algorithm than the traditional gradient algorithm to solve this SIR approximation power control (SAPC) problem.

2) Then, we consider the max-min weighted SIR problem. During the process of convexifying the problem, we discover an unexpected connection to the SAPC problem, which allows to come up with a fast two time-scale algorithm. In addition, we derive a closed-form expression for the optimal power levels which is of independent theoretical interest and also of practical use in small networks. In the equal maximum power case, we also derive a new algorithm for the max-min weighted SIR problem by exploiting a connection to the nonlinear Perron-Frobenius theory.

3) We derive a condition (a nonnegative matrix \(\mathbf{\bar{B}}\) exists) under which we are able to bound the ratio of the optimal weighted sum rate to the sum rate obtained under SAPC and max-min weighted SIR problem.

4) For joint power control and scheduling, we first study the geometry of the rate region. If the rate region is convex, then time-sharing is not necessary; otherwise, it is. We obtain a closed-form expression for the capacity region without time-sharing and show that if the interference channel gains are small or if the maximum powers are small, then the region is convex. Otherwise, it may not be.

5) When the capacity region without time-sharing is not convex, then it is interesting to understand if power control is useful at all. For this purpose, we study the impact of solving the weighted sum rate maximization (the same as the maxweight problem (e.g., in \([10]\)–[13]) if queue lengths are weights) using our max-min weighted

I. INTRODUCTION

The following problem of sum rate maximization through power control has been extensively studied in wireless and DSL network design (e.g., a partial list of recent work include [1]–[9]):

\[
\begin{align*}
\text{maximize} & \quad \sum_i w_i \log(1 + \text{SIR}_i(\mathbf{p})) \\
\text{subject to} & \quad 0 \leq p_i \leq \bar{p}_i \quad \forall i, \\
\text{variables:} & \quad \mathbf{p}_i \quad \forall i,
\end{align*}
\]

(1)

where \(p_i\) is the transmit power, \(w_i\) is some positive weight assigned by the network to the \(i\)th link (to reflect some long-term priority) and \(\text{SIR}_i(\mathbf{p})\) is the signal-to-interference noise ratio. Without loss of generality, we assume that \(\mathbf{w}\) is a probability vector. Denote the optimal solution to (1) by \(\mathbf{p}^*\).

This fundamental problem is difficult to solve or to analyze:

1) It is a nonconvex optimization problem and the resulting rate region (obtained by varying \(\mathbf{w}\) and solving for \(\mathbf{p}^*\) accordingly) is in general a nonconvex set. Moreover, (1) may even be hard to approximate \([8]\). This paper aims at answering the following questions, interesting in their own right as well as for their importance in understanding cross-layer optimization involving transmit powers:

- Can some related efficiently-solvable problems provide provable approximation ratios to (1)? We show how to solve (1) by (i) approximating the function that describes rate as a function of SIR and (ii) by solving the max-min SIR problem in Section III (Theorem 1, 3, 4). We derive algorithms that are much faster than existing algorithms for these two extensively studied problems, and then quantify their approximation ratios with respect to \(\mathbf{p}^*\) of (1) in Section IV (Theorem 6,7).
- Can we completely characterize the resulting rate region obtained by solving (1)? We provide the answer in a closed-form expression in Section V (Theorem 8), which quantifies the intuition that power-controlled rate region is convex for sufficiently weak interference channels or sufficiently small maximum powers.

Overall, the contributions of the paper are as follows:

1) We start with the weighted sum rate maximization problem. Then, as in past work based on Geometric Programming (GP) \([1]\), we use convex approximation by approximating \(\log(1 + \text{SIR})\) by \(\log\text{SIR}\), followed by a change of variables. At this point, we make a departure and show that there exists a faster algorithm than the traditional gradient algorithm to solve this SIR approximation power control (SAPC) problem.

2) Then, we consider the max-min weighted SIR problem. During the process of convexifying the problem, we discover an unexpected connection to the SAPC problem, which allows to come up with a fast two time-scale algorithm. In addition, we derive a closed-form expression for the optimal power levels which is of independent theoretical interest and also of practical use in small networks. In the equal maximum power case, we also derive a new algorithm for the max-min weighted SIR problem by exploiting a connection to the nonlinear Perron-Frobenius theory.

3) We derive a condition (a nonnegative matrix \(\mathbf{\bar{B}}\) exists) under which we are able to bound the ratio of the optimal weighted sum rate to the sum rate obtained under SAPC and max-min weighted SIR problem.

4) For joint power control and scheduling, we first study the geometry of the rate region. If the rate region is convex, then time-sharing is not necessary; otherwise, it is. We obtain a closed-form expression for the capacity region without time-sharing and show that if the interference channel gains are small or if the maximum powers are small, then the region is convex. Otherwise, it may not be.

5) When the capacity region without time-sharing is not convex, then it is interesting to understand if power control is useful at all. For this purpose, we study the impact of solving the weighted sum rate maximization (the same as the maxweight problem (e.g., in \([10]\)–[13]) if queue lengths are weights) using our max-min weighted
SiR algorithm. Using standard IEEE 802.11 values for RTS and CTS SiR thresholds, we numerically show that \( \mathbf{I} \) exists in most network scenarios. Further, we apply our results to analyze the connection-level stability of flows in a wireless network when the sending rates are subject to congestion control for utility maximization.

This paper is organized as follows. We present the system model in Section II. We look at two power control problems and derive fast algorithms to solve them in Section III. In Section IV, we sharpen and apply a variety of tools from nonnegative matrix theory to derive performance guarantees of our algorithms. The characterization of the rate region (treating interference as noise) and its region geometry at different SNR regimes are presented in Section V. In Section VI, we apply our results and algorithms to investigate the connection-level stability of a utility maximization problem with congestion control. We conclude with a summary in Section VII. All the proofs can be found in the Appendix.

The following notation is used. Boldface uppercases denote matrices, boldface lowercase denote column vectors, italics denote scalars, and \( \mathbf{u} \geq \mathbf{v} \) denotes componentwise inequality between vectors \( \mathbf{u} \) and \( \mathbf{v} \). We also let \( \langle \mathbf{By} \rangle \) denote the \( l \)th element of \( \mathbf{By} \). Let \( \mathbf{x} \circ \mathbf{y} \) denote the Schur product of the vectors \( \mathbf{x} \) and \( \mathbf{y} \), i.e., \( \mathbf{x} \circ \mathbf{y} = [x_1y_1, \ldots, x_my_m] \). Let \( \| \cdot \|_\infty \) be the weighted maximum norm of the vector \( \mathbf{w} \) with respect to the weight \( \mathbf{x} \), i.e., \( \| \mathbf{w} \|_\infty = \max_i w_i / x_i, x_i > 0 \). We write \( \mathbf{B} \succeq \mathbf{F} \) if \( B_{ij} \geq F_{ij} \) for all \( i, j \). The Perron-Frobenius eigenvalue of a nonnegative matrix \( \mathbf{F} \) is denoted as \( \rho(\mathbf{F}) \), and the Perron (right) and left eigenvector of \( \mathbf{F} \) associated with \( \rho(\mathbf{F}) \) are denoted by \( \mathbf{x}(\mathbf{F}) \) and \( \mathbf{y}(\mathbf{F}) \geq 0 \) (or simply \( \mathbf{x} \) and \( \mathbf{y} \) when the context is clear) respectively. Recall that the Perron-Frobenius eigenvalue of \( \mathbf{F} \) is the eigenvalue with the largest absolute value. Assume that \( \mathbf{F} \) is a nonnegative irreducible matrix. Then \( \rho(\mathbf{F}) \) is simple and positive, and \( \mathbf{x}(\mathbf{F}), \mathbf{y}(\mathbf{F}) > 0 \) [14]. We will assume the normalization: \( \mathbf{x}(\mathbf{F}) \circ \mathbf{y}(\mathbf{F}) \) is a probability vector. The super-script \( (\cdot)\top \) denotes transpose. We denote \( e_l \) as the \( l \)th unit coordinate vector and \( \mathbf{I} \) as the identity matrix.

II. REVIEW OF A STANDARD SYSTEM MODEL

We consider a wireless network where the channel is interference-limited, and all the \( L \) links (equivalently, transceiver pairs) treat interference as white noise and do not use multiser detection. Such a model of communication has also been previously used (in, e.g., [1], [2], [4], [6], [7], [15]) for cellular and ad hoc networks. Figure 1(a) shows the model for a 2-user interference-limited channel.

The transmit power for the \( l \)th link is denoted by \( p_l \) for all \( l \). Assuming a matched-filter receiver, the SiR for the \( l \)th receiver is given by

\[
\text{SiR}_l(p) = \sum_{j \neq l} G_{lj} p_j + n_l,
\]

where \( G_{lj} \) are the channel gains from transmitter \( j \) to receiver \( l \) and \( n_l \) is the additive white Gaussian noise (AWGN) power for the \( l \)th receiver. The channel gain matrix \( \mathbf{G} \) takes into account propagation loss, spreading loss and other transmission modulation factors. Assuming a fixed bit error rate (BER) at the receiver, the Shannon capacity formula can be used to deduce the achievable data rate of the \( l \)th user as [16]:

\[
\log \left( 1 + \frac{\text{SiR}_l(p)}{I} \right) \text{ nats/symbol},
\]

where \( I \) is the gap to capacity, which is always greater than 1. We absorb \((I / I)\) into \( G_{lj} \), for all \( l \), and write the achievable data rate as \( \log(1 + \text{SiR}_l(p)) \).

Next, we define a nonnegative matrix \( \mathbf{F} \) with entries:

\[
F_{lj} = \begin{cases} 0, & \text{if } j = l; \\ \frac{G_{lj}}{G_{LL}}, & \text{if } l \neq j. \end{cases}
\]

and

\[
\mathbf{v} = \left( \frac{n_1}{G_{11}}, \frac{n_2}{G_{22}}, \ldots, \frac{n_L}{G_{LL}} \right)\top.
\]

Moreover, we assume that \( \mathbf{F} \) is irreducible, i.e., each link has at least an interferer. In matrix notation, the \( l \)th user has \( \text{SiR}_l(p) = p_l / (\langle \mathbf{F} \rangle + n_l) \) and signal-to-noise ratio (SNR) \( p_l / n_l \).

III. FAST POWER CONTROL ALGORITHMS

In this section, we consider two widely studied power control problems. The first one we consider is a direct approximation to (1), i.e., maximize a weighted sum of link rates by first approximating the link rate expression \( \log(1 + \text{SiR}) \) by \( \log(\text{SiR}) \), perform a change of variables to convexify the problem, and then use a gradient projection technique to solve the problem [1], [7]. We call such an approximation method SAPC. Here, we show that the gradient approach can be replaced by an alternative procedure that leads to faster convergence. Specifically, we show that the approximation to the link rate expression can be viewed in terms of standard interference function introduced in [17], which allows us to use greedy algorithm. Both theoretical analysis and numerical results confirm that this approach leads to much faster convergence than the gradient approach.

Another problem we consider in this section is the widely-studied max-min weighted SiR power control problem where the goal is to maximize the minimum weighted SiR at any receiver in the network. By studying its Lagrangian dual problem, we observe an interesting connection between the
max-min weighted SIR problem and SAPC, which allows us to use the interference function approach mentioned earlier as an intermediate step to solve the max-min weighted SIR problem as well. In the special case when the weights in the max-min SIR problem are equal and when all nodes in the network have the same maximum transmit power constraint, we develop an even faster power-control algorithm.

We also derive theoretical results which are of interest in their own right. We first derive a closed-form expression for the max-min weighted power control problem and then derive conditions under which the max-min SIR problem and SAPC have the same solution. These results are shown by exploiting a connection between power-control problems and nonnegative matrix theory. Figure 1(b) overviews the connections between these main optimization problems in the respective sections.

A. SIR approximation power control (SAPC)

In this section, we consider the following SIR approximation to (1):

\[
\text{maximize } \sum_j w_j \log \text{SIR}_j(p) \quad \text{subject to } 0 \leq p_l \leq \tilde{p} \quad \forall l, \quad \text{variables: } p_l \quad \forall l.
\]

(6)

It is well-known that problem (6) can be turned into a GP [1], [7]. We denote \( \tilde{p} \) as the optimal solution to (6). In particular, by making a change of variable, i.e., \( \tilde{p}_l = \log p_l \) for all \( l \), (6) is convex in \( \tilde{p} \) and thus \( \tilde{p}_l = \log \tilde{p}_l \) for all \( l \) [7]. To solve (6) optimally, there are gradient-based algorithms (requiring step-size tuning) in [7], which are based on the dynamical system \( \partial f / \partial t = f(p) \) for some function \( f \) after changing the variables in \( \tilde{p} \) to \( p \).

A different approach is to derive a fixed point iteration, i.e., \( p = g(p) \) for some function \( g \) based on the Karush-Kuhn-Tucker (KKT) optimality conditions (see [18]). Now, it can be verified that the KKT conditions of the convex form of (6) in the variables \( \tilde{p} \) can be rewritten in terms of \( p \) as \( p = \min(I(p), \tilde{p}) \) for some vector function \( I \). Furthermore, it can be shown that \( I(p) \) is a standard interference function [17]. By leveraging the interference function results in [17], this leads us to propose the following (step-size free) algorithm that converges to the optimal solution of (6), and geometrically fast when the initial point is \( \tilde{p} \). This algorithm is also used as a building block for the max-min weighted SIR power control later.

\textbf{Algorithm 1 (SAPC Algorithm):}

1) Update \( p(k+1) \):

\[
p(k+1) = \min \left\{ \frac{w_j E_{G_{ij}} \text{SIR}_j(p(k))}{p_j(k)} \right\} \tilde{p}.
\]

(7)

for all \( l \), where \( k \) indexes discrete time slots.

\textbf{Remark 1:} The information required for computation in (7) can be obtained by distributed message passing: For \( j \neq l \), the \( j \)th user first computes \( w_j E_{G_{ij}} \text{SIR}_j(p(k)) / (G_{ij} p_j) \) and measures \( G_{ij} \) by pilot signal transmitted from the \( l \)th user. Then, \( G_{ij} w_j \text{SIR}_j(p(k)) / (G_{ij} p_j) \) is broadcasted to the \( l \)th user for computation.

Theorem 1: Starting from any initial point \( p(0) \), \( p(k) \) in Algorithm SAPC converges to \( \tilde{p} \) asymptomatically under asynchronous update.

Corollary 1: Starting from the initial point \( p(0) = \tilde{p} \), \( p(k) \) in Algorithm SAPC converges synchronously at a geometric rate to \( \tilde{p} \).

Remark 2: The proofs of Theorem 1 and Corollary 1 follow the standard interference function approach [17], and the key is to show that (7) is a standard interference function.

Remark 3: The choice of \( p(0) = \tilde{p} \) does not limit the application of Corollary 1, because it can be shown that, at optimality of (6), some users transmit at maximum power. Hence, some users will not need further power update since they are already optimal at the initial step.

Example 1: We evaluate the performance of Algorithm SAPC and the subgradient algorithm in [7]. We adopt the path loss model with a path loss exponent of 3.7, log-normal shadowing with standard deviation of 8.9dB, and we assume slow fading. Ten users are distributed uniformly in a cell. All users have the same maximum power of 33mW. There are many standard choices of step-size \( \alpha(k) \) used in the subgradient method [19]. We use a (sufficiently small) constant stepsize and \( \alpha(k) = (m+1)/(m+k) \) for \( m = 5, 500 \) [19] for the subgradient approach in [7]. Figure 2 shows the evolution of SIR, which illustrate that Algorithm SAPC converges much faster than the subgradient algorithm in [7], regardless of the subgradient stepizes. In most of our simulations, the convergence speed of Algorithm SAPC is several orders magnitude faster than the subgradient algorithm.

B. Max-min weighted SIR power control

In this section, we study constrained max-min weighted SIR power control. Let \( \beta \) be a priority vector similar to the use of \( w \) in (6). Consider the following problem:

\[
\text{maximize } \min_l \frac{\text{SIR}_l(p)}{\beta_l} \quad \text{subject to } p \leq \tilde{p}, \quad \text{variables: } p.
\]

(8)

![Fig. 2. Performance comparison of Algorithm SAPC and the subgradient approach in [7] that uses different stepizes \( \alpha \) of i) \( \alpha(k) = (5 + 1)/(5 + k) \), ii) \( \alpha(k) = (50 + 1)/(50 + k) \) and iii) \( \alpha(k) = 0.01 \). Evolution of SIR.](image)
By exploiting a connection between nonnegative matrix theory and the algebraic structure of max-min SIR power control, we can give a closed form solution to (8).

**Theorem 2:** An optimal solution to (8) is such that the weighted SIR for all users are equal. This weighted SIR is given by

\[ \gamma^* = \frac{1}{\rho(\text{diag}(\beta)(\mathbf{F} + (1/\bar{p}_i)\mathbf{v}_i^\top))}, \]

where

\[ i = \arg\min_t \frac{1}{\rho(\text{diag}(\beta)(\mathbf{F} + (1/\bar{p}_i)\mathbf{v}_i^\top))}. \]

Further, all links \( i \) that achieve the minimum in (10) transmit at maximum power \( \bar{p}_i \) and the rest do not. The optimal \( \mathbf{p} \), denoted by \( \mathbf{p}^* \), is \( \mathbf{t}\times(\text{diag}(\beta)(\mathbf{F} + (1/\bar{p}_i)\mathbf{v}_i^\top)) \) for a constant \( t = \bar{p}_i/x_i \).

**Remark 4:** Suppose \( \beta = 1 \) and we let \( \bar{p}_i \to \infty \) for all \( l \) or let \( \mathbf{v} = 0 \) in Theorem 2, we obtain as a special case the result in [20], where the additive white Gaussian noise and maximum power constraints are assumed not to be present.

While the closed form solution in Theorem 2 is only useful when the eigenvector \( \mathbf{x}(\text{diag}(\beta)(\mathbf{F} + (1/\bar{p}_i)\mathbf{v}_i^\top)) \) can be computed centrally, we next discuss how to solve (8) for any number of users distributively. The proof of Theorem 2 computes an optimal dual function that has a form similar to the SAPC problem in (6) (cf. (45) and (46) in Appendix), which allows us to evaluate this dual function, possibly asynchronously, with fast convergence (cf. Theorem 1 and Corollary 1). This leads us to solve (8) using the following max-min weighted SIR Algorithm.

**Algorithm 2 (Max-min Weighted SIR Algorithm):**

1. Initialize an arbitrarily positive \( \mathbf{w}(0) \) and small \( \epsilon, \alpha(1) \).
2. Set \( \mathbf{p}(0) = \bar{\mathbf{p}} \). Repeat
   
   \[ p(k + 1) = \min \left\{ \frac{w_i(t)}{\sum_{j \neq i} w_j(t)/\text{SIR}_j(p(k))}, \bar{p}_i \right\} \]
   
   until \( \|\mathbf{p}(k + 1) - \mathbf{p}(k)\| \leq \epsilon \).
3. Compute \( u_i(t + 1) = \max\{u_i(t) + \alpha(t)\sum_j w_j(t) \log(\text{SIR}_j(p(k+1))/\beta_j) - \log(\text{SIR}_j(p(k+1))/\beta_j), 0\} \) for all \( l \), where \( t \) indexes discrete time slots much larger than \( k \).
4. Normalize \( \mathbf{w}(t + 1) \) so that \( \mathbf{1}^\top \mathbf{w}(t + 1) = 1 \). Go to Step 2.

**Theorem 3:** Starting from any initial point \( \mathbf{w}(0) \) and small \( \epsilon \), if the positive step-size \( \alpha(t) \) is strictly less than

\[ 2(-\log \rho(\text{diag}(\beta)(\mathbf{F} + (1/\bar{p}_i)\mathbf{v}_i^\top)) - \sum_j u_j(t)/\text{SIR}_j(p(\mathbf{w}(t))))/\beta_j) - \log \frac{\text{SIR}_j(p(\mathbf{w}(t)))}{\beta_j} \]

where \( i \) is given in (10) and \( g(t) \) is the following faster (single timescale and step-size free) Algorithm in special case

\[ p(k + 1) = \frac{\beta_i}{\text{SIR}_i(p(k))}, \forall l. \]

2. Normalize \( p(k + 1) \):

\[ p_i(k + 1) \leftarrow p_i(k + 1) \cdot \bar{p}_i / \max_j p_j(k + 1), \forall l. \]

**Algorithm 3 (Equal Power Max-min Weighted SIR Algorithm):**

1. Update power \( p_i(k + 1) \):

\[ p_i(k + 1) = \frac{\beta_i}{\text{SIR}_i(p(k))}, \forall l. \]

2. Normalize \( p_i(k + 1)\):

\[ p_i(k + 1) \leftarrow p_i(k + 1) \cdot \bar{p}_i / \max_j p_j(k + 1), \forall l. \]

**Remark 5:** For a given \( \mathbf{w}(t) \) at Step 2 of the Max-min SIR Algorithm, Algorithm SAPC is used as an intermediate iterative method, whose geometrical convergence rate is guaranteed by Corollary 1.

Now, we connect Algorithm SAPC and Algorithm 2. The following result illustrates that if the weight vector \( \mathbf{w} \) in (6) is chosen in a particular form as a function of \( \beta \), the solution obtained by Algorithm SAPC is equivalent to that of (8).

**Corollary 2 (Connecting Max-min Weighted SIR and SAPC):** Let \( \mathbf{x} \) and \( \mathbf{y} \) be the Perron and left eigenvectors of \( \text{diag}(\beta)(\mathbf{F} + (1/\bar{p}_i)\mathbf{v}_i^\top) \) respectively, where \( i \) is defined in (10). Recall that \( \mathbf{t}\times \) is also the optimal power vector for the Max-min SIR problem for some \( t > 0 \). Now consider the SAPC with \( \mathbf{w} = \mathbf{x} \circ \mathbf{y} \). Then, \( \mathbf{t}\times \) is also the optimal power vector for the SAPC.

From Corollary 2, we deduce that the optimal SIR allocation in (6) is a weighted geometric mean of the optimal SIR in (8), where the weights are the Schur product of the Perron and left eigenvectors of \( \text{diag}(\beta)(\mathbf{F} + (1/\bar{p}_i)\mathbf{v}_i^\top) \):

\[ \prod_i (\text{SIR}_i(\mathbf{t}\times))^{\alpha_i} = 1/\rho(\text{diag}(\beta)(\mathbf{F} + (1/\bar{p}_i)\mathbf{v}_i^\top)), \]

where \( i \) is given in (10) and \( t = \bar{p}_i/x_i \).

**C. A faster max-min weighted SIR Algorithm in special case**

Algorithm (2) is a (two-timescale) primal-dual algorithm that solves (8) for any maximum power constraint, but requires a step-size that is adapted iteratively. Under the special case when \( \bar{p}_i = \bar{p} \) for all \( l \), we give the following faster (single timescale and step-size free) algorithm to solve (8), also with the added advantage of backward compatibility with existing CDMA power control.

**Algorithm 4:** Starting from any initial point \( \mathbf{p}(0) \), \( \mathbf{p}(k) \) in Algorithm 3 converges geometrically fast to \( \mathbf{x}(\text{diag}(\beta)(\mathbf{F} + (1/\bar{p}_i)\mathbf{v}_i^\top)) \) (unique up to a scaling constant).

**Remark 6:** Interestingly, Step 1 of Algorithm 3 is simply the DPC algorithm in [21], where the \( l \)th user has a virtual SIR threshold of \( \beta_i \). DPC algorithm can be implemented using a local feedback from the \( l \)th receiver to the \( l \)th transmitter after a local SIR measurement. The only global coordination is for computing \( \max_j p_j(k + 1) \) in Step 2.

Using the numerical simulation model in Example 1, Figure 3 shows that Algorithm 3 converges much faster than Algorithm 2. Algorithm 2 however applies to the case where individual power constraint can differ and is thus more general.
IV. APPROXIMABILITY AND PERFORMANCE GUARANTEES

The two problems (6,8) in the last section can be solved efficiently and sometimes distributedly. But back to our original question, how well do they approximate the difficult problem of sum rate maximization (1)? In this section, we will use related results in [22] to compute the approximation ratios of Algorithm 1 and the max-min weighted SIR power control (using Algorithm 2 or Algorithm 3) under certain conditions. In addition, conditions under which the fast algorithms in Section III solve (1) optimally are also given.

A. Approximation ratio

First, we rewrite (1) in matrix notation as

$$\begin{align*}
\text{maximize} & \sum_i w_i \log((\mathbf{I} + \mathbf{F}) \mathbf{p} + \mathbf{v})_i / ((\mathbf{Fp} + \mathbf{v})_i) \\
\text{subject to} & \mathbf{p} \leq \overline{\mathbf{p}} \\
\text{variables:} & \mathbf{p}.
\end{align*}$$

(13)

Next, we consider a relaxed (but still nonconvex) version of (13):

$$\begin{align*}
\text{maximize} & \sum_i w_i \log((\mathbf{I} + \mathbf{F}) \mathbf{p} + \mathbf{v})_i / ((\mathbf{Fp} + \mathbf{v})_i) \\
\text{subject to} & \mathbf{1}^{\top} \mathbf{p} \leq 1 \\
\text{variables:} & \mathbf{p}.
\end{align*}$$

(14)

The optimal objective of (14) is larger than that of (13). Despite its nonconvexity, the optimal solution to (14) can be obtained analytically under a sufficient condition on the problem parameters [22]. We describe this condition in the following.

Let \( \mathbf{B} \) be a nonnegative matrix with entries:

$$\mathbf{B} = \mathbf{F} + \sum_i \frac{1}{\mathbf{F}^{\top} \mathbf{v}_i}. \quad (15)$$

Following [22], we introduce the notion of quasi-invertibility of a nonnegative matrix in [23], particularly of \( \mathbf{B} \) in (15), which will be useful in solving (14) optimally.

Definition 1 (Quasi-invertibility): A square nonnegative matrix \( \mathbf{B} \) is a quasi-inverse of a square nonnegative matrix \( \mathbf{B} \) if \( \mathbf{B} - \mathbf{B} = \mathbf{BB} = \mathbf{BB} \). Furthermore, \( (\mathbf{I} - \mathbf{B})^{-1} = \mathbf{I} + \mathbf{B} \) [23].

The existence of \( \mathbf{B} \) can be related to the different SNR regimes. Roughly speaking, in the high SNR regime or when interference (off-diagonals of \( \mathbf{F} \)) is very large, \( \mathbf{B} \) does not exist. On the other hand, in the sufficiently low SNR regime or when interference is small, \( \mathbf{B} \) exists. In the following, we assume that \( \mathbf{B} \) exists. Let \( \mathbf{z} = (\mathbf{I} + \mathbf{B}) \mathbf{p} \), we can rewrite (14) in terms of \( \mathbf{z} \) [22]:

$$\begin{align*}
\text{maximize} & \sum_i w_i \log\left(\frac{z_i}{(\mathbf{Bz})_i}\right) \\
\text{subject to} & (\mathbf{I} - \mathbf{B}) \mathbf{z} \geq 0.
\end{align*}$$

(16)

Our main result is to derive a tight upper bound to (16) in terms of the problem parameters, which in turn upper bounds (1), and also gives insight into the relationship between (1) and max-min SIR power control. We state this result as follows.

Theorem 5: If \( \mathbf{B} \) exists, then

$$\begin{align*}
\sum_i w_i \log(1 + \text{SIR}_i(\mathbf{p}^*)) \leq \|\mathbf{w}\|_\infty \log(1 + 1/\rho(\mathbf{B})),
\end{align*}$$

(17)

where \( \mathbf{w}, \mathbf{y} \) are the Perron and left eigenvectors of \( \mathbf{B} \) respectively.

Equality is achieved if and only if \( \mathbf{w} = \mathbf{x} \circ \mathbf{y} \) and \( \rho(\mathbf{B}) = \rho(\mathbf{F} + (1/\mathbf{p}) \mathbf{v}_i^\top) \). In this case, \( \mathbf{p}^* = \mathbf{x}(\mathbf{B}) \) (unique up to a scaling constant).

Remark: Theorem 5 illustrates a sufficient condition under which (8) (with \( \beta = 1 \)) optimally solves (1).

Based on Theorem 5, we now give the approximation ratio to (1) by solving the max-min weighted SIR problem (with \( \beta = \mathbf{w} \)) using Algorithm 2.

Theorem 6: Suppose \( \mathbf{B} \) exists. Let

$$\eta = \frac{\sum_i w_i \log(1 + w_i/\rho(\text{diag}(\mathbf{w})(\mathbf{F} + (1/\mathbf{p}) \mathbf{v}_i^\top)))}{\|\mathbf{w}\|_\infty \|\mathbf{B}\|_\infty \log(1 + 1/\rho(\mathbf{B}))}, \quad (18)$$

where \( i \) is given in (10).

Then, \( \eta \) is an approximation ratio to (1) by solving (8).

Similarly, the approximation ratio of Algorithm SAPC in solving (1) can be stated as follows.

Theorem 7: Suppose \( \mathbf{B} \) exists. Let

$$\eta = \frac{\sum_i w_i \log(1 + \text{SIR}_i(\mathbf{p}^*'))}{\|\mathbf{w}\|_\infty \|\mathbf{B}\|_\infty \log(1 + 1/\rho(\mathbf{B}))}, \quad (19)$$

where \( \mathbf{p}^* \) is the optimal solution to (6). Then, \( \eta \) is an approximation ratio to (1) by solving (6).

B. General approximability

The results in Section IV-A depend on the existence of a nonnegative \( \mathbf{B} \). If this sufficient condition is not satisfied, we show how the results in Section IV-A can still be used to construct useful upper bounds to (1). First, we define a transmission configuration as a set of links \( \mathcal{C} = \{l | l = 1, \ldots, L\} \) with \( |\mathcal{C}| \leq L \). Users in \( \mathcal{C} \) transmit with positive power. In general, a transmission configuration can be used to construct a scheduling policy in which users that belong to \( \mathcal{C} \) transmit one at a time. Clearly, \( |\mathcal{C}| + \bar{\mathcal{C}} = L \). For example, when \( \mathbf{B} \) exists and all users have positive optimal power, we need to only consider \( \mathcal{C} \) such that \( |\mathcal{C}| = L \). A scheduling policy determines how users in \( \mathcal{C} \) and \( \bar{\mathcal{C}} \) are time-shared in
time-division multiple access. When there are $L$ users, there are
\[\sum_{l=1}^{L} \binom{L}{l} \frac{2^L - L}{1} = 2^L - L\]
(20)

possible transmission configurations. Now, for any transmission configuration $C$ and any $p$, we have

\[\sum_{l=1}^{L} w_l \log(1 + \text{SIR}_l(p)) \leq \sum_{l \in C} w_l \log(1 + \text{SIR}_l(p)) + \sum_{l \notin C} w_l \log(1 + \tilde{p}_l / \nu_l),\]
(21)

where $\text{SIR}_l(p), l \in C$, in the first summand on the right-hand side of (21) contains only interference terms coming from users in $C$. Therefore, it can be deduced that

\[\sum_{l=1}^{L} w_l \log(1 + \text{SIR}_l(p^*)) \leq \max_{0 \leq \rho \leq \rho_0} \sum_{l \in C} w_l \log(1 + \text{SIR}_l(p)) + \sum_{l \notin C} w_l \log(1 + \tilde{p}_l / \nu_l).
\]
(22)

Let $B_C$ denote a submatrix obtained from $B$ by deleting those rows and columns whose indices belong to $C$. If $B_C$ is a quasi-inverse of a nonnegative matrix, we denote that matrix as $\tilde{B}_C$. Hence, for any transmission configuration $C$, if $\tilde{B}_C$ exists, we can use the previous results to deduce the bound:

\[\sum_{l=1}^{L} w_l \log(1 + \text{SIR}_l(p^*)) \leq \max_{0 \leq \rho \leq \rho_0} \sum_{l \in C} w_l \log(1 + \frac{1}{\rho(\tilde{B}_C)}),\]
(23)

Subject to the existence of quasi-inverses, the tightest upper bounds to (1) can be found by searching all $2^L - L$ transmission configurations. Note that if $B$ exists, then $\tilde{B}_C$ exists for any $C$. On the other hand, as $|C|$ gets smaller, then it is more likely that $\tilde{B}_C$ exists, because the second summand in a submatrix of (15) tends to dominate the first summand.

**Example 2:** We now examine the implications of our results for practical interference-limited networks such as the IEEE 802.11b ad hoc network using a numerical example. We also consider the on-off scheduling algorithm, which finds the power vector that maximizes sum rates in which users either transmit at maximum or zero power, as a baseline for comparison with SAPC and max-min SIR. Table I records a typical numerical example for a ten-user network, where the maximum power is set as 33mW and 1W (the largest possible value allowed in IEEE 802.11b). We set $w = x(B) \circ y(B)$.

For each maximum power constraint, we average the percentage of instances where $\hat{B}$ exists, the sum rates, and the approximation ratios over 10,000 random instances. In the case where $\hat{B}$ does not exist, the tightest upper bound using the general approximability technique in Section IV-B is computed and the approximation ratios of SAPC, max-min weighted SIR (with $\beta = w$) and on-off scheduling are computed. Recorded in parentheses, the approximation ratios for max-min weighted SIR (with $\beta = w$) and SAPC are computed using Theorem 6 and 7 respectively, and the approximation ratio of on-off scheduling is computed using Theorem 5. Also shown, without parentheses in each cell, are the actual ratios of the achieved rates by the respective algorithms to the global optimal sum rates. As shown in Table I, the sufficient condition for the existence of $\hat{B}$ holds over a large proportion of time, and the relatively large approximation ratios using SAPC and max-min SIR indicate the usefulness of using power control to maximize sum rates in a typical IEEE 802.11b network.

**Table I**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Avg. % of $\hat{B}$ exists</th>
<th>SAPC (%)</th>
<th>Max-min SIR (%)</th>
<th>On-off sched. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_t = 33mW \forall l$</td>
<td>99 (0.93)</td>
<td>0.97 (0.93)</td>
<td>0.99 (0.96)</td>
<td>0.89 (0.84)</td>
</tr>
<tr>
<td>$p_t = 1W \forall l$</td>
<td>65 (0.82)</td>
<td>0.87 (0.84)</td>
<td>0.91 (0.83)</td>
<td>0.87 (0.82)</td>
</tr>
</tbody>
</table>

A TYPICAL NUMERICAL EXAMPLE IN A TEN-USER NETWORK WITH TWO MAXIMUM POWER CONSTRAINT SETTINGS: $p_t = 33mW$ OR $p_t = 1W$ FOR ALL $l$. THE PERCENTAGE OF INSTANCES WHERE $\hat{B}$ EXISTS IS RECORDED.

V. GEOMETRY OF POWER CONTROLLED RATE REGION

We now move from the nonconvexity of the sum rate maximization problem to the nonconvexity of the resulting rate region. We first present an equivalent problem to (1) in [24] that is useful to characterize the geometry of the achievable rate region in different SNR regimes. This quantifies the intuition that if the interference channel gains or SNR are sufficiently small, the rate region is amenable to analysis as has been shown in the previous section.

**Lemma 1:** Consider the following maximization problem:

\[\text{maximize} \quad \sum_l w_l \log(1 + \gamma_l) \quad \text{subject to} \quad \rho(\text{diag}(\gamma) (F + (1 / \tilde{p})) v) = 1, \quad \forall l \quad (24)\]

variables $\gamma_l, \forall l$.

The optimal SIR vector $\gamma^*$ in (24) is related to $p^*$ in (1) by:

\[p^* = (I - \text{diag}(\gamma^*) F)^{-1} \text{diag}(\gamma^*) v.\]
(25)

Further, there exists a link $i$ such that

\[\rho(\text{diag}(\gamma^*) (F + (1 / \tilde{p})) v) \leq 1, \quad \forall l \quad (26)\]

for all $l$.

**Remark 8:** From (24), we deduce that $p^*$ can be interpreted as the Perron eigenvector (unique up to a scaling constant) of a nonnegative matrix $\text{diag}(\gamma) (F + (1 / \tilde{p})) v$ for some $i$ with the Perron-Frobenius eigenvalue of 1. This index $i$ is however not easy to determine a priori.

One of the advantages gained through (24) is to characterize the power controlled rate region, which is defined as all possible points $r_t = \log(1 + \text{SIR}_t(p))$ satisfying the constraints in (1) for all $l$ obtained by varying $w$. We next investigate the geometry of the rate region using power control and time-sharing as the SNR of each user varies.

**Theorem 8:** The achievable rate region $R$ using power control is given by

\[\{r \in R^L : \rho(\text{diag}(\text{exp}(r) - 1) (F + (1 / \tilde{p})) v) \leq 1\}. \]
(27)
The achievable rate region $\mathcal{R}$ using power control and time-sharing is given by the convex hull of $\mathcal{R}$, i.e., $\text{Co}(\mathcal{R})$.

Remark 9: From Theorem 8, we deduce that $\text{Conv}\{\mathbf{r} \in \mathcal{R} : \rho(\text{diag}(\exp(\mathbf{r}) - 1) \mathbf{F}) \leq 1\}$ is the achievable rate region in the high SNR regime (as $\bar{p} \to \infty$).

In general, computing $\mathcal{R}$ explicitly is difficult. The following example illustrates the use of Lemma 1 to determine special case solution.

Example 3: When $L = 2$ and $w_1 = w_2$, and using (26), (24) is equivalent to the following optimization problem:

$$\begin{align*}
\text{maximize} & \quad \gamma_1 + \gamma_2 + \min \left\{ \frac{\bar{p}_1 v_1 + \bar{p}_2 v_2}{F_{12} v_1 + F_{12} v_2}, \frac{\bar{p}_2 v_1 + \bar{p}_2 v_2}{F_{21} v_1 + F_{21} v_2} \right\} \\
\text{subject to} & \quad \gamma_i \leq \bar{p}_i / v_i, \quad i = 1, 2.
\end{align*}$$

(28)

Now, the optimal solution to (28) can be determined by solving two linear programs (by reducing the objective in (28) to two affine functions), whose solution lie on the boundary of the feasible set. The one with the least optimal value yields the optimal solution to (28). This implies that one of the three vectors: $(\bar{p}_1 / v_1, 0)^T$, $(0, \bar{p}_2 / v_2)^T$ or $(\bar{p}_1 / (F_{12} \bar{p}_2 + v_1), \bar{p}_2 / (F_{21} \bar{p}_1 + v_2))^T$ solves (28) optimally or, equivalently, one of the three vectors: $(\bar{p}_1, 0)^T$, $(0, \bar{p}_2)^T$ or $(\bar{p}_1, \bar{p}_2)^T$ solves (1) optimally. This is consistent with recent findings in [3], [4]. Further, if $v_1 / \bar{p}_1 \leq v_2 / \bar{p}_2$ and $F_{12} v_2 / \bar{p}_1 \leq F_{21} v_1 / \bar{p}_2$, $\bar{p}_1 = \bar{p}_2$.

When $L = 2$, Theorem 8 simplifies as follows.

Corollary 3: When $L = 2$, the achievable rate region $\mathcal{R}$ using power control only is given by

$$\begin{align*}
\{ (r_1, r_2) \in \mathbb{R}^2 : r_1 \leq \min \{ & \log \left( 1 + \frac{\bar{p}_2 \left( v_2 - 1 \right)}{F_{21} \bar{p}_2 + v_2} \right), \\
& \log \left( 1 + \frac{\bar{p}_1 \left( v_1 - 1 \right)}{F_{12} \bar{p}_1 + v_1} \right) \} \}
\end{align*}$$

(29)

Furthermore, as $\bar{p}_i \to \infty$ for all $i$, the asymptotic rate region $\mathcal{R}$ is given by

$$\begin{align*}
\{ (r_1, r_2) \in \mathbb{R}^2 : r_1 \leq \log \left( 1 + \frac{1}{F_{21} \bar{p}_2 + v_2} \right) \}.
\end{align*}$$

(30)

We note that (29) is independently and contemporaneously derived in [6] using a different approach.

VI. APPLICATION TO CONNECTION-LEVEL STABILITY WITH END-TO-END CONGESTION CONTROL

So far we have assumed that each user has an infinite backlog of data to transmit rather than a dynamic arrival model of user for each connection. In this section, we illustrate how our previous results and algorithms can be applied to analyze the connection-level stability of end-to-end flows with congestion control. For technical simplicity, we focus only on the max-min weighted SIR algorithm (Algorithm 3). We consider a network with $L$ links and $S$ classes of users. Let $H_{is}$ be 1 if the path of users of class $s$ uses link $i$, and $H_{is} = 0$, otherwise. We are interested in the congestion-level stability as flows arrive stochastically to the network and their rates are regulated by congestion control. We assume that that users of class $s$ arrive to the network according to a Poisson process with rate $\lambda_s$ and that each user transmits a file whose size is exponentially distributed with mean $1 / \mu_s$. The load brought by users of class $s$ is then $\rho_s = \lambda_s / \mu_s$. Assuming Poisson arrivals and exponential file size distribution, the evolution of $n(t)$ will be governed by a Markov process, whose transition rates are given by

$$\begin{align*}
&n_s(t) \to n_s(t) + 1, \quad \text{with rate } \lambda_s, \\
&n_s(t) \to n_s(t) - 1, \quad \text{with rate } \mu_s n_s(t) x_s(t) \text{ if } n_s(t) > 0.
\end{align*}$$

We follow the optimization-based congestion control model in [10], [12], [13], where we assume that time in the system is divided into time slots of length $T$, and also assume that the $l$th link has an implicit cost $q_l$ that corresponds to the congestion cost (in fact, $q_l$ is a scalar multiple of the actual queue length at the $l$th link) [12], [13]. These implicit costs are updated at the end of each time slot. However, users may arrive and depart at any point within a single time slot. Let $q(kT)$ denote the implicit costs at time slot $k$. The definition of connection-level stability used in this section is given as follows.

Definition 2: The Markov process $\{n(t), q(t)\}$ is stable-in-the-mean, or simply stable if

$$\lim_{t \to \infty} \frac{1}{t} \int_{0}^{t} E \left[ \sum_{s=1}^{S} n_s(t) + \sum_{l=1}^{L} q_l(t) \right] dt < \infty.$$
portional fairness for all users (with weights \( \phi \)), i.e.,

\[
\begin{align*}
\text{maximize} & \quad \sum_{s=1}^{S} n_s \phi_s \log x_s \\
\text{subject to} & \quad H(n \circ x) \preceq r(p), \\
& \quad r(p) \in \text{co}(R), \quad 0 \leq p \leq \tilde{p}, \quad l = 1, \ldots, L, \\
\text{variables:} & \quad x, \ p.
\end{align*}
\]

(32)

- Find the associated scheduling policy that stabilizes the system, i.e., each file is served in finite time and the queue at each link is bounded.

The optimization-based congestion controller can be viewed as an iterative solution that solves (32), and is based on the “dual solution” of (32) [25]. More precisely, the optimization-based congestion controller is a dual algorithm with an appropriately chosen stepsize.

Now, finding the optimal scheduling policy to stabilize the system is difficult, since \( R \) in (27) is nonconvex. Instead of finding the optimal rate allocation at each time slot \( kT \), we use the results established in Section IV to determine a suboptimal schedule that provides an overall stability guarantee to the system (also known as imperfect scheduling first proposed in [12]). More precisely, at time slot \( kT \), we compute a schedule \( r \in R \) such that it satisfies the following definition.

**Definition 3 (Slot Dependent Imperfect Scheduling):** At time slot \( kT \), a schedule \( r \in R \) is called Slot Dependent Imperfect Scheduling if it satisfies

\[
\sum_{i} q_i(kT) r_i(kT) \geq \eta(kT) \cdot \max_{r \in R} \sum_{i} q_i(kT) r_i
\]

where \( \eta(kT) \in (0, 1] \).

When \( \eta(kT) \) is a constant independent of \( k \), this corresponds to the case studied in [12]. We say the schedule given by (33) is imperfect, because it attains only a factor \( \eta(kT) \) of the achievable rate region at time slot \( kT \).\(^1\) We now investigate the case when \( \eta(kT) \) in (33) is a function of the queue size \( q(kT) \), and \( r(kT) \) is given by (3). Note that the maximization on the right-hand side of (33) is performed over \( R \) only at time slot \( kT \).

We propose the following congestion control dual algorithm that finds a schedule given by (33).

**Algorithm 4:**

1. At the beginning of each time slot \( kT \), the link that has \( q_i(kT) > 0 \) uses Algorithm 3 to transmit at a rate

\[
r_i(kT) = \log \left( 1 + \frac{q_i(kT)}{\rho \text{diag}(q(kT)) (F + (1/\bar{p}) v \text{vec}(r_i))} \right)
\]

(34)

for all \( l \), where \( j^l \) denote the maximum power \( \alpha \) of the class \( i \).

2. At any time \( t \) in time slot \( kT \) for \( kT \leq t \leq (k + 1)T \), the data rate of users of class \( s \) is given by

\[
x_s(t) = x_s(kT) = \min \left\{ \frac{\phi_s}{H(q(kT))}, M_s \right\}
\]

(35)

\(^1\)Under a node exclusive interference model, e.g., in [12], when \( \eta(kT) = \eta_\text{avg} \) \( \forall t \), \( \eta_\text{avg} \) can be viewed as a parameter indicating the degree of complexity to approximate maximum weight matching: the complexity of finding a schedule satisfying (33) decreases for smaller \( \eta_\text{avg} \).

3. At the end of each time slot \( kT \), the implicit costs are updated by each link as

\[
q((k + 1)T) = \max \{ q(kT) + \\
\alpha_l \left( \sum_{s=1}^{S} H_{ls} r_{(k+1)T} n_s(t) x_s(kT) dt - Tr_k(kT) \right), 0 \}.
\]

Let \( \mathcal{S} = \max_{i} \sum_{s=1}^{S} H_{ls} \) denote the maximum number of classes using any link and let \( \mathcal{L} = \max_{i} \sum_{s=1}^{S} H_{ls} \) denote the maximum number of links used by any class.

We now apply Theorem 6 to show the stability performance guarantee of the scheduling policy where all the links solve the max-min weighted SIR problem at each time slot \( kT \) (can be done distributively using Algorithm 3).

**Theorem 9:** Let \( i = \arg\max_i F + (1/\bar{p}) v \text{vec} \) and \( \mathbf{B} = \mathbf{F} + (1/\bar{p}) v \mathbf{1} \). Suppose that \( \mathbf{B} \geq 0 \). If \( \max_i \alpha_l \leq (16T \mathcal{S} \mathcal{L})^{-1} \min_i \phi_i / (\rho_s M_s) \), then for any offered load \( \rho \) that resides strictly inside

\[
\min_i (\mathbf{u}(X) \circ y(Y)) \log (1 + 1/\rho \mathbf{F} + (1/\bar{p}) \text{vec}(\mathbf{v})) \Lambda^{-1}, (37)
\]

the system described by the Markov process \{\( \mathbf{n}(kT), \mathbf{q}(kT) \} \) is stable in the sense of Definition 2.

In summary, Theorem 9 states that, if the offered load lies strictly in a given subset (a priori determined based only on the channel parameters and maximum power) of the stability region, Algorithm 4 (with link transmission rates given by Algorithm 3) stabilizes the queue size and the number of flows for an appropriately chosen stepsize at each link (chosen independently of the offered load).

**A. Numerical Example**

In this section, we present simulation results to illustrate the use of Algorithm 4 for imperfect scheduling. We continue to use the set-up of Example 3 for three users (classes) to demonstrate the flow performance of using Algorithm 4 in a simple one-hop network (\( \mathbf{H} = \mathbf{I} \)). The maximum power is set to \( 1W \) for all users. Algorithm 3 is used in Step 1 of Algorithm 4. We use the global optimization algorithm in [24] to solve for the optimal power control allocation on the righthand-side of (33).

We simulate the case when there are dynamic arrivals and departures of the users, whereby users of each class arrive to the network according to a Poisson process with rate \( \lambda \). Each user brings with it a file to transfer whose size is exponentially distributed with mean \( 1/\mu = 1 \) unit. We first set the arrival rate \( \lambda = 0.12 \) for each class. Figure 5 shows the trajectory of the number of flows in each class. As observed from Figure 5, the penalty of imperfect scheduling is the increase in number of flows per class (hence the increase in sojourn time). As compared to the computation time of the optimal power control allocation in [24], the fast computation time of Algorithm 4 (orders of magnitude faster) outweighs the penalty in increased sojourn time in the system.

Next, we vary the arrival rate \( \lambda \) thereby changing the traffic intensity \( \rho = \lambda/\mu \). The average total number of flows in the system is recorded for both Algorithm 4 and the optimal power control allocation. Figure 6 plots the average total number of flows versus \( \rho = \lambda/\mu \). The black dotted line indicates the
Fig. 5. Comparison of the number of flows in the three classes when Algorithm 4 and the optimal power control allocation are used.

Fig. 6. Average number of flows versus traffic intensity $\rho = \lambda / \mu$ when Algorithm 4 and the optimal power control allocation are used.

worst-case $\rho = 0.272$ as a fraction of $\Lambda$ computed by (37) in Theorem 9. As seen from Figure 6, for $\rho$ that lies outside this region, the Markov process $\{n(kT), q(kT)\}$ is still stable.

VII. CONCLUSION

Sum rate maximization is a hard problem in power control and any cross-layer design involving transmit power. But large approximation ratios can be obtained through two related problems: SIR approximation power control (SAPC) and max-min weighted SIR optimization. These two problems have also been extensively studied before, and now we have faster algorithms for them with geometric convergence rate, often independent of stepsize. These results are derived based on new techniques from nonnegative matrix theory.

We characterize the achievable rate region of sum rate maximization by power control and its convexity property. We then apply our results and algorithm to analyze the connection-level stability of utility maximization in a network, where users arrive and depart randomly and are subject to congestion control. In particular, we determine the stability region when all the links solve the max-min weighted SIR problem (using instantaneous queue size as weights) for congestion-controlled utility maximization.

ACKNOWLEDGEMENT

We acknowledge helpful discussions with Steven Low at Caltech, Shmuel Friedland at UIC and Kevin Tang at Cornell. This research has been supported in part by ONR grant N00014-07-1-0864, NSF CNS 0720570 and ARO MURI Award W911NF-08-1-0233.

APPENDIX

A. Proof of Theorem 1

The objective function in (6) can be made convex by a change-of-variable technique: Let $\tilde{p}_l = \log p_l$ for all $l$. First, ignoring the maximum power constraints in (6), we let $\partial (\sum_j w_j \log \text{SIR} (\tilde{p})) / \partial \tilde{p}_l = 0$ for all $l$ to deduce

$$ e^{\tilde{p}_l} = w_l \left( \frac{\sum_{j \neq l} w_j F_{ij} \tilde{p}_j + v_j}{\sum_{j \neq l} F_{ij} e^{\tilde{p}_j} + v_j} \right) \forall l. $$

Clearly, $\tilde{p}_l = \log p_l$ for all $l$. Since $\mathbf{p}' \leq \tilde{p}$, we state the following lemma to show that the KKT conditions of (6) are equivalent to projecting (38) to $[-\infty, \log \tilde{p}]$ for all $l$.

Lemma 2 (Box-constrained projection) [26], pp. 520):

Consider the problem $\min_{z \leq \tilde{z}} f(z)$. The KKT conditions are equivalent to the condition $P_{\tilde{z} \leq z} \partial f(z) / \partial z = 0$, where $P_{\tilde{z} \leq z}$ is the projection of the vector $g$ onto the box $[\tilde{z}, z]$ defined by $P_{\tilde{z} \leq z} g_l = \min \{ 0, g_l \}$ if $z_l = \tilde{z}_l$, $(P_{\tilde{z} \leq z} g)_l = g_l$ if $z_l \leq \tilde{z}_l$, and $(P_{\tilde{z} \leq z} g)_l = \max \{ 0, g_l \}$ if $z_l > \tilde{z}_l$.

Applying Lemma 2 to (6) in the $\tilde{p}$ domain (let $z = -\infty, \tilde{z} = \log \tilde{p}$ in Lemma 2) and then converting it back to the $\mathbf{p}$ domain, we deduce

$$ \tilde{p}_l = \min \left\{ w_l \left( \frac{\sum_{j \neq l} w_j F_{ij} \tilde{p}_j + v_j}{\sum_{j \neq l} F_{ij} e^{\tilde{p}_j} + v_j} \right), \tilde{p}_l \right\}, $$

(39)

where $\tilde{p}_l$ is unique since $\sum_j w_j \log \text{SIR}(\tilde{p})$ is strictly convex in $\tilde{p}$ and the transformation $\tilde{p}_l = \log p_l$ is one-to-one.

Next, we establish the convergence of Algorithm SAPC. The convergence proof of (7) is based on the standard interference function approach [17], which is summarized as follows.

Definition 4 (Standard interference function): $I(\mathbf{p})$ is a standard interference function if it satisfies [17]:

1) (monotonicity) $I(\mathbf{p}') > I(\mathbf{p})$ if $\mathbf{p}' > \mathbf{p}$.
2) (scalability) Suppose $\beta > 1$. Then, $\beta I(\mathbf{p}) > I(\beta \mathbf{p})$.

Definition 5: A standard interference function $I(\mathbf{p})$ is feasible if $\mathbf{p} \geq I(\mathbf{p})$ has a feasible solution [17].

Lemma 3 (Theorem 2 and 4 in [17]): If $I(\mathbf{p})$ is feasible, then $\mathbf{p}(k+1) = I(\mathbf{p}(k))$ converges synchronously and totally asynchronously to the unique fixed point from any initial point $\mathbf{p}(0)$.

First, it can be easily verified that

$$ (I(\mathbf{p})); = w_l \left( \frac{\sum_{j \neq l} w_j F_{ij} \tilde{p}_j + v_j}{\sum_{j \neq l} F_{ij} e^{\tilde{p}_j} + v_j} \right) $$

(40)

is standard for all $l$. Furthermore, if $I(\mathbf{p})$ is standard, then $\min \{ I(\mathbf{p}));, \tilde{p}_l \}$ is standard (See Theorem 7 in [17]). Hence, the iterative equation in (7) is standard. Clearly, from the KKT conditions, $\mathbf{p} = I(\mathbf{p})$ has a feasible solution, thus
min(I(\(p\)), \(\tilde{p}\)) is feasible. Therefore, by Lemma 3, \(p(k+1)\) converges to the fixed point of (39) and hence the optimal solution of (6).

**B. Proof of Corollary 1**

Let \((I(p))^n\) be the power vector produced by \(n\) iterations of the standard interference function \(I(p)\) starting from an initial point \(\tilde{p}\). We first state the following lemma from [17].

**Lemma 4 (Lemma 1 in [17]):** If \(p \geq I(p)\), then \((I(p))^n\) is a monotone decreasing sequence of feasible power vectors that converges to a unique fixed point \(p^*\).

Starting from \(p(0) = \tilde{p}\), \(p(k)\) in Algorithm SACP is a decreasing sequence by Lemma 4, i.e., \(p(k+1) \leq p(k)\). This means \(p^* = p(k+1) \leq p(k) \leq \tilde{p}\). Thus, there exists a constant \(a \in [0,1]\) such that \(\|p(k+1) - p^*\| \leq a \|p(k) - p^*\|\), or equivalently, \(\|I(p(k)) - p^*\| \leq a \|p(k) - p^*\|\). This means \(I(p(k))\) is a pseudo-contraction mapping, which guarantees a geometric convergence rate to the unique fixed point \(p^*\). Combining with Theorem 1, Corollary 1 is proved.

**C. Proof of Theorem 2**

Introducing an auxiliary variable \(\tau\) to (8), we rewrite (8) as

\[
\text{maximize} \quad \tau
\]
\[
\text{subject to} \quad \tau \leq \frac{\text{SIR}_i(p)}{\beta_i} \quad \forall i, \quad p \leq \tilde{p}, \quad \tau > 0. \tag{41}
\]

Making a change of variables: \(\tilde{\tau} = \log \tau\) and \(\tilde{p}_i = \log p_i\) for all \(i\), (41) is equivalent to the following convex problem:

\[
\text{maximize} \quad \tilde{\tau}
\]
\[
\text{subject to} \quad \tilde{\tau} \leq \log (\text{SIR}_i(\tilde{p})/\beta_i) \quad \forall i, \quad \tilde{p} \leq \tilde{p}_i \quad \forall i. \tag{42}
\]

Next, introducing the dual variables \(\lambda\), the partial Lagrangian of (42) is given by

\[
L(\{\tilde{\tau}, \tilde{p}\}, \{\lambda\}) = \tilde{\tau} (1 - \sum \lambda_i) + \sum \lambda_i \log (\text{SIR}_i(\tilde{p})/\beta_i). \tag{43}
\]

In order for (43) to be bounded, we must have \(\sum \lambda_i = 1\). Hence, the dual problem of (42) is given by

\[
\text{maximize} \quad \sum \lambda_i \log (\text{SIR}_i(\tilde{p}(\lambda))/\beta_i).
\]
\[
\text{subject to} \quad 1^T \lambda = 1, \quad \lambda > 0. \tag{44}
\]

At optimality, the optimal objective to (42), i.e., \(\tilde{\tau}^*\), is equal to the optimal dual function of (44) given by

\[
\tilde{\tau}^* = \sum \lambda_i \log (\text{SIR}_i(p(\lambda))/\beta_i). \tag{45}
\]

Next, we compute an upper bound to \(\tilde{\tau}\) in (41). For any feasible \(\lambda\) in (45), we have

\[
\tilde{\tau} \leq \tilde{\tau}^* \leq \sum \lambda_i \log (\text{SIR}_i(p(\lambda))/\beta_i). \tag{47}
\]

First, we prove the following result:

**Lemma 5:** For any power vector \(p\),

\[
\sum x_i y_i \log (\text{SIR}_i(p)/\beta_i) \leq -\log \rho(\text{diag}(\beta)(F + (1/\tilde{p}_i)ve_i^T)), \tag{48}
\]

where \(i\) is defined in (10). Equality is achieved in (48) if and only if \(p = \alpha x(\text{diag}(\beta)(F + (1/\tilde{p}_i)ve_i^T))\), where \(\alpha = \tilde{p}_i / x_i\). In this case, \(p = \tilde{p}_i\).

First, let \(i\) be given as in (10), and we state the following key result in [27].

**Theorem 10 (Theorem 3.1 in [27]):** For any irreducible nonnegative matrix \(A\),

\[
\prod_i ((Az)_i/z_i)^{x_i/y_i} \geq \rho(A) \tag{49}
\]

for all strictly positive \(z\), where \(x\) and \(y\) are the Perron and left eigenvectors of \(A\) respectively. Equality holds in (49) if and only if \(z = \alpha x\) for some positive \(\alpha\).

Theorem 10 leads to the following result.

**Lemma 6:** If \(z = x\) in (49), then \((Az)_i/z_i = (Az)_{i}/z_i \forall j \neq i\).

Since \(\text{diag}(\beta)(F + (1/\tilde{p}_i)ve_i^T)\) is irreducible, we let \(x\) and \(y\) be the Perron and left eigenvectors of \(\text{diag}(\beta)(F + (1/\tilde{p}_i)ve_i^T)\) respectively, and write, for all \(p\) in (6),

\[
\prod_i ((Az)_i/z_i)^{x_i/y_i} = \prod_i \left( \frac{p_i}{(\text{diag}(\beta)(F + (1/\tilde{p}_i)ve_i^T))} \right)^{x_i/y_i} \leq \frac{1}{\rho(\text{diag}(\beta)(F + (1/\tilde{p}_i)ve_i^T))}, \tag{50}
\]

where, in (50), (a) is due to \(p_i \leq \tilde{p}_i\) and (b) is due to letting \(A = \text{diag}(\beta)(F + (1/\tilde{p}_i)ve_i^T)\) in (49). Thus, we deduce (48).

Lemma 5 is proved by showing that the inequality in (48) becomes an equality if and only if \(p = x(\text{diag}(\beta)(F + (1/\tilde{p}_i)ve_i^T))\) (unique up to a scaling constant) as follows. First, we note that (b) in (50) becomes an equality if and only if \(p = x(\text{diag}(\beta)(F + (1/\tilde{p}_i)ve_i^T))\) (unique up to a scaling constant) using Theorem 10. Equality of (a) in (50) follows from the fact that for the max-min SIR problem, \(p_i = \tilde{p}_i\) and thus \(p_i/\tilde{p}_i = 1\).

Letting \(\lambda = x \circ y\) in (47) (which is clearly feasible in (44)), we deduce that

\[
\tilde{\tau}^* \leq \sum x_i y_i \log (\text{SIR}_i(p(x \circ y))/\beta_i) \leq -\log \rho(\text{diag}(\beta)(F + (1/\tilde{p}_i)ve_i^T)), \tag{51}
\]

where (a) in (51) is due to Lemma 5, thus proving that \(\tau \leq 1/\rho(\text{diag}(\beta)(F + (1/\tilde{p}_i)ve_i^T))\). Further, from Lemma 5, equality of (a) in (51) is achieved if and only if \(p = ax\) for some positive \(a\) (such that \(p_i = \tilde{p}_i\)). From Lemma 6, this implies \(\text{SIR}_i(ax)/\beta_i = \text{SIR}_i(ax)/\beta_j \forall j \neq i\). This means that a feasible power allocation \(p = ax\) for some positive \(a\) solves (8) optimally. Hence, we deduce that \(\tau^* = 1/\rho(\text{diag}(\beta)(F + (1/\tilde{p}_i)ve_i^T))\). This completes the proof of Theorem 2.
D. Proof of Theorem 3

Algorithm 2 is based on the subgradient method that maximizes the partial Lagrangian given by (43) of (42) using Lagrange dual decomposition (cf. Proof of Theorem 2). Since the duality gap is zero, solving the Lagrange dual problem is equivalent to solving (42) (thereby solving (8)).

First, we compute a subgradient of $\text{max}_{\lambda, \bar{\lambda}} L_i \{ \hat{\lambda} \}$ for any feasible $\lambda$ in (44). Observe the following chain of inequalities in (52). In (52), (a) is due to choosing a feasible $\hat{\lambda}$ and

$$\hat{\lambda} \equiv \arg \max_{\lambda \leq \bar{\lambda}} L_i \{ \hat{\lambda} \}$$

such that

$$\log(S\text{IR}_{\lambda}(\hat{\lambda}))/\beta_\lambda \geq \hat{\tau}$$

for a feasible $\lambda$, (b) is due to $\sum \lambda_i \log(S\text{IR}_{\lambda}(\hat{\lambda}))/\beta_\lambda \geq \hat{\tau}$, which is implied by (54), (c) is due to the definition of $\hat{\lambda}$ in (53) and, in (d), we have

$$\log(L_i \{ \hat{\lambda} \}) = \sum \lambda_i \log(S\text{IR}_{\lambda}(\hat{\lambda}))/\beta_\lambda - \log(S\text{IR}_{\lambda}(\hat{\lambda}))/\beta_\lambda.$$ 

We thus deduce that $-g$ is a subgradient of $\log(S\text{IR}_{\lambda}(\hat{\lambda}))/\beta_\lambda$ at $\hat{\lambda}$. The subgradient method generates a sequence of dual feasible points according to the iteration

$$\lambda(t + 1) = \max\{\lambda(t) + \alpha(t)g(k(t)), 0\} \forall t,$$

where $g(t)$ is the subgradient at $\lambda(t)$ and $\alpha(t)$ is a positive scalar step-size. For a given feasible $\lambda$, the Lagrange dual function is evaluated by $p(\lambda(t))$ or equivalently $p(\lambda(t))$, which is simply equivalent to solving (6) by letting the weight vector $w = \lambda(t)$. There are many standard convergence results for the subgradient method, depending on the choice of step-size used [19]. For example, if $\alpha(t) = \epsilon_0$, then (56) converges to within some range of the optimal value, i.e., $L_i\{\hat{\lambda}(t), \hat{p}(\hat{\lambda})\} - \hat{\tau} < \epsilon_0$ for a small positive $\epsilon_0$. If $\alpha(t) \rightarrow \infty$ satisfies $\lim_{k \rightarrow \infty} \alpha(t) = 0$ and $\sum_{k=1}^\infty \alpha(t) = \infty$, then $\lambda(t)$ is a unique solution to (53).

Lemma 7 (Proposition 6.3.1 in [19]): If $\lambda(t)$ is not optimal, then for every dual optimal solution $\lambda^*$, we have $\|\lambda(k + 1) - \lambda^*(t)\| < \|\lambda(t) - \lambda^*(t)\|$.

Now, since the optimal Lagrange dual function in (44) is $\hat{\tau} = -\log(\rho(\text{diag}(\beta)(F + 1/\bar{p})v_\gamma))$, we apply Lemma 7 to our problem to deduce that if $\alpha(t)$ in Algorithm 2 is strictly less than (11) (let $w(t) = \lambda(t)$ in (11)) where $g(t)$ is the subgradient function given by (55) at $\lambda(t)$, then Algorithm 2 converges linearly. Taken together with Corollary 1 (geometric convergence of $p(\lambda(t))$, this implies the overall geometric convergence of $p(k)$ to the solution of (8).

E. Proof of Corollary 2

Using Lemma 5 in the proof of Theorem 2, we show that $\lambda_k^* = x_ky_k$ for all $k$. From (45),

$$\lambda_k^* \log(S\text{IR}_{\lambda_k^*}(p(\lambda_k^*))) = \hat{\tau} = \log\rho(\text{diag}(\beta)(F + 1/\bar{p})v_\gamma))$$

Now, by the Perron-Frobenius theorem, $x$ and $y$ are unique up to a scaling constant. Also, $\lambda_k^*$ is unique by strict convexity of the constraint set in (42). Combining (58) and the equality case in (48), we have $\lambda_k^* = x \circ y$. Considering SAPC in (6) with $w = x \circ y$ thus yields the solution $p = x(\text{diag}(\beta)(F + 1/\bar{p})v_\gamma))$ (unique up to a scaling constant).

F. Proof of Theorem 4

Our proof is based on the nonlinear Perron-Frobenius theory [28]. We let the optimal max-min weighted $S\text{IR}(p^*)$ in (8) be $\tau^*$. A key observation is that all the $S\text{IR}$ constraints are tight at optimality. This implies, at optimality of (8),

$$\frac{\sum_{j \neq i} F_{ij}(\lambda^*)}{\sum_{j \neq i} F_{ji}(\lambda^*)} = \tau^*$$

for all $i$. Letting $s^* = (1/\bar{p})p^*$, (59) can be rewritten as

$$1/\tau^* s^* = \text{diag}(\beta)F s^* + (1/\bar{p})\text{diag}(\beta)v.$$

We first state the following result:

**Lemma 8 (Conditional eigenvalue [28], Corollary 14):** Let $A$, $B$ be a nonnegative matrix and $b$ be a nonnegative vector. If $\rho(A + be_i) > \rho(A)$, where $i = \arg\min_1 \rho(\lambda + be_i)$, then the conditional eigenvalue problem $\lambda = \text{As} + b$, $\lambda \in \mathbb{R}$, $s \geq 0$, $\max s_i = 1$, has a unique solution given by $\lambda = \rho(A + be_i)$ and $s$ being the unique normalized Perron eigenvector of $A + be_i$.

Letting $\lambda = 1/\tau^*$, $A = \text{diag}(\beta)F$, $b = (1/\bar{p})\text{diag}(\beta)v$ in Lemma 8 and noting that $\max_{\lambda, s}_i s^* = 1$ shows that $p^* = (\lambda_1/\bar{p}) \text{diag}(\beta)F + (1/\bar{p})\text{diag}(\beta)v$ is a fixed point of (60) as it should be (cf. Theorem 2). Now, the fixed point in Lemma 8 is also a unique fixed point of the following iteration [28]:

$$s(k + 1) = \frac{As(k) + b}{\|As(k) + b\|_\infty}.$$

Applying the power method in (61) to the system of equations in (60) and letting $p(k + 1) = s(k + 1)p$ yield Algorithm 3. This completes the proof of Theorem 4.

G. Proof of Theorem 5

Our proof is to find an upper bound to (16), which in turn upper bounds (1). Let $x = x(\mathbf{B})$ and $y = y(\mathbf{B})$. Using $z = (I + \mathbf{B})p$, we have the following chain of inequalities:

$$\sum_{i} x_i y_i \log(1 + S\text{IR}(p)) = \sum_{i} x_i y_i \log(z_i/\|\mathbf{B}z_i\|) \leq \log \rho(\mathbf{B}) \leq \log(1 + 1/\rho(\mathbf{B})),$$
where (a) is due to letting $A = \hat{B}$ in (49), and using the fact that $x(B) = x(B)$ and $y(B) = y(B)$, and (b) is due to $\rho(\hat{B}) = \rho(\hat{B}) / (1 + \rho(\hat{B}))$ [22].

We give another proof to establish (a) using the following result in [27].

**Theorem 11** ([27]): Let $\gamma \geq 0$ and let $A \in \mathbb{R}^{L \times L}_{+}$ be an irreducible matrix. Then,

$$\prod_{i} \gamma_{\hat{i}}(A) \rho(A) \leq \rho(\text{diag}(\gamma) A).$$

Equality holds for the lefthand-side if and only if $\gamma_{i}$ is equal for all $l$.

It is shown in [22] that the optimal solution to (16) can be determined by solving the following system of equations:

$$z_{l}^{\ast} = \frac{u_{l}}{\sum_{j} w_{j} B_{l} / (B z_{l}^{\ast})}$$

for all $l$ and satisfies $1^{T} z^{\ast} - 1^{T} \hat{B} z^{\ast} = \hat{p}$. By making a change of variables $q_{l} = u_{l} / z_{l}^{\ast} > 0$ for all $l$ and using the fact that $z_{l}^{\ast} / (B z_{l}^{\ast}) = 1 / \gamma_{l}$, (64) can be rewritten as $\hat{B} \text{diag}(1 + \gamma) q = q$. By the Perron Frobenius Theorem, this implies $\rho(\hat{B}) \text{diag}(1 + \gamma) \gamma = \rho(\text{diag}(1 + \gamma) \hat{B}) = 1$. Using (63) in Theorem 11, we have

$$\prod_{l} (1 + \gamma_{l}) (y^{\times y}) \rho(\hat{B}) \leq \rho(\text{diag}(1 + \gamma) \hat{B}) = 1.$$ (65)

Since $\rho(\hat{B}) = \rho(\hat{B})$, by taking the logarithm on both sides of (65), this establishes (a) in (62).

Next, in (49), if $z \geq A z$, then for any positive vector $w$, (49) can be extended to

$$\prod_{l} (\frac{(A z_{l})}{z_{l}^{\ast}})^{w_{l}} \geq (\rho(A))^{\|w\|_{\infty}^{2}}.$$ (66)

Clearly, (66) reduces to (49) when $w = x \circ y$. (66) can be deduced by noting that $\|w\|_{\infty_{x}}^{2} \leq y_{l} / w_{l}$ and thus

$$\prod_{l} (\frac{(A z_{l})}{z_{l}^{\ast}})^{w_{l}} \geq \prod_{l} (\frac{(A z_{l})}{z_{l}^{\ast}})^{x_{l} / y_{l}} (\|w\|_{\infty}^{2}) \geq (\rho(A))^{\|w\|_{\infty}^{2}}.$$ (67)

Now, we apply (66) to (62) by letting $A = \hat{B}$ to deduce

$$\sum_{l} w_{l} \log(z_{l} / (\hat{B} z_{l}^{\ast})) \leq -\|w\|_{\infty_{x}}^{2} \log \rho(\hat{B}) = \|w\|_{\infty_{x}}^{2} \log(1 + 1 / \rho(\hat{B})).$$ (68)

Clearly, equality holds in (68) if and only if $w = x(B) \circ y(B)$ by considering $z = x(B)$ and using (49).

Next, we prove the condition for equality in (5). We first prove the inequality $\rho(B) \geq \rho(F + (1 / \hat{p}) v e_{i})$ for $x(B)$ to be feasible in (8).

**Lemma 9**: The vector $x(B)$ is feasible in (8) with $\beta = 1$ if and only if

$$\rho(B) \geq \rho(F + (1 / \hat{p}) v e_{i})$$ (69)

for all $l$. In this case, the optimal value of (8) is $1 / \rho(B)$.

To prove (69), we use the following result.

**Lemma 10** (Lemma 4 in [28]): Let $A \geq 0$ and $b, c$ be nonnegative vectors. If $\rho(A + b c^{\top}) \geq \rho(A)$, then for any nonnegative vector $d$,

$$\text{sign} \rho(A + b c^{\top}) = \text{sign}(c^{\top} x - d^{\top} x).$$ (70)


By solving (8) with $\beta = w$, the $l$th user achieves the SIR value $w_{l} / \rho(\text{diag}(w))(F + (1 / \hat{p}) v e_{i})$, where $i$ is given in (10). Substituting this quantity into the objective function of (1), the sum rate is thus given by $\sum_{l} w_{l} \log(1 + w_{l} / \rho(\text{diag}(w))(F + (1 / \hat{p}) v e_{i})).$

**H. Proof of Theorem 6**

By solving (8) with $\beta = w$, the $l$th user achieves the SIR value $w_{l} / \rho(\text{diag}(w))(F + (1 / \hat{p}) v e_{i})$, where $i$ is given in (10). Substituting this quantity into the objective function of (1), the sum rate is thus given by $\sum_{l} w_{l} \log(1 + w_{l} / \rho(\text{diag}(w))(F + (1 / \hat{p}) v e_{i})).$
Now, suppose $\tilde{B} \geq 0$. Then by using the upper bound to the optimal objective in Theorem 5, the quantity
\[
\eta = \frac{\sum_{i} u_{i} \log(1 + w_{i} / \rho \text{diag}(w)(F + (1/\tilde{p}_{i})v_{i}))}{\|w\|_{\infty}(B)^{y}} \log(1 + 1/\rho(B)),
\]
(71)
is an approximation ratio to (1) by solving (8) with $\beta = w$.

I. Proof of Theorem 7

Theorem 7 is proved similarly to Theorem 6, where the optimal solution $p^*$ to (6) is used instead. Hence, the quantity
\[
\eta = \frac{\sum_{i} u_{i} \log(1 + \text{SI}(p^*))}{\|w\|_{\infty}(B)^{y(B)}} \log(1 + 1/\rho(B)),
\]
(72)
is an approximation ratio to (1) by solving (6).

J. Proof of Theorem 8

(27) is obtained by substituting $\gamma_{l} = e^{r_{l}} - 1$ for all $l$ in the constraint set of (24). Time-sharing operation then convexifies this nonconvex constraint set to yield a convex achievable rate region $\text{Conv}(\mathcal{R})$.

K. Proof of Corollary 3

The first part of Corollary 3 given by (29) is proved easily by substituting $\gamma_{l} = e^{r_{l}} - 1, l = 1, 2$, in the constraint set of (28). For the second part, we let $\tilde{p}_{i} \to \infty$ in (29) and use L'Hôpital's rule to reduce both terms on the right-hand side in (29) to $\log(1 + 1/(F_{12}^{2}F_{21}^{2}e^{r_{l}} - 1))$, which bounds the achievable rate for the first user.

L. Proof of Theorem 9

We will prove Theorem 9 using Theorem 6 in Section IV and Proposition 3 in [13] (with extension to time-varying capacity found in Section IV-B of [13] and similar results are also found in [12]). Our proof approach follows the argument in Proposition 3 in [13]. The key to proving Proposition 3 in [13] is to define the Lyapunov function (omitting the time slot index $kT$) $V(n, q) = V_{n}(n) + V_{q}(q)$, where $V_{n}(n) = \sum_{l = 1}^{L} \phi_{n} n_{l}^{2}$ and $V_{q}(q) = \sum_{l = 1}^{L} \phi_{q} n_{l}^{2}$, and show that $V(n, q)$ has a negative drift for an appropriately chosen $\alpha_{l}$ for all $l$.

First, we note that $R(n)$ in (27) and hence $\text{Co}(\mathcal{R})$ is closed and bounded for all time slots. Following [13], we can show that
\[
\mathbb{E}[V(n((k+1)T), q((k+1)T)) - V(n(kT), q(kT))|n(kT), q(kT)] \\

to T \sum_{l=1}^{L} q_{l} \eta_{l} + E_{1}(k) + E_{2}(k),
\]
where $E_{1}(k)$ and $E_{2}(k)$ are error terms roughly on the order of $|p_{k} - n_{k}(kT)|x_{k}(kT)|$ and $|\sum_{l=1}^{L} H_{l} n_{l}(kT)x_{l}(kT) - r_{l}(kT)|^{2}$, respectively.

Now, suppose $\rho$ lies strictly inside $\eta(kT)\Lambda$, then there exists some $\epsilon > 0$ such that
\[
(1 + \epsilon) \sum_{s=1}^{S} H_{s} \rho_{s} \in \eta(kT)\Lambda,
\]
provided $\lim_{k \to \infty} \eta(kT)$ is bounded away from zero when the maximum entry of $q(kT)$ is very large. This is the case in [12] when $\lim_{k \to \infty} \eta(kT)$ is a constant independent of $k$.

We now show that by solving a max-min weighted $\text{SI}$R problem ($q(kT)$ as weights) using Algorithm 3 to determine $r(kT)$, we obtain $r(kT) = \log(1 + q_{i}(kT)/\rho(\text{diag}(q(kT))) (F + (1/\tilde{p}_{i})v_{i}))$, where $i = \arg\max_{i} F + (1/\tilde{p}_{i})v_{i}$. Then, $\eta(kT)$ is given by (18) of Theorem 18 (set $w = q(kT)$). Further, this $\eta(kT)$ is bounded away from zero when $|q(kT)|$ is large. More precisely, we shall find a lower bound to $\eta(kT)$ as $\max_{i} q_{i}(kT)$ (let $m = \arg\max_{i} q_{i}(kT)$) tends to a large value. Let $i = \arg\max_{i} F + (1/\tilde{p}_{i})v_{i}$. Observe the following chain of inequalities (omitting the time slot index $kT$):
\[
\lim_{\eta_{m} \to \infty} \eta = \lim_{\eta_{m} \to \infty} \sum_{j} \frac{q_{j} / \eta}{1 - \rho(\text{diag}(q) (F + (1/\tilde{p}_{i})v_{i}))} \log(1 + 1/\rho(B)) \\
\geq \lim_{\eta_{m} \to \infty} \sum_{j} \frac{q_{j} / \eta}{1 - \rho(\text{diag}(q) (F + (1/\tilde{p}_{i})v_{i}))} \log(1 + 1/\rho(B)) \\
= \left(\lim_{\eta_{m} \to \infty} \min_{l} (x_{B}(x_{B}) \circ y_{B})_{l} \frac{1}{q_{j}} \eta \log(1 + 1/\rho(B)) \right) \\
\lim_{\eta_{m} \to \infty} \sum_{j} \frac{q_{j}}{1 - \rho(\text{diag}(q) (F + (1/\tilde{p}_{i})v_{i}))} \log(1 + 1/\rho(B)) \\
\geq \left(\lim_{\eta_{m} \to \infty} \min_{l} (x_{B}(x_{B}) \circ y_{B})_{l} \frac{1}{q_{j}} \eta \log(1 + 1/\rho(B)) \right) \\
\lim_{\eta_{m} \to \infty} \sum_{j} \frac{q_{j}}{1 - \rho(\text{diag}(q) (F + (1/\tilde{p}_{i})v_{i}))} \log(1 + 1/\rho(B))
\]
(75)
where (a) is due to the fact that the index $j^{k}$ is the optimal index of the max-min $\text{SI}$R problem (since we consider only $q_{i}(kT) > 0$) at time slot $kT$, (b) is due to the inequality $\rho(\text{diag}(z)A) \leq \max_{i} z_{i} \rho(A)$ for any nonnegative and irreducible square matrix $A$ (e.g., see [27]), and (c) is due to the L'Hôpital's rule.

Now, assume that $\rho$ lies strictly inside
\[
\min_{l} (x_{B}(x_{B}) \circ y_{B})_{l} \frac{1}{\log(1 + 1/\rho(B))} \log(1 + 1/\rho(F + (1/\tilde{p}_{i})v_{i})) \Gamma,
\]
(76)
then there exists some $\epsilon > 0$ such that
\[
(1 + \epsilon) \sum_{l=1}^{L} H_{l} \rho_{s} \in \min_{l} (x_{B}(x_{B}) \circ y_{B})_{l} \frac{1}{\log(1 + 1/\rho(B))} \log(1 + 1/\rho(F + (1/\tilde{p}_{i})v_{i})) \text{Co}(\mathcal{R}).
\]
(77)
Substituting (77) into (73), we have
\[
\mathbb{E}[V(n((k+1)T), q((k+1)T)) - V(n(kT), q(kT))|n(kT), q(kT)] \\
\leq -T \sum_{l=1}^{L} q_{l}(kT) \sum_{l=1}^{L} H_{l} \rho_{s} + E_{1}(k) + E_{2}(k),
\]
(78)
thus showing that $V(\mathbf{n}(kT), \mathbf{q}(kT))$ drifts to zero when queue sizes are large and when the error terms $E_1(k)$ and $E_2(k)$ are bounded [13]. This concludes the proof of Theorem 9.

REFERENCES


