Distributed Non-Autonomous Power Control through Distributed Convex Optimization

S. Sundhar Ram and V. V. Veeravalli
ECE Department and Coordinated Science Lab
University of Illinois at Urbana-Champaign
Email: {ssriniv5,vvv}@illinois.edu

A. Nedić
Department of Industrial and Enterprise Systems Eng.
University of Illinois at Urbana-Champaign
Email: angelia@illinois.edu

***Abstract*— We consider the uplink power control problem where mobile users in different cells are communicating with their base stations. We formulate the power control problem as the minimization of a sum of convex functions. Each component function depends on the channel coefficients from all the mobile users to a specific base station and is assumed to be known only to that base station (only CSIR). We then view the power control problem as a distributed optimization problem that is to be solved by the base stations and propose convergent, distributed and iterative power control algorithms. These algorithms require each base station to communicate with the base stations in its neighboring cells in each iteration and are hence non-autonomous. Since the base stations are connected through a wired backbone the communication overhead is not an issue. The convergence of the algorithms is shown theoretically and also verified through numerical simulations.**

I. Introduction

The multi-access nature of the wireless channel enables multiple transmitters to communicate with their receivers simultaneously using the same channel. This presents a classic resource allocation problem. Each transmitter would prefer to increase its transmission power to improve the signal at its receiver. The higher the signal power, the higher is the interference caused by a transmission at the other receiver locations. The power control problem is to assign power levels to the transmissions so that the signal strengths at *all* the receivers are satisfactory.

We will study the inter-cell uplink power control problem in cellular networks. In this setting, a mobile user (MU) from each cell communicates with its base station (BS) on a common channel. The optimal power allocation will depend on the channel between each MU and each BS. In general, the channel is estimated at the base station using pilot signals transmitted by the MUs. Therefore, a BS has information about the channel between the MUs and itself, but has no information about the channel between the MUs and other BSs. Thus, distributed power control algorithms in which each BS uses only locally available information to determine the optimal powers are of interest.

There are different mathematical formulations for the power control problem. We refer the reader to [1] for an extensive survey and we only summarize some points of interest. The first formulation that was studied was to minimize the total power subject to a constraint on the signal to interference and noise ratio (SINR) at each base station [2]. The power control problem was viewed as a closed loop control problem and distributed control algorithms were proposed. In these, each MU iteratively refines its power. In each iteration, all the MUs transmit at the current iterate value. The SINR is measured at the BS and is communicated to the MU. Based on the SINR value the old iterate is refined to obtain the new iterate. A general framework was later defined in [3] to study such control theoretic solutions. The second approach is to view the power control as a game between non-cooperative users [4], [5]. In this, a utility function is defined for each MU-BS pair that captures their satisfaction as a function of the SINR at the BS and the power used by the MS. The third approach is to view the power control problem as an open loop global optimization problem [6], [7].

All the studies mentioned above do not require any message passing between the MUs and only require communication between each BS and MU. Thus, they are not just distributed but also *autonomous*. In this paper, we view the power control problem as a deterministic convex optimization problem that is to solved by the base stations through information sharing with neighbors. As we will see, the optimization problem has a special structure where the objective function can be written as the sum of convex functions, each of which is completely known to a BS. Thus the power control problem is a special case of the distributed optimization problem studied in [8], [9], [10]. These algorithms will require the BSs to iteratively exchange information with their immediate neighbors to solve this optimization problem. Since the base stations are connected through a wired backbone this communication overhead may not be an issue. Once the optimal power value (or, a sufficiently good approximation) is determined each BS communicates the optimal power value to the MU in its cell. Thus these algorithms are non-autonomous as the base stations collaborate in solving the problem.

The rest of the paper is organized as follows. In Section II, we introduce the notation and mathematically formulate the power control problem. In Section III, we discuss the distributed power control algorithms and state convergence results from [10], [8], [9]. In Section IV, we perform numerical simulations to test the performance of the algorithms. We finally conclude in Section V.

This research was supported by a Vodafone Graduate Fellowship, NSF grant CMMI 07-42538 and NSF Award CNS-0831670, through the University of Illinois.

II. Problem Formulation

We will consider a finite cellular network. There are m mobile users (MU) in neighboring cells communicating with their respective base station using a common wireless channel. Any other interfering source is treated as noise. Some of the MUs may be communicating with the same physical base station (BS). However, we will find it convenient to assign a unique index to each MU's receiver. Specifically, we use BS i to denote the base station with which MU i is communicates.

We assume that the channel is static and the number of users do not change over the time scales studied in this paper. We denote the channel coefficient between MU j and BS i by $h_{i,j}$. It includes both the effects of large scale and small scale variations. We will assume that BS i has a good estimate of $h_{i,j}$, for all j. MUs that are in cells that are not immediate neighbors cause negligible interference. Therefore, they can be taken to be 0. For MUs in cell i and in the immediate neighborhood, BS i can estimate the channel coefficient from pilot signals that are sent by the mobiles.

Let p_i denote the power used by MU i and σ_i^2 be the receiver noise variance. Define $\bar{p}$ to be the vector with the j-th component equal to p_j, and $\bar{h}_i$ to be the vector with the j-th component equal to $\bar{h}_{i,j}$. The total received SINR at BS i is given by

$$\gamma_i\left(\bar{p}, \bar{h}_i\right) = \frac{p_i h_{i,i}^2}{\sigma_i^2 + \sum_{j\neq i} p_j h_{i,j}^2}.$$

Let U_i denote the utility function that captures the satisfaction of BS i as a function of its received SINR.[1] Depending on the nature of traffic (voice, multimedia or data) between MU i and BS i, the form of the function U_i could be different. The power control problem is to operate at the optimal point on the utility versus power curve. Formally, we have the following optimization problem:

$$\max_{\bar{p}} \sum_i U_i\left(\gamma_i\left(\bar{p}, \bar{h}_i\right)\right) - \sum_i V\left(p_i\right)$$
$$\text{subject to} \quad 0 \leq p_i \leq p_t, \quad \forall i. \tag{1}$$

Here V is a convex and increasing, and captures the cost of the power and p_t is a threshold on the maximum power that a MU can use.

In general there are no closed form solutions for the optimization problem in (1) and iterative algorithms have been used. When the problem is non-convex, the iterative algorithms may converge to a local maximum, rather than a global maximum. To avoid this, we impose additional restrictions on the utility functions U_i resulting in convex problem (1). These functions have the following property

$$-\frac{xU_i''(x)}{U_i'(x)} \geq 1, \qquad \forall x \in X_i,$$

where X_i is some convex constraint set (see page 53 of [1]). We will focus on the case when the function $U_i(x) = \log(x)$, although nothing prevents the algorithms developed in this paper to be used for other utility functions such as the α-fair utility [1]. There is a natural motivation for the log utility function. First, it is the standard proportional fairness function used in literature [6], [11]. Second, a common choice for a mobile user utility function is the channel throughput achieved by the user. For each mobile user i, the throughput of the user is modeled as (see [12], [13])

$$T_i(\bar{p}) = \log(1 + \eta\gamma_i(\bar{p})),$$

where η is a constant determined by the modulation scheme that is used. The expression for $T_i(\bar{p})$ is a very good approximation for both additive white Gaussian channels and Rayleigh fading environments. When $\eta = 1$ this is also Shanon's capacity. As such the throughput is a non-convex function of the powers. A commonly used technique to deal with the non-convexity [12], [14], [13] is to approximate

$$T_i(\bar{p}) \approx \log(\eta\gamma_i(\bar{p})).$$

This approximation is valid in the high SINR regime. In summary, the problem that is of interest is

$$\max_{\bar{p}} \sum_i \left[\log\left(\frac{p_i h_{i,i}^2}{\sigma_i^2 + \sum_{j\neq i} p_j h_{i,j}^2}\right) - V\left(p_i\right)\right]$$
$$\text{subject to} \quad 0 \leq p_i \leq p_t, \quad \forall i. \tag{2}$$

Using the substitution $p_i = e^{x_i}$ in (2) we can rewrite the optimization problem as

$$\min_x \sum_{i=1}^m \left[\log\left(\sigma_i^2 h_{i,i}^{-2} e^{-x_i} + \sum_{j\neq i} h_{i,i}^{-2} h_{j,i}^2 e^{x_j - x_i}\right) + V(e^{x_i})\right]$$
$$\text{subject to} \quad x \in X. \tag{3}$$

Here x is the vector with the i-th component equal to x_i and X is the set $\{x : x_i \leq \log(p_t) \ \ \forall i\}$. The constraint set X is convex. Furthermore, since the log of a sum of exponentials is convex, the objective function is convex. Thus the problem in (3) is a convex optimization problem. Define

$$f_i(x; \bar{h}_i) = \log\left(\sigma_i^2 h_{i,i}^{-1} e^{-x_i} + \sum_{j\neq i} h_{i,i}^{-1} h_{j,i} e^{x_j - x_i}\right) + V(e^{x_i}).$$

Then, the problem in (3) can be written as

$$\min_x \sum_{i=1}^m f_i(x; \bar{h}_i)$$
$$\text{subject to} \quad x \in X. \tag{4}$$

Our assumption that the channel coefficients $h_{i,j}$ between users j and base station i are known only to base station i translates to the function $f_i(x; \bar{h}_i)$ being known only to base station i. Therefore, the problem in (3) has to be solved in a distributed manner by the base stations, which communicate locally over the wired backbone that connects them.

[1] Typically, U_i is assumed to be an increasing function of the SINR.

The problem in (4) is a special case of a general distributed optimization problem studied in [8], [10], [15], which has the following form

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{m} f_i(x) \\ \text{subject to} \quad & x \in X. \end{aligned} \tag{5}$$

The function f_i is known only to agent i and the agents are connected through a peer-to-peer network. In the context of the power control problem, the agents are the base stations and the peer-to-peer network is the wired backbone that connects the base stations. We next describe the algorithms considered in [8], [10], [15]. Let x_k denote the iterate vector at the end of the k-th iteration. In the $(k+1)$-st iteration, the base station (BS) having the iterate x_k communicates the iterate to a neighbor. If $s(k+1)$ is the index of the BS that receives the iterate, then the new iterate is evaluated *incrementally* according to

$$x_{k+1} = \mathcal{P}_X\left[x_k - \alpha_{k+1}\nabla f_{s(k+1)}(x_k;\bar{h}_{s(k+1)})\right]. \tag{6}$$

Here α_{k+1} is the stepsize, ∇ denotes the gradient, and P_X denotes projection onto the set X. In the cyclical incremental gradient algorithm, the index $s(k+1)$ is given by $s(k+1) = \bmod\ (k+1, m)$, i.e., the iterates are cycled and passed from one BS to the next sequentially [10]. In the Markov incremental gradient algorithm, the BS having the iterate x_k randomly selects a neighbor that will update next [15]. By a neighbor of BS i, we mean a BS to which BS i is connected directly in the wired backbone. The convergence of the algorithm for the case when the network topology changes with time (thus, the neighbors change with time) is studied in [8].

The incremental algorithms can be used by the network of base stations to solve the power control problem in (4). An important criterion that decides the effectiveness of distributed power control algorithms is the time to convergence. If the time to convergence is slow, the channel and the number of users may change by the time the optimal power is determined. In the incremental algorithms, the base stations sequentially process the iterates. A possible way to improve the convergence time would be to use a parallel and distributed algorithm. In the next section, we propose one such optimization algorithm to solve (5) over a general, possibly time-varying, network topology and discuss its convergence.

III. Parallel and Distributed Optimization Algorithms

There are m agents that are indexed using the set $V = \{1,\ldots,m\}$ in an arbitrary manner. The network objective is to

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{m} f_i(x) \\ \text{subject to} \quad & x \in X. \end{aligned} \tag{7}$$

Here, x is a parameter vector of dimensions n, X is a parameter set in $\Re^n$ and f_i is a function from $\Re^n$ to $\Re$ that is known only to agent i. We define

$$f(x) = \sum_{i=1}^{m} f_i(x), \qquad f^* = \min_{x\in X} f(x),$$
$$X^* = \{x \in X : f(x) = f^*\}.$$

We make the following basic assumptions on the constraint set X and the functions f_i.

Assumption 1: The set $X \subseteq \Re^n$ is closed, convex and bounded. For each $i \in V$, the function $f_i : \Re^n \to \Re$ is convex.

Since each function f_i is convex over $\Re^n$, each f_i is continuous (see Theorem 10.1 in [16], or Proposition 1.4.6 in [17]). As an immediate consequence of this and the fact that the constraint set X is compact, we have that the optimal value f^* is finite and the set X^* of optimal solutions is nonempty. Note that we are making no assumptions on the differentiability of f_i. We define ∇f_i as a subgradient of the function f_i, which satisfies the following relation

$$\nabla f_i(x)^T(y-x) \le f_i(y) - f_i(x) \qquad \text{for all } x, y \in X. \tag{8}$$

For each $i \in V$ and every $x \in X$, a subgradient $\nabla f_i(x)$ exists since f_i is convex over $\Re^n$ (see Theorem 23.4 in [16], or Proposition 4.2.1 in [17]).

In the distributed and parallel optimization algorithm that we study, each agent maintains and updates an iterate. Each iteration of the algorithm has three phases.

Phase 1: Each agent receives the current iterate of a subset of its neighboring agents. We denote the set of agents whose iterates are available to agent i in the $(k+1)$-th iteration by $N_i(k+1)$, where $i \in N_i(k+1)$. We make the following assumptions on the sets $N_i(k+1)$. Define (V, E_{k+1}) to be the directed graph with directed edges $E_{k+1} = \{(j,i) : j \in N_i(k+1),\, i \in V\}$.

Assumption 2: There exists a positive integer Q such that the graph $(V, \cup_{l=1,\ldots,Q} E_{k+l})$ is strongly connected for all $k \ge 0$.

This assumption ensures that every agent influences the iterate value of every other agent, directly or indirectly, every Q iterations.

Phase 2: Each agent calculates a weighted sum of all the iterates that it has access to. We denote the weighted sum calculated by agent i as $v_{i,k}$. Therefore,

$$v_{i,k} = \sum_{j\in N_i(k+1)} a_{i,j}(k+1) w_{j,k},$$

where $a_{i,j}(k+1)$ are non-negative weights that satisfy the following assumption.

Assumption 3: For all $k \ge 0$ and $i \in V$,

1) $a_{i,j}(k+1) = 0$ when $(j,i) \notin N_i(k+1)$,
2) $\sum_{j=1}^{m} a_{i,j}(k+1) = 1$,
3) There exists scalar η, $0 < \eta < 1$, such that $a_{i,j}(k+1) \ge \eta$ when $j \in N_i(k+1)$,
4) $\sum_{i=1}^{m} a_{i,j}(k+1) = 1$.

Assumptions 3.1 and 3.2 state that each agent calculates a weighted average of all the iterates it has access to. Assumption 3.3 ensures that each agent gives sufficient weight to its current iterate and all the iterates it receives. We remark that agents need not be aware of the common bound η.[2]

Assumption 3.4 ensures that all the agents are equally influential in the long run. For the weights to satisfy Assumption 3.4, the agents need to coordinate with each other. Some schemes are discussed in [18], [8].

Phase 3: Agent i generates its new iterate from $v_{i,k}$ according to the following rule

$$w_{i,k+1} = \mathcal{P}_X \left[v_{i,k} - \alpha_{k+1} \left(\nabla f_i \left(v_{i,k} \right) \right) \right]. \qquad (9)$$

Here α_{k+1} is the stepsize and $\mathcal{P}_X$ is the Euclidean projection onto the set X. This algorithm is a constrained version of the algorithm that was proposed in [18]. The projection in (9) complicates the analyses of the algorithm by introducing a non-linearity in the system dynamics. Further, [18] considers only the case of constant stepsizes, while we study the convergence of the algorithm for diminishing stepsizes. Thus, we succeed in showing convergence to the global optimum, as opposed to convergence to a region around the optimum. We next state a theorem that guarantees convergence of the algorithm. The proof is available in [9], where a more general setting is considered to include stochastic errors in the gradient terms.

Theorem 1: Let Assumptions 1, 2 and 3 hold. Let $\{\alpha_k\}$ be a positive non-increasing sequence such that $\sum_k \alpha_k = \infty$ and $\sum_k \alpha_k^2 < \infty$. Then, the iterate sequences $\{w_{i,k}\}$ of the agents converge to the same point in X^*, i.e., there is a vector $\tilde{x} \in X^*$ such that

$$\lim_{k \to \infty} \|w_{i,k} - \tilde{x}\| = 0, \qquad \forall\ i.$$

Note that the theorem cannot be directly applied to the power control problem. This is because the theorem requires the set X to be bounded, while in the power control problem the set X is unbounded (it extends to $-\infty$ along each coordinate direction). However, we can overcome this by imposing the additional constraint that the power used by each MU must be above a certain threshold.

IV. Simulation results

Intuitively, one would expect the parallel and distributed algorithm in (9) to converge faster than the incremental algorithms. This is because each agent updates at the same time. We next verify this through simulations.

We consider a cellular network of 25 square cells. Each cell is of dimensions 10×10. Within each cell, the MU is randomly located and the base station is at the center of the cell. The network is shown in Fig. 1. The channel coefficient is assumed to decay as the fourth power of the distance between the transmitter and receiver. The shadow fading is assumed to be lognormal with variance 0.1. The receiver noise variance is taken to be 0.01. The cost of the power is modeled as $V(p_i) = 10^{-3} p_i$. The stepsize is taken to be $\alpha_k = \frac{10}{n_{s(k)}}$, where n_i is the number of times agent i receives the iterates in the incremental gradient projection

[2]The actual value of η is only used in our analysis.

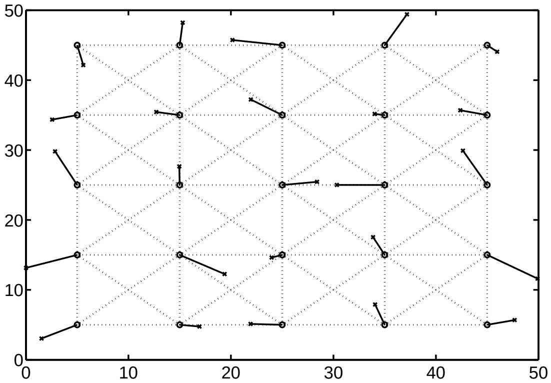

Fig. 1. The circles denote the base stations. The dotted lines denote the noiseless communication links between adjacent BSs. The cross denotes the MUs. The thick lines connect each MU to its respective base station.

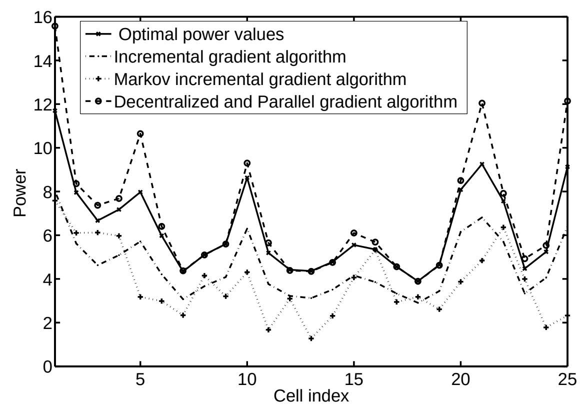

Fig. 2. The final iterate values after 100 cycles of the cyclical incremental gradient projection algorithm, after 2500 iterations of the Markov incremental gradient projection algorithm, and after 2500 iterations of the decentralized and parallel gradient projection algorithm.

algorithms. For the consensus gradient projection algorithm the stepsize is taken to be $\alpha_k = \frac{7}{k^{0.7}}$. First, observe from Fig. 2 that the three algorithms choose power values that are close to the optimal powers. Next, observe from Fig. 3 that a good estimate is obtained after about 500 iterations of the decentralized and parallel gradient projection algorithm. We also observed that it took 20 cycles (or, 500 iterations) of the cyclical incremental gradient projection algorithm and 1000 iterations of the Markov incremental gradient projection algorithm to obtain estimates of the same quality. One can expect a similar behavior when the cells are hexagonal and the the channel coefficients decay differently with distance.

V. Discussion

In this paper, we viewed the uplink power control algorithm as an optimization problem that is to be solved by the base stations in a distributed manner. We formulated the power control problem as the minimization of a sum of convex functions, each of which is known only to a specific base

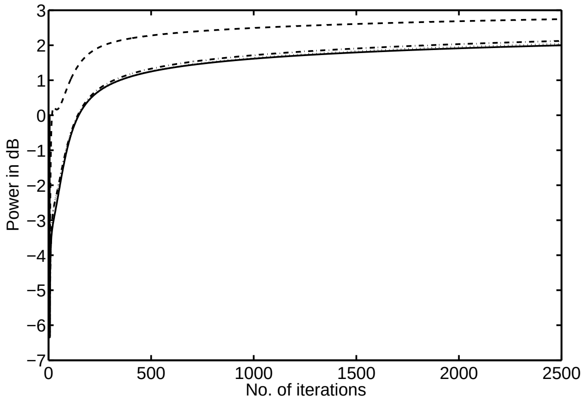

Fig. 3. The iterate sequence generated at a randomly chosen agent in the decentralized and parallel algorithm. Only a few components of the parameter vector x are plotted.

station. We have proposed three distributed power control algorithms that exploit the wired backbone connecting the base stations for localized message passing between neighbors. We also verified the convergence of the algorithms to the optimal value through simulations. The distributed algorithms have also been studied in the presence of stochastic errors in [8], [9]. In the power control problem context, these errors may be due to errors in the channel information that is available at the receiver. Further, the effect of quantizing the iterates before they are communicated is also studied in [8], [9]. This is especially relevant when the power control algorithm proposed here is used in ad hoc networks.

An important criterion that decides the effectiveness of distributed power control algorithms is the time to convergence. If the time to convergence is slow, the channel and the number of users may change by the time the optimal power is determined. The autonomous algorithms discussed earlier require each MU in each iteration to transmit a signal with power equal to the current iterate value. In addition, there is feedback from the BS to the MU. In contrast, the non-autonomous approach proposed here only requires communication between neighboring base stations over wired links. Thus, the energy consumption at the MUs is lower in the non-autonomous algorithms. Further, the non-autonomous algorithms may potentially converge faster. As a part of our future research, we plan to compare the rate of convergence through simulations.

While we consider only the uplink of the wireless cellular networks, the algorithm and the discussion can be extended to the downlink channel. More generally, the non-autonomous algorithms proposed here can be used whenever communication between the receivers is cheaper than communication between the receiver and the transmitter. This would be the case in ad hoc wireless networks when the receivers are physically close to each other.

VI. Acknowledgments

The authors would like to thank R. Srikant and Che Lin for helpful discussions.

References

[1] M. Chiang, P. Hande, T. Lan, and C. Tan, "Power control in wireless cellular networks," *Foundations and Trends in Networking*, vol. 2, no. 4, pp. 381–533, 2008.

[2] G. Foschini and Z. Milanjic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Transactions on Vehicular Technology*, vol. 42, no. 4, pp. 641–646, 1993.

[3] R. Yates, "A framework for power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341—1347, 1995.

[4] T Alpcan, T Basar, R. Srikant, and E Altman, "CDMA uplink power control as a noncooperative game," *Wireless Networks*, vol. 8, no. 6, pp. 659—670, 2004.

[5] D. Falomari, N. Mandayam, and D. Goodman, "A new framework for power control in wireless data networks: Games utility and pricing," in *Proceedings of Allerton Conference on Communication, Control and Computing,*, 1998, pp. 546–555.

[6] P. Hande, S. Rangan, and M. Chiang, "Distributed uplink power control for optimal SIR assignment in cellular data networks," in *Proceedings of 25th IEEE INFOCOM*, 1006, pp. 1–13.

[7] J. Huang, R. Berry, and M. Honig, "Distributed interference compensation for wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 5, pp. 1074–1085, 2006.

[8] S. Sundhar Ram, A. Nedić, and V. V. Veeravalli, "Incremental stochastic sub-gradient algorithms for convex optimization," Submitted to SIAM Journal of Optimization. Available at http://arxiv.org/abs/0806.1092, 2008.

[9] S. Sundhar Ram, A. Nedić, and V. V. Veeravalli, "Distributed stochastic subgradient algorithms for convex optimization," http://arxiv.org/abs/0811.2595, 2008.

[10] A. Nedić and D. P. Bertsekas, "Incremental subgradient method for nondifferentiable optimization," *SIAM Journal of Optimization*, vol. 12, 2001.

[11] J. Mo and J. Walrand, "Fair end-to-end window based congestion control," *IEEE/ACM Transactions on Networking*, vol. 8, no. 5, pp. 556–567, 2000.

[12] X. Qui and K. Chawla, "On the performance of adaptive modulation in cellular systems," *IEEE Transactions on Communications*, vol. 47, no. 6, pp. 884—895, 1999.

[13] M. Chiang, "Geometric programming for communication systems," *Foundations and Trends in Communications and Information Theory*, vol. 2, no. 1-2, pp. 1—156, 2005.

[14] D. Julian, M. Chiang, D. O'Neill, and S. Boyd, "Qos and fairness constrained convex optimization of resource allocation for wireless cellular and ad hoc networks," in *Proceedings of the 21st IEEE INFOCOM*, 2002, vol. 2, pp. 477—486.

[15] B. Johansson, M. Rabi, and M. Johansson, "A simple peer-to-peer algorithm for distributed optimization in sensor networks," in *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007, pp. 4705–4710.

[16] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[17] D. P. Bertsekas, A. Nedić, and A. Ozdalgar, *Convex Analysis and Optimization*, Athena Scientific, 2003.

[18] A. Nedić and A. Ozdaglar, "Distributed sub-gradient methods for multi-agent optimization," *To appear in Transactions on Automatic Control*, 2008.