

Towards a Fair Non-convex Resource Allocation in Wireless Networks

G. Tychoiorgos, A. Gkelias and K. K. Leung

Electrical and Electronic Engineering

Imperial College

London SW7 2AZ, UK

{g.tychoiorgos, a.gkelias, kin.leung}@imperial.ac.uk

Abstract—This paper presents a non-convex optimization framework for the Network Utility Maximization problem in Wireless Networks, which incorporates the interference among links and introduces a power penalty term in the objective function to assure both convergence and energy efficiency of the method. Moreover, a distributed gradient based algorithm is proposed that converges to the optimal solution for problems with zero duality gap and a *fair-allocation* heuristic is presented to resolve user oscillations when they occur. Finally, numerical results regarding the performance of the heuristic and the distributed approach are presented.

I. INTRODUCTION

The problem of optimal resource allocation in communication networks using Optimization Theory has found a lot of interest since the *Network Utility Maximization - NUM* framework proposed by Kelly et al. in their seminal paper [1]. NUM models the allocation of network resources as an optimization problem and solves it distributedly using a set of differential equations. Later, [2] proposed an alternative distributed iterative algorithm based on *Duality Theory* to solve the same resource allocation problem. This algorithm makes use of the *Gradient* method to calculate its optimal solution

Both methods are proven to converge to the optimal rate allocation under the assumption of logarithmic utility functions, which are ideal for “elastic” TCP-like traffic, and assuming that all links in the network are wired and therefore their capacity is fixed throughout the optimization process. However, current communication networks usually consist of some (and possibly all links are) wireless links. In these cases, the capacity of a link is not constant but a function of its transmission power and that of the neighboring nodes. Moreover, the properties of the traffic in such networks vary significantly from the traditional TCP-like traffic. For instance, a significant percentage of current network traffic comprises real-time applications, such as video and audio streaming, and

their “inelastic” properties cannot be modeled using logarithmic utilities.

These characteristics of current communication networks have attracted increased effort from the research community to extend the NUM framework to model “inelastic” traffic. Fazel et al. in [3] remove the restricting assumption of concave utilities and propose a centralized method based on sum-of-squares relaxations to calculate approximations of the optimal solution along with some performance bounds to evaluate the approximation error. Moreover, [4] and [5] have extended the NUM framework to model inelastic traffic using sigmoidal utilities and have proposed distributed algorithms for wired networks, which however are not guaranteed to converge to the optimal solution in case of oscillations in the network.

On the other hand, there have been some attempts to extend the NUM framework to model networks with wireless links. Chiang in [6] extends the framework by proving the concavity of an approximation of Shannon’s capacity formula in high SINR environments. In [7], we have proposed an alternative NUM formulation to jointly perform *resource allocation* and *power control* in wireless networks by incorporating the interference among wireless links. Finally, in [8] we propose a general necessary and sufficient condition to identify the subset of non-concave optimization problems that can be solved distributedly and describe a non-concave NUM formulation in *hybrid* ad-hoc networks. Nevertheless, all these works attempt to either tackle the “inelastic” traffic modeling or the existence of wireless links in a network. The lack of a complete distributed optimization framework that accounts for both issues while at the same time resolves any oscillations (caused by non-convexity) in a *fair* manner is our motivation behind this work.

The contribution of this paper is threefold. We propose a non-convex optimization framework for wireless networks that, firstly, incorporates the main characteristic of the wireless medium, namely the interference among links, and secondly, introduces a power penalty term in the objective function to ensure convergence and energy efficiency of the power control subproblem. Finally, we propose the *fair-allocation* heuristic to resolve user oscillations when they occur.

The rest of the paper is organized as follows: First, the non-convex formulation is presented in Section II. Consequently in Section III, a distributed gradient based algorithm is proposed to solve the problem and the *fair-allocation* heuristic is pro-

Research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

posed to resolve cases of oscillation due to the non-concavity of the optimization problem. Section IV presents numerical results comparing our heuristic against other algorithms in literature and Section V concludes our results and outlines our future work.

II. PROBLEM FORMULATION

Consider a multihop wireless network where nodes can operate either as sources, destinations or relays (i.e., intermediate nodes that do not generate any new traffic, just forward traffic to their neighbors) and let $\mathbf{r} = [r_1, r_2, \dots, r_M]^T$ represent the rate vector that includes the transmission rates of all M source nodes in the network. A link l is defined as a tuple (T_l, R_l) , where T_l is a transmitting and R_l a receiving node, respectively, while $\mathbf{p} = [p_1, p_2, \dots, p_L]^T$ represents the transmission power vector comprises the transmission powers of the L involved wireless links. Let G be the path loss matrix of size $L \times L$ which depends on the physical characteristics of the wireless links, where G_{ij} is the path loss coefficient from the transmitter of link i to the receiver of link j .

Each source node i receives a *utility* when sending traffic at rate r_i through the network according to a *utility function* $U_i(r_i)$ and indicates its degree of satisfaction. At the same time, each wireless link j has a cost for transmitting at a specific power based on a convex cost function $V_j(p_j)$, which represents the cost of using the limited power resources of the wireless medium and assures the energy efficiency of the solution.

The optimization of the network's performance can be formulated as a maximization problem of the form:

$$\begin{aligned} \max_{\mathbf{r}, \mathbf{p}} \quad & \sum_{i=1}^M U_i(r_i) - \alpha \sum_{j=1}^L V_j(p_j) \\ \text{s. t.} \quad & \sum_{i \in Z(j)} r_i \leq C_j(\mathbf{p}), \quad \forall \text{ links } j \end{aligned} \quad (1)$$

where $Z(j)$ represents the set of traffic flows that are passing through link j . The rates r_i and powers p_j are positive variables, α is a non negative constant and the capacity of a wireless link is a function of the transmission powers of all wireless links using Shannon's capacity formula, $C_j(\mathbf{p}) = B \cdot \log_2(1 + \text{SINR}_j)$. Shannon's formula is evidently a non-concave function of powers which would prevent any gradient based distributed algorithm to converge to the optimal power vector. However, in case that $\text{SINR}_j \gg 1$, i.e. in high SINR environments where $G_{jj} \gg G_{kj}$, $k \neq j$, the link capacity formula can be approximated sufficiently well by (2)

$$\begin{aligned} C_j(\mathbf{p}) &= B \log_2(\text{SINR}_j) \\ &= B \log_2 \left(\frac{p_j G_{jj}}{\sum_{k \neq j} p_k G_{kj} + n_j} \right) \end{aligned} \quad (2)$$

which can be proven to be a concave function of powers after a log transformation of the power vector [6]. For the remainder of this paper, and in order to gain concavity of the link capacity function, we consider *high SINR* environments where the capacity $C_j(\mathbf{p})$ of link j in (1) will be approximated

by (2). This is a fair approximation, used in several other works such as [6], [9] and [10]. When $\alpha > 0$, the optimization process will try to balance the profit of the link with the cost of transmission and thus leading to more energy efficient solutions. Therefore, the value of α and consequently the relative importance of the power cost term in the optimization objective consists a trade-off between the performance and the energy efficiency of the network. Moreover, $V_j(p_j)$ plays an important role in the convergence of the power control part of the distributed algorithm, as explained later.

The utility function of each user $U_i(r_i)$ represents the user satisfaction with respect to the allocated rate to user i . The application generating the traffic at user i determines the shape of the utility function. For instance, applications such as file transfer or browsing are usually modeled by logarithmic utilities, i.e. concave functions, while real-time applications are modeled using sigmoidal functions, which are non-concave. The choices for utility $U_i(r_i)$ in problem (1) are not restricted to concave functions as in the traditional NUM framework, making problem (1) non-convex. In this case, the *duality gap* in non-convex problems can be positive and therefore these problems can not be solved distributedly in general. However, Theorem 1 in [8] can be used to identify a subset of non-convex optimization problems for which the duality gap is zero and the distributed gradient based algorithm can converge to the optimal solution.

The proposed formulation, extends the NUM framework for *wireless* networks by allowing non-concave utility functions while considering mutual interference among links and by using a power penalty term to ensure the convergence of the distributed power control algorithm.

III. DISTRIBUTED ALGORITHM

Problem (1) optimizes the allocation of resources in an ad-hoc network and therefore the applicability of any solution relies on the ability to develop a distributed algorithm with minimum message overhead among nodes. Such an algorithm can be developed using *Duality Theory* and to this purpose, we first define the *Lagrangian* function as:

$$\begin{aligned} L(\mathbf{r}, \mathbf{p}, \boldsymbol{\lambda}) &= \sum_{i=1}^M \left\{ U_i(r_i) - r_i \cdot \left(\sum_{j \in S(i)} \lambda_j \right) \right\} \\ &+ \sum_{j=1}^L \left\{ \lambda_j B \log \left[\frac{p_j G_{jj}}{\sum_{k \neq j} p_k G_{kj} + n_j} \right] - \alpha V_j(p_j) \right\}, \end{aligned} \quad (3)$$

where $S(i)$ is the set of links that source i is using in order to send its traffic to the destination. The *Lagrangian* function described in (3) incorporates two major terms, one that refers to the sources and one that refers to the links. Regarding the physical meaning of these terms, $U_i(r_i)$ is the "profit" that source i will make for sending its traffic at rate r_i , quantity $r_i \cdot (\sum_{j \in S(i)} \lambda_j)$ represents the total cost for source i in order to send r_i b/s of traffic through the network. Then,

term $\lambda_j B \log\left[\frac{p_j G_{jj}}{\sum_{k \neq j} p_k G_{kj} + n_j}\right]$ represents the total “profit” that link j will make by charging each unit of its capacity with λ_j and term $\alpha V_j(p_j)$ represents the *weighted* cost for the link to achieve a capacity of $B \cdot \log\left[\frac{p_j G_{jj}}{\sum_{k \neq j} p_k G_{kj} + n_j}\right]$. After a careful observation of the Lagrangian function, one can see that the optimization process consists of two independent subproblems of the primal variables \mathbf{x} and \mathbf{p} coupled by the dual optimization variable vector $\boldsymbol{\lambda}$. The first subproblem is the *rate allocation*, maximizing the net revenue of each *source*, and the second is a *power control* problem, maximizing the net revenue of the *links*.

The gradient of the Lagrangian function with respect to the optimization variables r_i , p_j and λ_j are:

$$\frac{\partial L(\mathbf{r}, \mathbf{p}, \boldsymbol{\lambda})}{\partial r_i} = U'_i(r_i) - \sum_{j \in S(i)} \lambda_j \quad (4)$$

$$\begin{aligned} \frac{\partial L(\mathbf{r}, \mathbf{p}, \boldsymbol{\lambda})}{\partial \lambda_j} = & B \cdot \log_2 \left[\frac{p_j G_{jj}}{\sum_{k \neq j} p_k G_{kj} + n_j} \right] \\ & - \sum_{i \in Z(j)} r_i \end{aligned} \quad (5)$$

$$\begin{aligned} \frac{\partial L(\mathbf{r}, \mathbf{p}, \boldsymbol{\lambda})}{\partial p_j} = & -\alpha V'_j(p_j) \\ & + \frac{1}{p_j \ln(2)} \left[\lambda_j - \sum_{m \neq j} \lambda_m \frac{G_{jm} P_j}{\sum_{k \neq m} G_{km} P_k + n_m} \right] \end{aligned} \quad (6)$$

and the gradient based iterative algorithm consists of the following equations:

$$r_i(t+1) = r_i(t) + \gamma_r(t) \frac{\partial L(\mathbf{r}, \mathbf{p}, \boldsymbol{\lambda})}{\partial r_i} \quad (7)$$

$$\lambda_j(t+1) = \lambda_j(t) - \gamma_\lambda(t) \frac{\partial L(\mathbf{r}, \mathbf{p}, \boldsymbol{\lambda})}{\partial \lambda_j} \quad (8)$$

$$p_j(t+1) = p_j(t) + \gamma_p(t) \frac{\partial L(\mathbf{r}, \mathbf{p}, \boldsymbol{\lambda})}{\partial p_j}, \quad (9)$$

where $\gamma_r(t)$, $\gamma_\lambda(t)$ and $\gamma_p(t)$ are small positive constants. As mentioned earlier, $V_j(p_j)$ assures convergence of the distributed algorithm. In the absence of this cost function, i.e. when $\alpha = 0$, it is possible to fall in a situation that equation (6) is always positive, leading the distributed algorithm to infinite power vectors. This phenomenon has been first identified in [9] for a simple symmetric network topology but can also be justified in the general case using the following scenario when $\alpha = 0$. Assuming an arbitrary wireless network at some iteration t , the power vector is $\mathbf{p} = [p_j, j \in \{1, \dots, L\}]$ and the rate vector $\mathbf{r} = [r_i^t, i \in \{1, \dots, N\}]$. If the power vector is increased by a small percentage, let ϵ , the new power vector will become $\hat{\mathbf{p}} = (1 + \epsilon) \cdot \mathbf{p}$ and the new capacity of each link will be $C_j(\hat{\mathbf{p}}) = B \cdot \log_2 \left(\frac{(1+\epsilon) p_j G_{jj}}{(1+\epsilon) \cdot \sum_{k \neq j} p_k G_{kj} + n_j} \right) > C_j(\mathbf{p})$. Consequently, larger link capacities will lead to higher rate

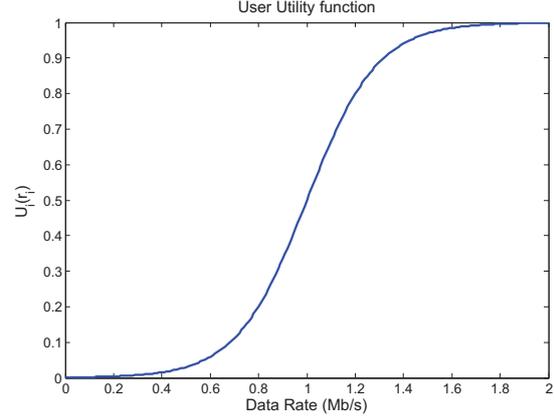


Fig. 1. Example of a sigmoidal utility function

allocation and ultimately to higher aggregate utility in the network. Hence, at the absence of a power penalty function, the distributed algorithm might lead to infinite transmission power. On the other hand, when $\alpha > 0$, there will be a finite power vector \mathbf{p}' at which any further increase would lead to a decrease in the net utility of the network and thus the transmission power will stop increasing.

In literature, this case of infinitely increasing power is often prevented by assuming a transmission power threshold p_j^{max} . Such an assumption, even though is reasonable in a practical system, causes distortion in the theoretical analysis since it creates artificial convergence points. Specifically, according to the *Brouwer Fixed Point Theorem* [11], a continuous mapping of the power vector within a closed range $[p_j^{min}, p_j^{max}]$ creates fixed points of an algorithm that might otherwise never converge. Therefore, the use of the penalty function $V_j(p_j)$ is a more natural way of assuring both energy efficiency and convergence of the distributed power control algorithm.

The dual optimization problem is always a convex problem of the variables λ_j as a point-wise supremum of a family of affine functions. Moreover, the power control subproblem is also convex assuming that the capacity of a link j is given by (2) and the penalty function $V_j(p_j)$ is a convex function of powers. Therefore, (8) and (9) will converge to the optimal solution. However, the rate allocation subproblem is not convex since the utility functions can be non-concave. In this case, the duality gap can be positive and the distributed algorithm can converge to a suboptimal solution. It is possible to use Theorem 1 in [8] to identify those non-convex optimization problems for which the duality gap is zero and the gradient based distributed algorithm can converge to the optimal solution. Theorem 1 relies on the knowledge of the continuity properties of the primal variables as a function of the dual ones. In the context of problem (1), the optimal transmission rate of a source, $r_i^*(\boldsymbol{\lambda})$ is given by:

$$r_i^*(\boldsymbol{\lambda}) = \arg \max \left[U_i(r_i) - r_i \cdot \left(\sum_{j \in S(i)} \lambda_j \right) \right], \quad (10)$$

which is a function of the aggregate price along the path that source i is using, $\lambda_s = \sum_{j \in S(i)} \lambda_j$. However, the continuity properties of (10) are not known for a general utility function

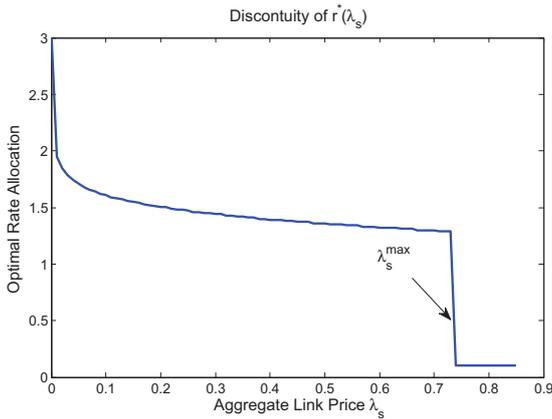


Fig. 2. Optimal Rate Allocation as a function of the aggregate price

and therefore we have to restrict the choices of utilities to more well-studied cases, such as to sigmoidal utilities, which are non-concave functions that model efficiently real-time applications such as video/audio streaming. An example of such a function is shown in Figure 1. For this family of utilities, $r_i^*(\lambda)$ is discontinuous at only one point [5]. This point represents the user's maximum willingness to pay, λ_s^{max} , above which the optimal user rate is always zero. The *maximum willingness to pay* of a user solely depends on its utility function and therefore the choice of utility function uniquely determines its maximum willingness to pay. For instance, Figure 2 shows the optimal rate and the discontinuity at λ_s^{max} for the sigmoidal utility in Figure 1. The optimal rates for aggregate prices less than λ_s^{max} are in the concave region of $U_i(r_i)$ and hence the optimal solution if non-zero will always be in the concave region. Then, according to Theorem 1 [8], iff $\lambda_s^* \neq \lambda_s^{max}$, the duality gap is zero and the distributed algorithm converges to the optimal primal solution following the convergence rate and properties, regarding the step sizes γ_r , γ_λ and γ_p , of the *Gradient Method* [12].

For sigmoidal utilities, and under specific conditions [5], it is possible that the algorithm oscillates between zero and positive values of rates when the aggregate price is around λ_s^{max} . In such cases, the users could run the *fair-allocation* heuristic to resolve oscillations.

The *fair-allocation* heuristic is based on the idea that an oscillating user will be allocated the minimum rate in the concave region of its utility function and will be removed from the rest of the optimization process to allow stability of the network. User oscillations indicate that the optimal rate allocation is non-zero, but due to the discontinuity at λ_s^{max} , the optimal rate is not possible to be calculated. According to the *fair-allocation* heuristic, when user i realizes oscillations between zero and positive rates around λ_s^{max} , user i will be allocated with rate equal to the inflection point of its utility function, as the minimum non-zero rate value that is a candidate optimal solution. Consequently, the optimization process will continue for the rest of the users and with the remaining capacity at each link. The resulting rate allocation for the remaining users is optimal while oscillating users get a non-zero suboptimal rate.

Removing the oscillating users is an obvious decision to

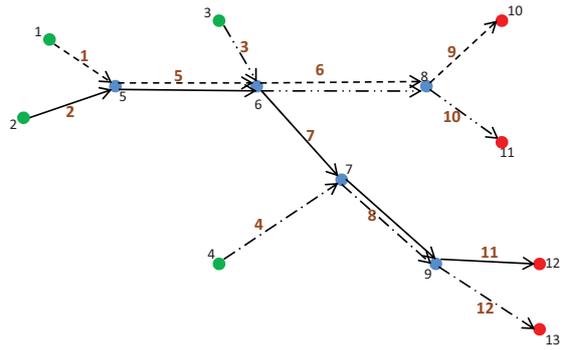


Fig. 3. Example Network Topology

ensure stability of the system but the question lies in the allocated rate to those users. Authors in [5] attempt to solve the oscillation problem by removing the oscillating users from the optimization process without allocating any rate. However, this approach is unfair and suboptimal for finite number of users. The *fair-allocation* heuristic has two advantages against this approach; first, it leads to a more fair resource allocation, since no users are excluded unless the optimal solution indicates zero rate, and, second, it generally leads to higher values of the objective function. The latter is justified as follows: Since both concave and sigmoidal utilities behave in a logarithmic manner around the optimal solution, allocating the total throughput to all users, instead of a subset of them, will yield higher total utility. After removing the oscillating users, if more of the remaining ones start oscillating as well, the *fair-allocation* heuristic will, of course, follow the same procedure to bring the network to a stable condition.

IV. NUMERICAL RESULTS

The *fair-allocation* heuristic was simulated in MATLAB for various network scenarios, an example of which is the network topology shown in Figure 3. The wireless network consists of 4 source nodes (in green), 5 intermediate nodes (in blue) and a set of 4 destination nodes (in red). Source nodes 1 and 4 are HTTP applications with concave utilities while source nodes 2 and 3 are real-time applications with sigmoidal utilities.

The performance of the *fair-allocation* heuristic is compared against two cases: the case that no oscillation resolving method is applied, and the case that the *self-regulating* heuristic [5] is applied in the network to resolve oscillation. Figures 4, 5 and 6 illustrate the performance of the three methods. During iteration 92, the aggregate price for user 2 exceeds its maximum "willingness to pay" and thus user 2 stops transmitting at iteration 93. This abrupt decrease in the transmitted traffic causes a decrease in the aggregate price, which now falls below the maximum "willingness to pay" value of user 2. This encourages user 2 to start sending at a positive rate which in turn initiates a new circle of oscillation. When such an oscillation is observed, the heuristics are evoked to resolve it. The *fair-allocation* heuristic sets the rate to a non-zero value, the inflection point (in this case to 1 Mb/s), and continues the optimization process, while according to the *self-regulating* heuristic the rate is set to zero. According to Figure 4, the decision for non zero rate for the oscillating users yields higher

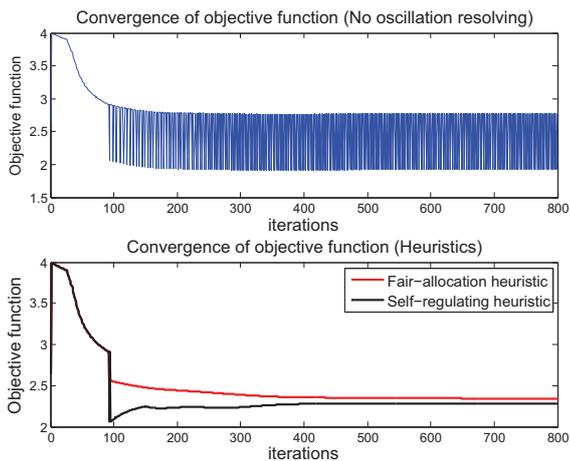


Fig. 4. Convergence of Objective Function

value of the objective function compared to the *self-regulating* heuristic, which is intuitively expected for the reason explained earlier. For brevity, Figure 6 shows the convergence of the power allocation of the first 4 links in the network, the ones initiated from the 4 source nodes.

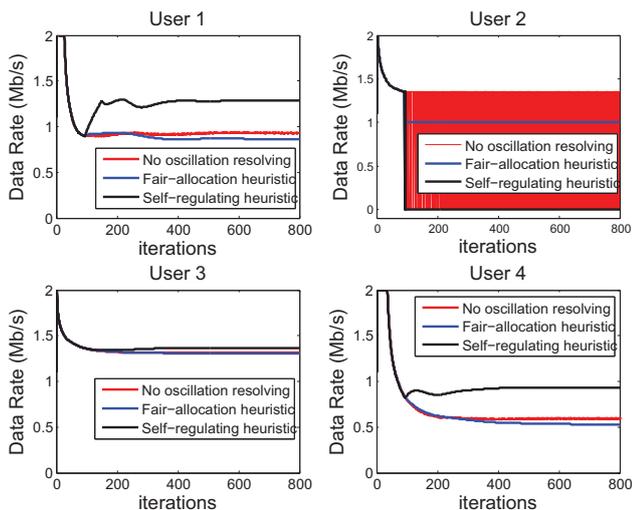


Fig. 5. Convergence of Rate Allocation

V. CONCLUDING REMARKS

This paper extends the theoretical work in [8] by presenting a non-convex optimization formulation for *Network Utility Maximization* in wireless networks and provides a distributed algorithm to solve it optimally for the cases that the duality gap is zero. The formulation includes a power penalty function to assure convergence and energy efficiency of the approach. Moreover, an oscillation resolving heuristic is proposed that assures network stability, leads to a *fair* resource allocation and results in higher aggregate utility compared to other heuristics.

The non-convex optimization framework for wireless networks and the *fair-allocation* heuristic lead to optimal solutions when the duality gap is zero and to an approximation of the optimal solution otherwise. It is within our research

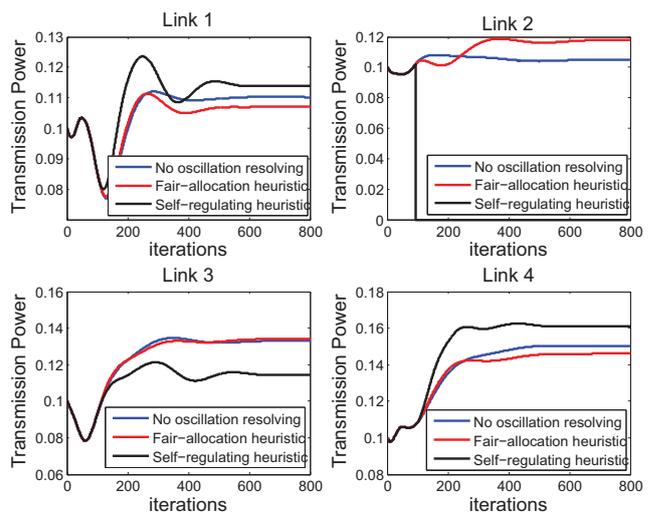


Fig. 6. Convergence of Transmission Power Allocation

interests to identify analytical performance bounds of the heuristic and work towards a distributed framework that would converge to the optimal solution in all cases.

REFERENCES

- [1] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks: Shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, pp. 237–252, 1998.
- [2] S. Low and D. Lapsley, "Optimization flow control. i. basic algorithm and convergence," *Networking, IEEE/ACM Transactions on*, vol. 7, no. 6, pp. 861–874, December 1999.
- [3] M. Fazel and M. Chiang, "Network utility maximization with nonconcave utilities using sum-of-squares method," in *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, December 2005, pp. 1867–1874.
- [4] P. Hande, S. Zhang, and M. Chiang, "Distributed rate allocation for inelastic flows," *Networking, IEEE/ACM Transactions on*, vol. 15, no. 6, pp. 1240–1253, December 2007.
- [5] J.-W. Lee, R. R. Mazumdar, and N. B. Shroff, "Non-convex optimization and rate control for multi-class services in the internet," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 4, pp. 827–840, August 2005.
- [6] M. Chiang, "Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control," *IEEE/ACM Transactions on Communications*, vol. 23, no. 1, pp. 104–116, January 2005.
- [7] G. Tychogiorgos, K. Leung, A. Misra, and T. LaPorta, "Distributed network utility optimization in wireless sensor networks using power control," in *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*, September 2008, pp. 1–6.
- [8] G. Tychogiorgos, A. Gkelias, and K. K. Leung, "A new distributed optimization framework for hybrid ad-hoc networks," submitted for publication.
- [9] Y. Hou, K. K. Leung, and A. Misra, "Mission-based joint optimal resource allocation in wireless multicast sensor networks," in *Annual Conference of ITA*, Maryland, USA, Sept. 2009.
- [10] M. Chiang, "To layer or not to layer: balancing transport and physical layers in wireless multihop networks," in *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 4, march 2004, pp. 2525–2536 vol.4.
- [11] J. S. P. V. Hutson and M. J. Cloud, *Applications of Functional Analysis and Operator Theory*. Academic Press, 1980.
- [12] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.