UNIVERSITY of GREENWICH

Greenwich Academic Literature Archive (GALA)
– the University of Greenwich open access repository
http://gala.gre.ac.uk

Citation for published version:

Liao, S., Sun, J., Chen, Y., Wang, Y. and Zhang, P. (2015) Distributed power control for wireless networks via the alternating direction method of multipliers. Journal of Network and Computer Applications, 55. pp. 81-88. ISSN 1084-8045 (doi:10.1016/j.jnca.2015.05.005)

Publisher's version available at:
http://dx.doi.org/10.1016/j.jnca.2015.05.005

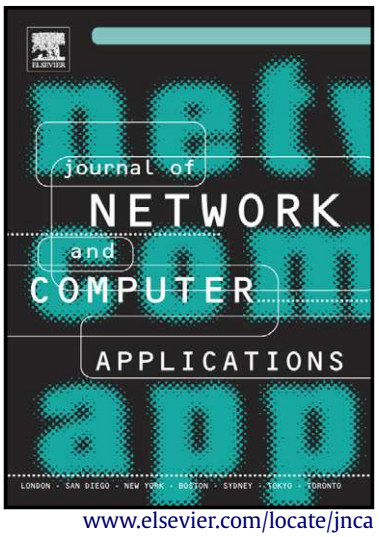
Please note that where the full text version provided on GALA is not the final published version, the version made available will be the most up-to-date full-text (post-print) version as provided by the author(s). Where possible, or if citing, it is recommended that the publisher's (definitive) version be consulted to ensure any subsequent changes to the text are noted.

Citation for this version held on GALA:

Liao, S., Sun, J., Chen, Y., Wang, Y. and Zhang, P. (2015) Distributed power control for wireless networks via the alternating direction method of multipliers. London: Greenwich Academic Literature Archive.
Available at: http://gala.gre.ac.uk/13644/

Contact: gala@gre.ac.uk

Author's Accepted Manuscript



www.elsevier.com/locate/jnca

Distributed Power Control for Wireless Networks via the Alternating Direction Method of Multipliers

S. Liao, J. Sun, Y. Chen, Y. Wang, P. Zhang

PII: S1084-8045(15)00105-8
DOI: http://dx.doi.org/10.1016/j.jnca.2015.05.005
Reference: YJNCA1394

To appear in: *Journal of Network and Computer Applications*

Received date: 20 May 2014
Revised date: 27 November 2014
Accepted date: 12 May 2015

Cite this article as: S. Liao, J. Sun, Y. Chen, Y. Wang, P. Zhang, Distributed Power Control for Wireless Networks via the Alternating Direction Method of Multipliers, *Journal of Network and Computer Applications,* http://dx.doi.org/10.1016/j.jnca.2015.05.005

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Distributed Power Control for Wireless Networks via the Alternating Direction Method of Multipliers

S. Liao[a,*], J. Sun[b], Y. Chen[c], Y. Wang[c], P. Zhang[d]

[a]*National Engineering Research Center, Huazhong Normal University, China 430079*
[b]*Faculty of Engineering and Science, University of Greenwich, Central Avenue, Chatham Maritime, Kent, ME4 4TB, U.K.*
[c]*Department of Computer Science, Huazhong Normal University, Wuhan, China 430079,*
[d]*School of Geodesy and Geomatics, Wuhan University, Wuhan, China 430072*

Abstract

Utility-based power control in wireless networks has been widely recognized as an effective mechanism to managing co-channel interferences. It is based on the maximization of system utility subject to power constraints, which is referred to as power control optimization problem. Global coupling between the mutual interference of wireless channels increases the difficulty of searching global optimum significantly. In this paper, we decouple the optimization problems with concave and non-concave utility functions; and transform them into a global consensus problem by introducing locally slack variables. We then propose two distributed iterative optimization algorithms for the global consensus problems with concave and non-concave objective functions, respectively, based on an alternating direction method of multipliers. Furthermore, we prove that both algorithms converge to the global optimum of the total network utility. Simulation results show the effectiveness of the algorithms. Comparison experiments show that the developed algorithms compare favourably against some other well-known algorithms.

Keywords: Power control, network utility maximization, alternating direction method, dual decomposition, penalty factor

Accepted manuscript

*Corresponding author
Email addresses: liaoshengbin@gmail.com (S. Liao), j.sun@gre.ac.uk (J. Sun), jennycy@126.com (Y. Chen), wri617@163.com (Y. Wang), zhangpeng.hust@gmail.com (P. Zhang)

ACCEPTED MANUSCRIPT

1. Introduction

Interference occurs due to simultaneous transmissions in the same wireless channel. This will significantly deteriorate the performance of wireless networks. To mitigate interference in wireless networks, power control is one of the most-widely used basic techniques (Xiao et al. , 2003). In power control, the interference is treated as noise and the signal-to-interference plus noise ratio (SINR) was mapped to a utility; and the total system utility is to be maximized subject to power constraints which is referred to as the power control optimization problem. It can be considered as a network utility maximization (NUM) problem with coupling power variables (Chiang et al. , 2007). By distributively solving the power control problem, all network users continually adjust their transmission powers until the total utility approaches the optimal, this can achieve the purpose of interference control. In this paper, we address the distributed power control optimization problems where concave and non-concave utility functions are used. In both cases, utility functions depend on the received SINR. The objective is to optimize the overall performance measured in terms of total network utility.

Research work has been done for power control and interference management in wireless networks based on NUM. In Wei et al. (2014), a distributed power control policy for wireless networks based on game-theoretic approaches was proposed. It can only guarantee an optimal power allocation for strictly concave utility functions. In Hussein and Tobias (2012) and Liping et al. (2009), globally optimal power control schemes were developed. However, these algorithms are not suitable for practical implementations in distributed networks. Campos et al. (2008) proposed a distributed power control approach that can guarantee sub-optimality. Hande et al. (2008) developed a distributed power control algorithm which can converge to the global optimum in case of pseudo-linear utility functions. However, the proposed algorithm does not guarantee a global convergence for non-concave utility functions.

The goal of the paper is to develop distributed power control algorithms for both concave and non-concave utility functions, which is similar to the problems solved in Qian et al. (2012) and Lei et al. (2012). In Qian et al. (2012), the authors presented a distributed power control algorithm and its two enhanced versions. These algorithms can maximize any form of system utility and converge to the global optimal solution through Gibbs Sampling (Constantino , 1988). In these algorithms, each transmitter needs to update

Accepted manuscript

ACCEPTED MANUSCRIPT

its power setting according to a given Gibbs probability distribution. These algorithms are essentially heuristic, and their performances highly depend on the parameters of the Gibbs distribution. In practice, it is very challenging to set optimal Gibbs parameters for best algorithmic performances.

In Lei et al. (2012), the utility maximization problem is transformed into an equivalent max-min problem. The extended duality theory (Chen and Chen, 2010) is applied on the max-min problem to devise a distributed power control algorithm in which each transmitter stochastically adjusts its power to improve the total utility by simulated annealing (Kirkpatrick et al., 1983). The authors argued that the proposed distributed power control algorithm can guarantee global optimality but at the cost of slow convergence due to the simulated annealing used for transmitter's power adjustment.

In this paper, we propose to decouple the NUM problem with coupled utilities by introducing a set of local slack variables. This method is similar to the one presented in (Lei et al. , 2012) where a global slack variable is introduced. However, the original problem is transformed into a *global consensus problem* (Nedi and Ozdaglar , 2010) in this paper, while a max-min problem is created in Lei et al. (2012).

To design a distributed power control algorithm, we combine the extended duality theory with the alternating direction method of multipliers (ADMM) (Boyd et al., 2011). Specifically, in case of concave utility functions, we directly use the Lagrangian duality and ADMM to develop the distributed power allocation algorithm. On the other hand, in case of non-concave utility functions, we build up a new nonlinear Lagrangian function. It contains two kinds of penalty terms, i.e., an l_1^m-penalty term (Chen and Chen, 2010) and a quadratic penalty term. Our method combines the ideas of the extended duality theory and ADMM. The extended duality theory can guarantee zero duality gap between the primal and dual problems in case that a l_1^m-penalty function is used for non-concave problems. The quadratic penalty term used in ADMM can drive all local variables to reach a global consensus.

The rest of the paper is organized as follows. After a brief description of the ADMM in Section 2, the model and the considered problem are presented in Section 3. In Section 4, we apply the ADMM to obtain a distributed algorithm for concave utility functions. Section 5 presents a distributed algorithm for nonconcave utility functions based on the ADMM and the extended dual theory. Section 6 describes the simulation results. Section 7 concludes the paper.

Accepted manuscript

ACCEPTED MANUSCRIPT

2. Preliminaries

In this section, we briefly review the alternating direction method of multipliers (ADMM) and the distributed iterative algorithms based upon it, which form the basis of the distributed iterative algorithms presented in this paper.

Let's consider the following constrained optimization problem

$$\begin{array}{ll} \min & f(x)+g(z) \\ \text{subject to} & Ax+Bz=c \end{array} \tag{1}$$

where variables $x \in R^n$ and $z \in R^m$, $A \in R^{p\times n}$, $B \in R^{p\times m}$, $c \in R^p$, and $f : R^n \to R$ and $g : R^m \to R$. The augmented Lagrangian for (1) can be written as follows:

$$L_\rho(x,z,\mu) = f(x)+g(z)+\mu^T(Ax+Bz-c)+\frac{\rho}{2}\|Ax+Bz-c\|_2^2 \tag{2}$$

The corresponding augmented dual function is given by

$$D_\rho(\mu) = \min_{x,z} L_\rho(x,z,\mu) \tag{3}$$

where μ is the dual variable and $\rho \geq 0$ is the penalty parameter. The dual problem corresponding to the primal problem (1) is defined as

$$\max_{\rho} D_\rho(\mu) \tag{4}$$

Further, applying the dual ascent method to the dual problem yields the following iterative algorithm:

$$(x(k+1), z(k+1)) = \arg\min_{x,z} L_\rho(x,z,\mu(k)) \tag{5}$$

$$\mu(k+1) = \mu(k)+\rho(Ax(k+1)+Bz(k+1)-c) \tag{6}$$

where k is the iteration counter and ρ is the step size.[1] The algorithm (5)-(6) is called as the basic method of multipliers for solving (1). From the augmented Lagrangian (2), it can be seen that L_ρ is not separable. Thus, the minimization step w.r.t. (x, z) (Eq. (5)) cannot be carried out in parallel.

[1]The purpose for using ρ as the step size in the dual update is to guarantee that $(x(k+1), z(k+1), \mu(k+1))$ is dual feasible. For details, please refer to (Boyd et al., 2011)

Accepted manuscript

ACCEPTED MANUSCRIPT

This means that the above basic method of multipliers cannot be used for distributed optimization.

ADMM addresses this issue by separating the minimization over x and z into two steps. The ADMM steps can be summarized as follows:

$$x(k+1) = \arg\min_{x} L_\rho(x, z(k), \mu(k)) \quad (7)$$

$$z(k+1) = \arg\min_{z} L_\rho(x(k+1), z, \mu(k)) \quad (8)$$

$$\mu(k+1) = \mu(k) + \rho(Ax(k+1) + Bz(k+1) - c) \quad (9)$$

where the primary variables x and z are updated in an alternating mode, which accounts for the term *alternating direction*. In ADMM, the dual variable update uses a step size which is equal to the penalty parameter ρ.

The basic idea of ADMM is to eliminate some or all of the constraints by adding a penalty term (cf. (2)). Associated with the method is the penalty parameter ρ. ρ determines the severity of the penalty for infeasible solutions and the extent to which the resulting minima of Eq. (2) approximates the optimal solutions of the original constrained problem (1). Theoretically, along the increase of the penalty parameter ρ, the approximation becomes increasingly accurate. In this paper, we will use ADMM to produce acceptable suboptimal solutions for practical use when a high accuracy cannot be obtained.

3. System Model and Problem Formulation

We consider a wireless network with a set L of links with interference-limited channels. Each link consists of a transmitter and a receiver. We assume that multiple links may transmit data simultaneously, and the transmission from one link may be interfered by other concurrent transmitters.

Let p_l denote the transmission power of the transmitter of link l with P_l^{max} being its maximum power constraint and n_l be the receiving noise on link l. The channel is modeled by a channel gain matrix G, where G_{lk} is the power gain from the transmitter of link l to the receiver of link k. The channel model is shown in Fig.1.

The received signal-to-interference-plus-noise ratio (SINR) for link l can be obtained as follows:

$$\gamma_l(\mathbf{p}) = \frac{G_{ll}p_l}{\sum_{k\neq l} G_{kl}p_k + n_l} \quad (10)$$

Accepted manuscript

ACCEPTED MANUSCRIPT

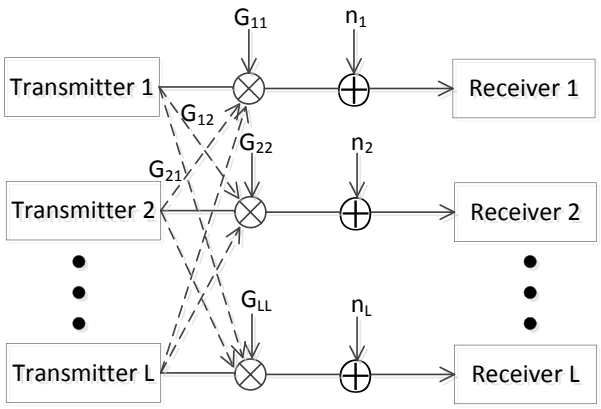


Figure 1: The channel model of a wireless network.

where $\mathbf{p} = (p_1, p_2, \ldots, p_L)$ is a vector of the transmission powers of all links. Associated with the SINR γ_l of each link l is a utility function $U_l(\gamma_l(\mathbf{p}))$. By maximizing the aggregate utility functions $\sum_l U_l(\gamma_l(\mathbf{p}))$ under the maximum power constraints of links, we can establish the corresponding network utility maximization in a wireless network as follows.

$$\max \sum_{l \in L} U_l(\gamma_l(\mathbf{p})) \text{ subject to } 0 \leq p_l \leq P_l^{max}, \forall l \in L \tag{11}$$

Remark 1 Many kinds of utility functions can be applied in different applications (Shenker , 1995; Liao and Campbell , 2001). In the next two sections, we present the concave and non-concave utility functions used in this paper.

4. Distributed power control for concave utility maximization

Assume that the utility functions in (11) are concave and the feasible SINR region is convex, then (11) is a convex optimization problem (Qian et al. , 2009). However, as a result of the power-coupled model, i.e., each utility depends not only its own transmission power but on others, solving the problem (11) requires coordination among different nodes. It is impractical in wireless ad hoc networks.

A key idea to tackle power-coupled utilities is to introduce slack variables and build up a *global consensus problem* (Nedi and Ozdaglar , 2010). In this section, we present a distributed algorithm for solving the proposed consensus problem in conjunction with ADMM.

Accepted manuscript

ACCEPTED MANUSCRIPT

By introducing slack variables $(\mathbf{x}_l, l \in L)$ and regarding primary variable $\mathbf{p}$ as global common variable, the optimization problem (11) can be formulated equivalently as

$$\begin{aligned} \min \quad & \sum_{l\in L} -U_l(\gamma_l(\mathbf{x}_l)) \\ \text{subject to} \quad & \\ & \mathbf{x}_l = \mathbf{p}, \forall l \in L \\ & 0 \le \mathbf{p} \le (P_1^{max}, \ldots, P_l^{max})^T \end{aligned} \tag{12}$$

Then, the augmented Lagrangian for (12) can be written as follows.

$$L_\rho(\mathbf{x}_1, \cdot, \mathbf{x}_{|L|}, \mathbf{p}, \boldsymbol{\mu}_l, \cdot, \boldsymbol{\mu}_{|L|}) = \sum_{l=1}^{|L|} \left(-U_l(\gamma_l(\mathbf{x}_l)) + \boldsymbol{\mu}_l^T(\mathbf{x}_l - \mathbf{p}) + \frac{\rho}{2}\|\mathbf{x}_l - \mathbf{p}\|_2^2 \right) \tag{13}$$

where $|L|$ denotes the number of links, $\boldsymbol{\mu}_l, l \in L$ are the variables of multipliers and ρ is a nonnegative penalty factor. Let $\boldsymbol{\lambda}_l = \frac{1}{\rho}\boldsymbol{\mu}_l$, and define the residual variables $\mathbf{r}_l = \mathbf{x}_l - \mathbf{p}, l \in L$. Then, from (13) we have

$$\begin{aligned} \mathbf{x}_l(k+1) &= \arg\min_{\mathbf{x}_l}(-U_l(\gamma_l(\mathbf{x}_l)) + \boldsymbol{\mu}_l^T(k)\mathbf{r}_l + (\rho/2)\|\mathbf{r}_l\|_2^2) && (14) \\ &= \arg\min_{\mathbf{x}_l}(-U_l(\gamma_l(\mathbf{x}_l)) + (\rho/2)[(2/\rho)\boldsymbol{\mu}_l^T(k)\mathbf{r}_l + \|\mathbf{r}_l\|_2^2) && (15) \\ &\quad + \frac{1}{\rho^2}\|\boldsymbol{\mu}_l\|_2^2] - \frac{1}{2}\rho\|\boldsymbol{\mu}_l\|_2^2 && (16) \\ &= \arg\min_{\mathbf{x}_l}(-U_l(\gamma_l(\mathbf{x}_l)) + (\rho/2)\|\mathbf{r}_l + \boldsymbol{\lambda}_l\|_2^2 - (\rho/2)\|\boldsymbol{\lambda}_l\|_2^2) && (17) \\ &= \arg\min_{\mathbf{x}_l}(-U_l(\gamma_l(\mathbf{x}_l)) + (\rho/2)\|\mathbf{r}_l + \boldsymbol{\lambda}_l\|_2^2) && (18) \end{aligned}$$

where the last step holds because $\frac{\rho}{2}||\boldsymbol{\lambda}_l||_2^2$ is not related to the variable $\mathbf{x}_l$. Similarly, we have

$$\mathbf{p}(k+1) = \Pi_{\mathcal{C}}\left(\frac{1}{|L|}\sum_{l=1}^{|L|}(\mathbf{x}_l(k+1) + \boldsymbol{\lambda}_l(k)) \right) \tag{19}$$

here $\mathcal{C} = \{\mathbf{c} = (c_1, \ldots, c_{|L|}) \big| 0 \le c_l \le P_l^{max}, \forall l \in L\}$ and $\Pi_{\mathcal{C}}(\mathbf{c})$ denotes the projection of vector $\mathbf{c}$ onto $\mathcal{C}$.

Based on previous derivations, the resulting consensus ADMM algorithm for solving (12) can be summarized in Alg. 1:

Accepted manuscript

ACCEPTED MANUSCRIPT

Algorithm 1 The consensus ADMM algorithm for solving (12).

1: Initialization: Choose the initial values of variables $\mathbf{x}_l(0), \boldsymbol{\lambda}_l(0), \mathbf{p}(0)$; and input a stopping criterion;
2: **for** $k = 0, 1, 2, \ldots,$ **do**
3: Locally updates dual variables:

$$\boldsymbol{\lambda}_l(k+1) = \boldsymbol{\lambda}_l(k) + \mathbf{x}_l(k) - \mathbf{p}(k), \forall l \in L \tag{20}$$

4: Solve sub-optimization problems for $l \in L$:

$$\mathbf{x}_l(k+1) = \arg\min_{\mathbf{x}_l} \left(-U_l(\gamma_l(\mathbf{x}_l)) + \frac{\rho}{2} \|\mathbf{r}_l + \boldsymbol{\lambda}_l\|_2^2 \right) \tag{21}$$

5: Global aggregation: averages local variables and broadcasts the result back to each link.

$$\mathbf{p}(k+1) = \Pi_{\mathcal{C}} \left(\frac{1}{|L|} \sum_l (\mathbf{x}_l(k+1) + \boldsymbol{\lambda}_l(k+1)) \right) \tag{22}$$

6: If stopping criteria is satisfied, then stop, otherwise go to 2.
7: **end for**

Accepted manuscript

ACCEPTED MANUSCRIPT

Remark 2. Averaging Eq. (20) gives

$$\overline{\boldsymbol{\lambda}}(k+1) = \overline{\boldsymbol{\lambda}}(k) + \frac{1}{|L|}\sum_{l=1}^{|L|}(\mathbf{x}_l(k) - \mathbf{p}(k))$$

Substituting Eq. (22) into the above equation shows that $\overline{\boldsymbol{\lambda}}(k+1) = 0$. This reveals that although the dual variables are distributively updated, they drive the primary variables into consensus. Meanwhile, for a higher penalty parameter ρ, the penalty $(\rho/2)\|\mathbf{r}_l + \boldsymbol{\lambda}_l\|_2^2$ in Eq. (21) pushes all of the primary variables toward their average value. This suggests that the sequence of iterations (Eq. 21) is a good approximation to the primary problem (11).

Remark 3. Note that there is no local information exchange between links in Algorithm 1. This differentiates significantly from other distributed algorithms in wireless networks. However, the processing centre needs to gather all local variables to update the global variable and broadcast it back to the links (cf. line 5).

Remark 4. The choice of the stopping criteria depends on many conditions. It is proved that when the primal and dual residuals[2] are small, the objective sub-optimality also must be small (Eckstein and Bertsekas , 1992; Boyd et al., 2011). Thus, a common sense to choose a reasonable stopping criterion is that the primal and dual residuals must be smaller than a predetermined threshold.

The following result can be proved using standard analysis techniques in ADMM's convergence analysis as in (Nedi and Ozdaglar , 2010; Eckstein and Bertsekas , 1992; Gabay , 1983; Boyd et al., 2011).

Proposition 4.1. *Assume that the utility function in Algorithm 1 is bounded, closed, concave, and when $\rho = 0$, the un-augmented Lagrangian L_0 has a saddle point, then the objective function in Algorithm 1 approaches the optimal of the primary problem.*

Proof. See Appendix A. □

[2]In Alg. 1, the primal and dual residuals are of the forms $\mathbf{x}_l(k+1) - \mathbf{p}(k+1)$ and $-\rho(\mathbf{p}(k+1) - \mathbf{p}(k))$, respectively.

Accepted manuscript

ACCEPTED MANUSCRIPT

5. Distributed power control for Non-concave Utility Maximization

Assume that the utility functions in the model (11) are non-concave, then (11) is a non-convex optimization problem. In practice, it is very difficult to find a global optimal solution distributively for the problem.

Algorithm 1 fails to find global optimal solution as the duality gap appears due to the non-concave utility functions. Actually, if we remove the last term in the augmented Lagrangian (13), it becomes the ordinary Lagrangian which will cause the duality gap for non-convex optimization. Fig. 2 shows the zero duality gap and the nonzero duality for the ordinary Lagrangian duality, where S denotes the set of constraint-objective pairs. The primal problem can be considered as the *Min Common Point Problem* (MCPP). That is, it requires to find the minimum intercept of S with the w-axis. On the other hand, the dual problem can be seen as the *Max Intercept Point Problem* (MIPP). That is, it requires to find the maximum point of interception of the w-axis from all the hyperplanes supporting S from below (Bertsekas , 1995). The bold black part denotes the intersection of S and the w-axis. It can be seen that the zero duality gap was achieved by the ordinary Lagrangian duality as shown in Fig.2(a). This means that the MCPP and the MIPP are the same; while the nonzero duality gap (i.e., the red part) appeared for the non-convex problem in Fig.2(b) means that there is no such hyperplane which passes through the minimal common point and supports S from blow.

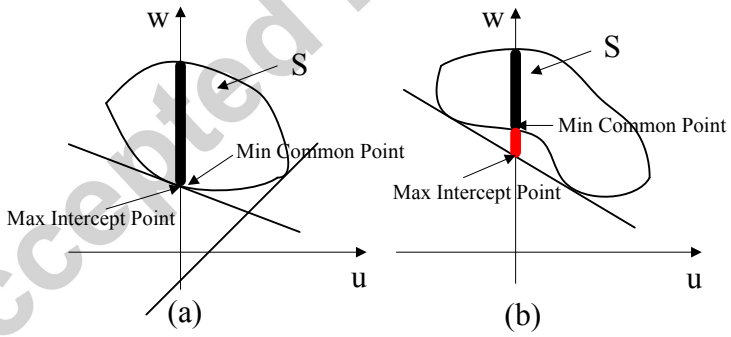


Figure 2: The geometric interpretation of the ordinary Lagrangian duality.

Considerable amount of research work has been conducted for reducing or eliminating the duality gap for non-convex problems (e.g. (Rockafellar and Wets , 1998; Rubinnov et al., 2002; Nedi and Ozdaglar, 2008; Chen and

ACCEPTED MANUSCRIPT

Chen, 2010)). In (Chen and Chen, 2010), an extended duality theory was proposed to achieve nonzero duality gap for non-convex problems based on a new l_1^m-penalty formulation which is a generalization of l_l-penalty. For problem (12), the l_1^m-penalty function is defined as follows:

$$L_{EX}(\mathbf{x}_1,\ldots,\mathbf{x}_l,\mathbf{p},\boldsymbol{\mu}_l)=\sum_{l=1}^{|L|}(-U_l(\gamma_l(\mathbf{x}_l))+\boldsymbol{\mu}_l^T|\mathbf{x}_l-\mathbf{p}|) \tag{23}$$

However, to reach a global consensus of all local variables, we modify the extended Lagrangian for problem (12) as

$$L_\rho(\mathbf{x}_1,\ldots,\mathbf{x}_l,\mathbf{p},\boldsymbol{\mu}_l)=\sum_{l=1}^{|L|}(-U_l(\gamma_l(\mathbf{x}_l))+\boldsymbol{\mu}_l^T|\mathbf{x}_l-\mathbf{p}|+(\rho/2)\|\mathbf{x}_l-\mathbf{p}\|_2^2) \tag{24}$$

The key ideas in (23) and (24) for eliminating the duality gap of the non-convex problems can be visually illustrated in Fig. 3.

The extended duality removes the duality gap by pushing different hyperplanes up to the minimum common point $(0, f^*)$ to collectively support S as shown in Fig.3(a). From equation (24), we can see that it consists of two parts, i.e., the extended Lagrangian and a quadratic penalty term. Therefore, our proposed Lagrangian can be regarded as a perturbation around the extended Lagrangian because of the presence of a quadratic penalty term, which will make the Lagrangian function nonlinear (cf. the dotted lines). The extended Lagrangian can guarantee zero duality gap between the primal and dual problems, while the quadratic penalty item used in ADMM can drive all local variables to reach a global consensus. These two properties can avoid the duality gap while preserving the global optimality as shown in Fig.3(b).

Based on the above analysis, we can use (24) as its extended Lagrangian for problem (11) with non-concave functions. This combines the ideas of the extend duality theory and the augmented Lagrangian method. Thus, we have

$$\begin{aligned}\mathbf{x}_l(k+1) &= \arg\min_{\mathbf{x}_l}(-U_l(\gamma_l(\mathbf{x}_l))+\boldsymbol{\mu}_l^T|\mathbf{x}_l-\mathbf{p}|+\frac{\rho}{2}\|\mathbf{x}_l-\mathbf{p}\|_2^2)\\ &= \arg\min_{\mathbf{x}_l}\Big(-U_l(\gamma_l(\mathbf{x}_l))+\boldsymbol{\mu}_l^T[(\mathbf{x}_l-\mathbf{p})\otimes\mathrm{sign}(\mathbf{x}_l-\mathbf{p})]+\frac{\rho}{2}\|\mathbf{x}_l-\mathbf{p}\|_2^2\Big)\end{aligned} \tag{25}$$

Accepted manuscript

ACCEPTED MANUSCRIPT

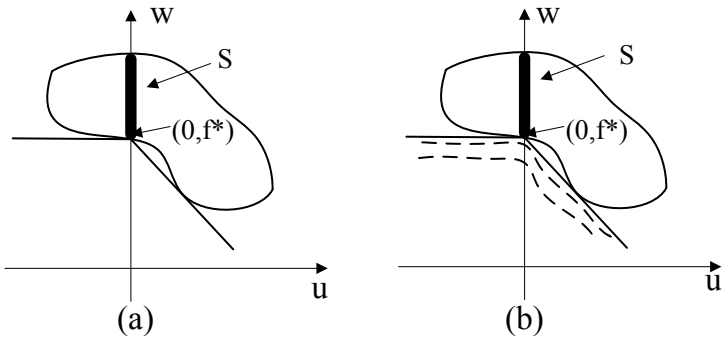


Figure 3: The geometric interpretation of the extended Lagrangian duality.

and

$$\mathbf{p} = \arg\min_{\mathbf{p}} L_\rho(\mathbf{x}_1, \ldots, \mathbf{x}_l, \mathbf{p}, \boldsymbol{\mu}_l) \quad (26)$$

$$= \arg\min_{\mathbf{p}} \sum_{l=1}^{|L|} (\boldsymbol{\mu}_l^T |\mathbf{x}_l - \mathbf{p}| + (\rho/2)\|\mathbf{x}_l - \mathbf{p}\|_2^2) \quad (27)$$

$$= \arg\min_{\mathbf{p}} \sum_{l=1}^{|L|} (\boldsymbol{\mu}_l^T [(\mathbf{x}_l - \mathbf{p}) \otimes \text{sign}(\mathbf{x}_l - \mathbf{p})] + (\rho/2)\|\mathbf{x}_l - \mathbf{p}\|_2^2) \quad (28)$$

where function $\text{sign}(\mathbf{y})$ returns the signs of the elements of $\mathbf{y}$; $\otimes$ denotes element-wise production of two vectors.

Similarly to the method adopted in Section 4, from Eq. (26), we have

$$\mathbf{p}(k+1) = \frac{1}{|L|} \sum_{l=1}^{|L|} (\mathbf{x}_l(k+1) + \frac{1}{\rho}\boldsymbol{\mu}_l(k+1) \otimes \text{sign}(\mathbf{x}_l(k+1) - \mathbf{p}(k))) \quad (29)$$

Adopting the same method in Section 3, an extended ADMM for solving the primary problem (11) with non-concave functions can be summarized in Alg. 2.

Remark 5. Solving problems (25) and (26) usually are difficult because of the term $\boldsymbol{\mu}_l^T|\mathbf{x}_l - \mathbf{p}|$. However, each link l or the processing centre in lines 4 and 5 of Algorithm 2 can locally determine the positive and negative signs of the components of $\mathbf{x}_l - \mathbf{p}$. Therefore, the absolute value signs in (25) and (26) can be removed as above.

Remark 6. The parameter ρ in Algorithm 2 has double roles, i.e., as a penalty factor and a step sizes (see (31),(30) and (32), respectively). In

Accepted manuscript

ACCEPTED MANUSCRIPT

Algorithm 2 Extended ADMM for solving (11) with non-concave functions.

1: Initialization: Choose the initial values of variables $\mathbf{x}_l(0), \boldsymbol{\mu}_l(0), \mathbf{p}(0)$; and input a stopping criterion;

2: **for** $k = 0, 1, 2, \ldots$ **do**

3: Update locally the dual variables:

$$\boldsymbol{\mu}_l(k+1) = \boldsymbol{\mu}_l(k) + \rho(|\mathbf{x}_l(k) - \mathbf{p}(k)|), \forall l \in L \tag{30}$$

4: Solve the following sub-optimization problem for all $l \in L$

$$\mathbf{x}_l(k+1) = \arg\min_{\mathbf{x}_l}(-U_l(\gamma_l(\mathbf{x}_l)) + \boldsymbol{\mu}_l^T[(\mathbf{x}_l - \mathbf{p}) \otimes \text{sign}(\mathbf{x}_l - \mathbf{p})] + \frac{\rho}{2}\|\mathbf{x}_l - \mathbf{p}\|_2^2) \tag{31}$$

5: Global aggregation: averages local variables and broadcasts the result back to each link.

$$\mathbf{p}(k+1) = (1/|L|)\sum_{l=1}^{|L|}(\mathbf{x}_l(k+1) + \frac{1}{\rho}\boldsymbol{\mu}_l(k+1) \otimes \text{sign}(\mathbf{x}_l(k+1) - \mathbf{p}(k))) \tag{32}$$

6: If stopping criteria is satisfied, then stop; otherwise go to 3.

7: **end for**

Accepted manuscript

ACCEPTED MANUSCRIPT

a general sense, an augmented Lagrangian algorithm needs to continuously increase the penalty factor, while an extend Lagrangian algorithm needs to reset the penalty factor after some consecutive iterations, to guarantee convergence. Algorithm 2 combines the different features by setting ρ a constant value which is larger than a particular threshold.

Proposition 5.1. *Suppose the utility functions of the primary problem (12) are continuous, the sequence $\boldsymbol{\mu}_l(k)$ in Algorithm 2 is bounded, and the penalty factor $\rho \geq \rho_0$ (ρ_0 is a nonnegative constant), then every limit of the sequence $(\boldsymbol{x}_1(k), \ldots, \boldsymbol{x}_{|L|}(k), \boldsymbol{P}(k))$ produced by Algorithm 2 is a global optimal solution to the primary problem (12).*

Proof. See Appendix B. □

6. Simulation

In this section, we justify empirically the effectiveness of the proposed algorithms by considering a wireless network with five links randomly distributed on a 10m-by-10m square area. The power gains G_{lk} are equal to d_{lk}^{-4}, where d_{lk} represents the distance between the transmitter of link k and the receiver of link k. We assume that $P_l^{max} = 1$ and $n_l = 10^{-4}$ for all $l \in L$, and consider one randomly generated realization of channel gains given by:

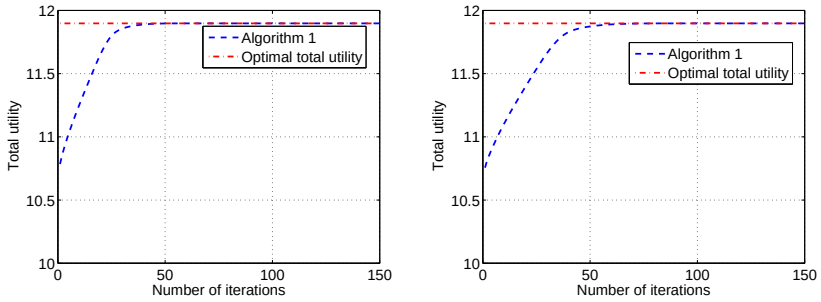
$$G = \begin{pmatrix} 0.4823 & 0.0397 & 0.0034 & 0.0007 & 0.0002 \\ 0.0338 & 0.9610 & 0.0368 & 0.0031 & 0.0007 \\ 0.0033 & 0.0397 & 0.6830 & 0.0282 & 0.0032 \\ 0.0007 & 0.0035 & 0.0368 & 0.2603 & 0.0309 \\ 0.0002 & 0.0007 & 0.0034 & 0.0282 & 0.3501 \end{pmatrix}$$

We first evaluate the convergence of Algorithm 1 with concave utility functions which were set as $U_l(\gamma_l(\mathbf{P})) = \log(\gamma_l(\mathbf{P}))$ for all $l \in L$. The convergence of the total utility over iterations is shown in Fig. 4 with different penalty factors $\rho = 30$ and $\rho = 45$, respectively. From the figure, we can see that the algorithm is able to approach the optimal utility very fast. However, with $\rho = 30$, the convergence speed is faster than that of $\rho = 45$. This is reasonable as a smaller penalty factor will cause a smaller deviation from the optimal utility compared to the higher penalty factor while the total utility is close to the neighbour of the optimal utility.

Subgradient-based algorithms are usually used to distributively solve a NUM problem with concave utility functions. For comparison we show the

Accepted manuscript

ACCEPTED MANUSCRIPT



(a) $\rho = 30$ (b) $\rho = 45$

Figure 4: The convergence of Algorithm 1 with $\rho = 30$ and $\rho = 45$.

convergence of the total utility in Fig. 5 for a subgradient-based algorithm as in (Chiang , 2005) with a stepsize of 0.01 and our algorithm 1 with $\rho = 45$. From Fig. 5, we can see that these two algorithms converge to the optimal value. However, the convergence speed of our algorithm 1 is much faster than the subgradient-based algorithm.

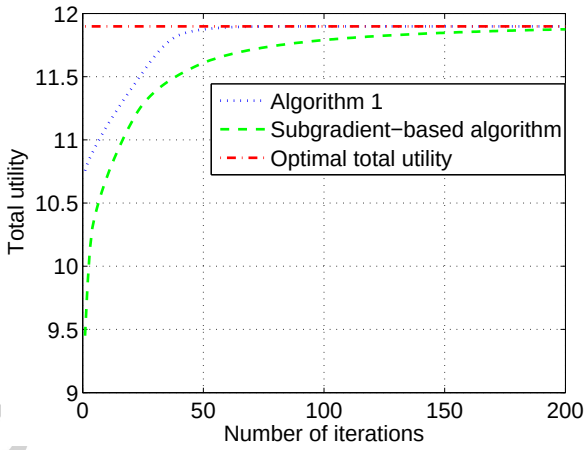


Figure 5: Comparison for the convergence speed of Algorithm 1 and subgradient-based algorithm

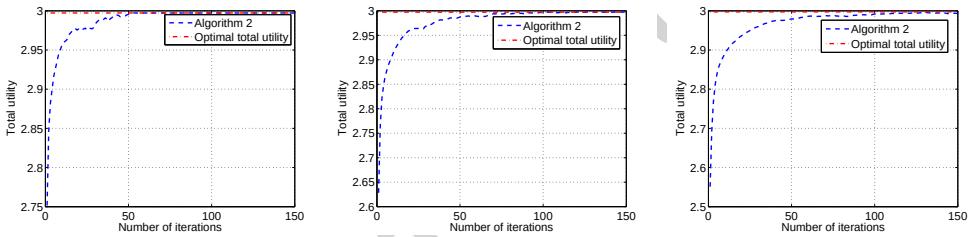
Next, we carry out simulation for the utility and convergence performance of Algorithm 2 with non-concave utility functions (Lee et al., 2006). That is, for each link l, we set the utility functions $U_l(\gamma_l(x_l)) = \frac{1}{1+\exp(\gamma_l - 7)}$. Fig. 6 shows the convergence of Algorithm 2 in terms of total utility on different

ACCEPTED MANUSCRIPT

penalty factor settings (ρ =15, 30, and 45), respectively.

From Fig. 6, we can see that the total utility converges to the optimal utility value, respectively. It can be seen from these figures that a larger ρ may lead to slower convergence rate. This is because a larger ρ will cause a smaller change to the global consensus variable $\mathbf{P}$ (cf. (25)). In our experiments, we also found that if ρ is too small, Algorithm 2 may generate many infeasible solutions. This is not surprising, because utility function $U_l(\bullet)$ is non-concave and it is easy to create infeasible points in some search area. However, a too small penalty factor is not enough to push infeasible points into feasible regions.

From the experiments, we observed that Algorithm 2 has good performance on convergence if $\rho \geq 11$. This shows consistency with Proposition 2 which requires that the penalty factor should set bigger than a constant threshold. Readers interested in the selection of the penalty factor ρ and its influence on convergence of algorithms please refer to (Boyd and Vandenberghe, 2004; Bertsekas , 1995; Huang and Yang, 2003; Chen and Chen, 2010)



(a) $\rho = 15$ (b) $\rho = 30$ (c) $\rho = 45$

Figure 6: The convergence of Algorithm 2 for different $\rho = 15, 30$ and 45.

In (Lei et al. , 2012), a distributed power control algorithm, namely the EDSPC algorithm, was presented to solve a NUM problem with non-concave utility functions. Fig. 7 presents the convergence behavior of the total utility for EDSPC algorithm and our algorithm 2 with $\rho = 45$. From Fig. 7, we can see that these two algorithms almost have the same convergence speed. From Fig. 6, we know that the convergence speed of our algorithm 2 decreases as the penalty factor ρ is increased. Therefore, with a smaller penalty factor, for example $\rho = 30$, our algorithm 2 will be faster than EDSPC.

We can also see that the total utility values in EDSPC have an oscillation when they are close to the optimal. Near the optimal, the oscillation will

ACCEPTED MANUSCRIPT

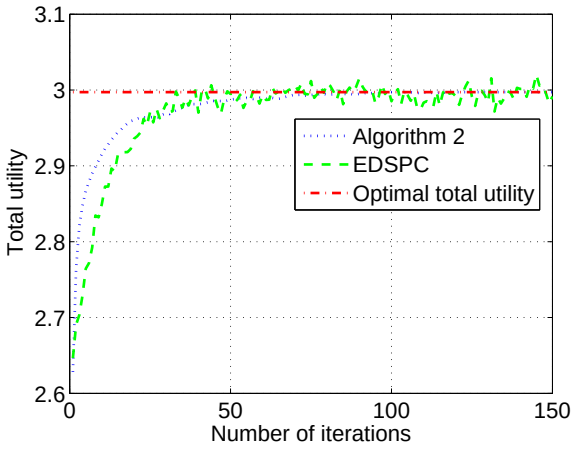


Figure 7: Comparison for the convergence speed of Algorithm 2 and EDSPC

worsen the network performance, thus, our algorithm 2 leads to a better performance compared with EDSPC.

7. Conclusion

In this paper, we studied the distributed power control problem based on NUM model with globally coupled utility functions in wireless networks. We first transformed the original power control problem into a global consensus problem by introducing multiple local slack variables to decouple the globally coupled objective function. Appealing to Lagrangian duality or the extended duality theory, we decomposed the primary problem into subproblems, and proposed two distributed power control algorithms for concave and non-concave primary problems where ADMM is adopted to solve the subproblems. The proposed two algorithms can efficiently converge to the global optimum in terms of the total network utility for concave and non-concave power control problems, respectively. Numerical results showed that the proposed algorithms are able to converge to global optimum, which is consistent with the theoretical analysis.

Appendix A. Proof of Proposition 1

Proof. Let U^* be the optimal solution of the primary problem (12) and U^k be the utilities produced by Algorithm 1 in the k-th iteration. Assume that

ACCEPTED MANUSCRIPT

$(\mathbf{x}_1^*, \ldots, \mathbf{x}_{|L|}^*, \mathbf{P}^*, \boldsymbol{\lambda}_1^*, \ldots, \boldsymbol{\lambda}_{|L|}^*)$ is a saddle point for Lagrange function L_0, and define

$$V^k = \frac{1}{\rho} \sum_{l=1}^{|L|} \|\boldsymbol{\lambda}_l(k) - \boldsymbol{\lambda}_l^*\|_2^2 + \rho \|\mathbf{P}(k) - \mathbf{P}^*\|_2^2$$

The sequel analysis is based on the following two inequalities (see (Boyd et al., 2011), inequalities A.1 and A.2,P.107):

$$V^{k+1} \leq V^k - \rho \sum_{l=1}^{|L|} ||\mathbf{x}_l(k+1) - \mathbf{P}(k+1)||_2^2 - \rho ||\mathbf{P}(k+1) - \mathbf{P}(k)||_2^2 \quad \text{(A.1)}$$

$$\begin{aligned} U^{k+1} - U^* \leq & \sum_{l=1}^{|L|} -\boldsymbol{\lambda}_l^T(k+1)(\mathbf{x}_l(k+1) - \mathbf{P}(k+1)) - \rho(\mathbf{P}(k+1) - \mathbf{P}(k))^T \times \\ & \sum_{l=1}^{|L|} -(\mathbf{x}_l(k+1) - \mathbf{P}(k+1)) + \mathbf{P}(k+1) - \mathbf{P}^* \end{aligned} \quad \text{(A.2)}$$

From inequality A.1, we have

$$\rho \sum_{k=0}^{\infty} \left(\sum_{l=1}^{|L|} \|\mathbf{x}_l(k+1) - \mathbf{P}(k+1)\|_2^2 + \rho \|\mathbf{P}(k+1) - \mathbf{P}(k)\|_2^2 \right) \leq V^0 \quad \text{(A.3)}$$

Through assumptions, it follows that V^0 is bounded. Inequality A.3 implies that $\mathbf{x}_l(k+1) - \mathbf{P}(k+1) \to 0$ and $\mathbf{P}(k+1) - \mathbf{P}(k) \to 0$ as $k \to \infty$. Therefore, the righthand side of A.2 goes to zero as $k \to \infty$, because $\mathbf{P}(k+1) - \mathbf{P}^*$ is bounded and both $\mathbf{x}_l(k+1) - \mathbf{P}(k+1)$ and $\mathbf{P}(k+1) - \mathbf{P}(k)$ go to zero. Hence, we have

$$U^{k+1} - U^* \leq 0, \text{ as } k \to \infty \quad \text{(A.4)}$$

Since $(\mathbf{x}_1^*, \ldots, \mathbf{x}_{|L|}^*, \mathbf{P}^*, \boldsymbol{\lambda}_1^*, \ldots, \boldsymbol{\lambda}_{|L|}^*)$ is a saddle point of the Lagrange function L_0, we have $L_0(\mathbf{x}_1^*, \ldots, \mathbf{x}_{|L|}^*, \mathbf{P}^*, \boldsymbol{\lambda}_1^*, \ldots, \boldsymbol{\lambda}_{|L|}^*) \leq L_0(\mathbf{x}_1(k+1), \ldots, \mathbf{x}_{|L|}(k+1), \mathbf{P}(k+1), \boldsymbol{\lambda}_1^*, \ldots, \boldsymbol{\lambda}_{|L|}^*)$.

By setting $\mathbf{x}_l^* = \mathbf{P}^*, l = 1, 2, \ldots, |L|$ in the above inequality, we have

$$U^* - U^{k+1} \leq \sum_{l=1}^{|L|} \boldsymbol{\lambda}_l^{*T}(\mathbf{x}_l(k+1) - \mathbf{P}(k+1))$$

Accepted manuscript

ACCEPTED MANUSCRIPT

The righthand side of the above inequality goes to zero as $k \to \infty$. Therefore, we have

$$U^* - U^{k+1} \leq 0, \text{ as } k \to \infty \tag{A.5}$$

Combining Eqs. A.4 and A.5, it follows that U^k converges to U^* when k tends to infinity. □

Appendix B. Proof of Proposition 2

We need the following Lemma Appendix B.1 to finish our proof. It is similar to Theorem 3.1 presented in (Chen and Chen, 2010).

Lemma Appendix B.1. *Suppose* $(\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_{|L|}^*, \boldsymbol{P}^*)$ *is a global optimal solution and* U^* *is the optimal value of the primary problem (12). Suppose* $(\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_{|L|}^*, \boldsymbol{P}^*)$ *satisfies the constraint qualification*[3]*, then there exist a set of positive* $\boldsymbol{\mu}_i^* \geq 0, 1 \leq i \leq |L|$ *such that*

$$U^* = \min_{\boldsymbol{x}_l, \boldsymbol{P}} L_\rho(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_{|L|}, \boldsymbol{P}, \boldsymbol{\mu}_1, \cdots, \boldsymbol{\mu}_{|L|}), \text{ for any } \boldsymbol{\mu_l} \geq \boldsymbol{\mu}_l^*, l \in L.$$

Proof. Let $(\overline{\mathbf{x}}_1, \ldots, \overline{\mathbf{x}}_{|L|}, \overline{\mathbf{P}}) = \lim_{k\to\infty}(\mathbf{x}_1(k), \cdots, \mathbf{x}_{|L|}(k), \mathbf{P}(k))$. For the simplicity of the notation, let Ψ denote the feasible region, and $(\mathbf{x}_1, \ldots, \mathbf{x}_{|L|}, \mathbf{P}, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{|L|})$ is abbreviated to $(\mathbf{x}_l, \mathbf{P}, \boldsymbol{\mu}_l)$. Based on Lemma Appendix B.1 and Algorithm 2, we have

$$L_\rho(\mathbf{x}_l(k), \mathbf{P}(k), \boldsymbol{\mu}_l(k)) \leq L_\rho(\mathbf{x}_l, \mathbf{P}, \boldsymbol{\mu}_l(k)), \forall \mathbf{x}_l, \mathbf{P} \in \Psi \tag{B.1}$$

and

$$\begin{aligned} U^* &= \min_{\mathbf{x}_l = \mathbf{P}, \mathbf{x}_l, \mathbf{P} \in \Psi} \sum_{l=1}^{|L|} (-U_l(\gamma_l(\mathbf{x}_l)) \\ &= \min_{\mathbf{x}_l = \mathbf{P}, \mathbf{x}_l, \mathbf{P} \in \Psi} \sum_{l=1}^{|L|} (-U_l(\gamma_l(\mathbf{x}_l) + \boldsymbol{\mu}_l^T(k)|\mathbf{x}_l - \mathbf{p}| + (\rho/2)\|\mathbf{x}_l - \mathbf{p}\|_2^2)) \\ &= \min_{\mathbf{x}_l = \mathbf{P}, \mathbf{x}_l, \mathbf{P} \in \Psi} L_\rho(\mathbf{x}_l, \mathbf{P}, \boldsymbol{\mu}_l(k)) \end{aligned} \tag{B.2}$$

[3]For the definition of the constraint qualification, please refer to (Chen and Chen, 2010)

Accepted manuscript

ACCEPTED MANUSCRIPT

Therefore, by taking the infimum of the righthand side of B.1 and combining with B.2, we have

$$L_\rho(\mathbf{x}_l(k), \mathbf{P}(k), \boldsymbol{\mu}_l(k)) = \sum_{l=1}^{|L|}(-U_l(\gamma_l(\mathbf{x}_l(k)) + \boldsymbol{\mu}_l^T(k)|\mathbf{x}_l(k) - \mathbf{p}(k)| + (\rho/2)\|\mathbf{x}_l(k) - \mathbf{p}(k)\|_2^2)) \leq U^* \quad \text{(B.3)}$$

The sequence $\boldsymbol{\mu}_l(k)$ is bounded and thus we can assume $\boldsymbol{\mu}_l(k) \to \overline{\boldsymbol{\mu}}_l$. By taking the limit superior in B.3 and using the continuity of utility functions in the primary problem (12), we have

$$\sum_{l=1}^{|L|}(-U_l(\gamma_l(\overline{\mathbf{x}}_l) + \overline{\boldsymbol{\mu}}_l^T|\overline{\mathbf{x}}_l - \overline{\mathbf{p}}| + \limsup_{k\to\infty}(\rho/2)\|\mathbf{x}_l(k) - \mathbf{p}(k)\|_2^2) \leq U^* \quad \text{(B.4)}$$

Since $\limsup_{k\to\infty} \|\mathbf{x}_l(k) - \mathbf{p}(k)\|_2^2 \geq 0$, it follows that $\mathbf{x}_l(k) - \mathbf{p}(k) \to 0$ and $\overline{\mathbf{x}}_l - \overline{\mathbf{p}} = 0$. Otherwise, the left hand side of B.4 would be bigger than U^* for a sufficiently large ρ.

From B.4, we obtain $\sum_{l=1}^{|L|}(-U_l(\gamma_l(\overline{\mathbf{x}}_l))) \leq U^*$. This proves that $(\overline{\mathbf{x}}_1, \ldots, \overline{\mathbf{x}}_{|L|}, \overline{\mathbf{P}})$ is a globally optimal solution. □

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China (Grant No. 61072051 and No. 61202470), the key science and technology project of Wuhan (Grant NO.2014010202010108) and Doctoral Program of Higher Education of China (Grant No. 20110141120046).

References

Xiao Mingbo, Ness B. Shroff, and Edwin K.P. Chong. A utility-based power control scheme in wireless cellular systems. IEEE/ACM tranactions on networking 2003;11(2):210-221.

Chiang M, Low S H, Calderbank A R, Doyle J C. Layering as optimization decomposition: A mathematical theory of network architectures. Proceedings of the IEEE 2007;95(1):255-312.

Accepted manuscript

ACCEPTED MANUSCRIPT

Wei Zheng, Tao Su, Haijun Zhang, Wei Li, Xiaoli Chu, Xiangming Wen. Distributed power optimization for spectrum-sharing femtocell networks: A fictitious game approach. Journal of Network and Computer Applications 2014;37:315-322.

Liping Qing, Yingjun Zhang, Jianwei Huang. MAPEL: achieving global optimality for a nonconvex wireless power control problem. IEEE transactions on wireless communications 2009;8(3):1553-1563.

Hussein Al-Shatri, Tobias Weber. Achieving the Maximum Sum Rate Using DC Programming in Cellular Networks. IEEE Transactions on signal processing 2012;60(3):1331-1341.

Campos-Delgado D U, Luna-Rivera J M, Martinez-Lopez F J. A suboptimal LQ power control algorithm for a CDMA wireless system. In: IEEE wireless communications and networking conference, WCNC 2008; 2008. p.380-385.

Hande P, Rangan S, Chiang M, Wu X. Distributed uplink power control for optimal SIR assignment in cellular data networks. IEEE/ACM Transactions on Networking 2008;16(6):1430-1443.

Qian L P, Zhang Y J, Chiang M. Distributed nonconvex power control using Gibbs sampling. IEEE Transactions on Communications 2012;60(12):3886-3898.

Lei Yang, Yalin E Sagduyu, Junshan Zhang, Jason H Li. Distributed stochastic power control in ad hoc networks: a nonconvex optimization case. EURASIP Journal on Wireless Communications and Networking 2012; 2012:231. doi:10.1186/1687-1499-2012-231

Constantino T. Possible generalization of BoltzmaanGibbs statistics. Journal of statistical physics 1988;52:479487.

Kirkpatrick S, Gelatt C D, Vecchi J M P. Optimization by simulated annealing. Science 1983;220(4598):671-680.

Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine Learning 2011;3(1):1-122.

Accepted manuscript

ACCEPTED MANUSCRIPT

Shenker S. Fundamental design issues for the future Internet. IEEE Journal on Selected Areas in Communications 1995;13(7):1176-1188.

Liao R, Campbell A. A utility-based approach to quantitative adaptation in wireless packet networks. ACM Journal on Wireless Networks 2001;7(5):541-557.

Qian L, Zhang Y J, and Huang J W. MAPEL: achieving global optimality for a non-convex power control problem, IEEE Transaction on Wireless Communication 2009;8(3):1553-1563.

Nedić A, Ozdaglar A. Cooperative distributed multi-agent optimization. Convex optimization in signal processing and communications, Cambridge University Press; 2010. p. 340-386

Eckstein J, Bertsekas D. P. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. Mathematical programming 1992;55:293-318.

Gabay D. Applications of the method of multipliers to variational inequalities. Augmented Lagrangian methods: Applications to the solution of Boundary-value problems, North-Holland: Amsterdam, 1983.

Bertsekas D P. *Nonlinear Programming*. Blemont, MA, Athena Scientific, 1995.

Rockafellar R. T, Wets R. *Variational Aanlysis*. Springer-Verlag, New YorK, 1998.

Rubinnov A M, Huang X X, Yang X Q. The zero duality gap property and lower semicontinuity of the perturbation. Mathematics of operations research 2002;27(4):775-791.

Nedić A, Ozdaglar A. A geometric framework for nonconvex optimization duality using augmented Lagrangian functions. Journal of global optimization 2008;40(4):545-573.

Yixin Chen, Minmin Chen. Extended duality for nonlinear programming. Computational optimization and applications 2010;47(1):33-59.

Accepted manuscript

ACCEPTED MANUSCRIPT

M. Chiang, Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control. IEEE Journal on Selected Areas in Communications 2005;23(1):104116.

Jang-Won Lee, Ravi R. Mazumdar, and Ness B. Shroff Opportunistic Power Scheduling for Dynamic Multi-server Wireless Systems. IEEE transactios on wireless communications 2006;5(6):1506-1515.

Boyd S, Vandenberghe L. *Convex Optimization*. Cambridge University Press: Cambridge, United Kingdom, 2004

X. Huang and X. Yang. A unified augmented Lagrangian approach to duality and exact penalization. Mathematics of operations research 2003;28(3):533-552.

Accepted manuscript