

# On distributed power control and transceiver optimization in wireless networks\*

[Invited Paper]

Sławomir Stańczak, Michał Kaliszan and Mario Goldenbaum

Heinrich-Hertz-Lehrstuhl für Informationstheorie und theoretische Informationstechnik  
Technische Universität Berlin, Einsteinufer 25, 10587 Berlin, Germany  
{slawomir.stanczak, michal.kaliszan}@hhi.fraunhofer.de,  
mario.goldenbaum@mk.tu-berlin.de

## ABSTRACT

This invited paper outlines some recent results on the max-min SIR balancing problem in wireless networks in which power control and beamforming are the only mechanisms for resource allocation and interference management. In addition, we describe several potential extensions and improvements to existing algorithmic solutions, as well as prove the convergence of a distributed algorithm for joint power control and receive beamforming to a global optimum of the max-min SIR balancing problem. Finally we briefly discuss a possibility of how to incorporate the optimization of transmit beamformers.

## Categories and Subject Descriptors

C.2.1 [Computer Systems Organization]: Computer-Communication Networks—*Network Architecture and Design*

## General Terms

Theory, Algorithms

## Keywords

Power Control, Beamforming, Max-Min Fairness, Utility Maximization, Saddle-Point Characterization, Algorithmic Solutions, Distributed Implementation

## 1. INTRODUCTION

Max-min fairness is the most common notion of fairness [7]. The max-min fair approach consists in treating all users as fairly as possible by making the rates allocated to the users as equal as possible. Under certain assumption, max-min

\*The work was supported by the German Research Foundation (DFG) under grant STA864/3-1 and by the German Federal Ministry of Education and Research (BMBF) under grant 01BU920.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

VALUETOOLS 2011, May 16-20, Paris, France

Copyright © 2011 ICST 978-1-936968-09-1

DOI 10.4108/icst.valuetools.2011.246069

fairness can be achieved by solving the so-called max-min SIR balancing problem, where the objective is to maximize the minimum signal-to-interference ratio (SIR). QoS-based power control is closely related and aims at satisfying given desired SIR levels (SIR targets) with a minimum total transmit power. Both approaches have been extensively studied and are fairly well understood [1, 36, 2, 63, 64, 19, 61, 62, 22, 26, 4, 39, 27, 58, 5, 65, 15, 17]. See also [14] for joint power control, scheduling and routing. Optimal power allocations in the sense of QoS-based power control can be found by means of iterative algorithms that allow distributed implementation, provided that the SIR targets are feasible (see for instance [19, 61, 5] and [25] for combined power control and cell-site selection). These strategies are in particular appropriate for applications that generate inelastic traffic, and therefore do not tolerate large delays.

In contrast, utility maximization problems aiming at maximizing some aggregate utility function of link rates (or other quantities) appears to fit better the needs and characteristics of some wireless communications applications with elastic traffic. Utility-based strategies implicitly use the relative delay tolerance of data applications as well as the network and channel dynamics to improve the network performance. At the same time, the use of monotonically increasing and strictly concave utility functions ensures the desired degree of (link-layer and end-to-end) fairness [56, 33, 38]. For these reasons, utility-based approaches to resource allocation and interference management in wireless networks have attracted a great deal of attention over the last decade. Theoretical work on utility-based power control or, more generally, cross-layer design includes [23, 45, 59, 32, 40, 12, 18, 13, 42, 55, 41, 53, 24, 57, 31, 6, 11]. Further references can be found in the papers listed above and in [21]. The power control problem has been also analyzed within the framework of game theory. See for instance [29, 30, 28, 3, 46].

In this paper, we consider a power-controlled wireless network in which power control and beamforming are the only mechanisms for resource allocation.<sup>1</sup> If the beamformers of all links are fixed, max-min fairness is a power control problem and the max-min fair rate allocation is achieved by the so-called max-min fair power allocation. This power allocation is usually obtained by solving the max-min SIR

<sup>1</sup>There is no scheduling and all users (transmitter-receiver pairs) share a common bandwidth.

balancing problem, which is a widely studied resource allocation problem for wireless networks (see, for instance, [65], [60], [54]) and references therein). The main challenge is to solve the max-min SIR balancing problem in a distributed manner.

This paper summarizes some of our recent results on the max-min SIR balancing problem in noisy wireless networks with general power constraints [54, 51, 50, 16, 52]. In particular, in [51, 52], we established a connection between the max-min SIR-balancing power control problem and the utility-based power control problem. A similar connection is known in the noiseless case [54; 49, Section 5.9] and constitutes the starting point for the analysis in [9]. The results lead to a saddle point characterization of certain aggregate utility functions [54, 51, 50], which was a basis in [52] for the development of a saddle-point algorithm converging to the solution of the max-min SIR balancing problem. The convergence rate and the initial convergence speed of the algorithm proposed in [52] are not satisfactory and the convergence behavior strongly depends on the choice of the step size. In this paper, therefore, we suggest using the max-min methods of [57] to significantly improve the convergence speed of the saddle-point algorithm. The power control iteration of [57] can be classified as a conditional Newton iteration (or reduced Lagrange-Newton iteration) that operates on a modified Lagrangian function and relies a specific min-max formulation of the power control problem. The combination of quadratic convergence and amenability to decentralized implementation makes the conditional Newton iteration predestined for efficient online application in real-world power control.

Further we combine the power control iteration with optimal receive beamforming and prove that the proposed iteration converges to a global optimal solution of the max-min SIR balancing problem over the joint space of transmit powers and receive beamformers. Although the approach is different from that in [9], the result in some sense extends the main statement of [9] to noisy channels with individual power constraints. In contrast to [16], the algorithm proposed in this paper does not exploit the knowledge of principal left and right eigenvectors of some nonnegative matrices, and therefore the iteration can be relatively efficiently implemented in decentralized wireless networks. The last section of the paper briefly discusses the possibility of incorporating transmit beamformers into the joint optimization process. The simulation results suggest huge potential for performance gains.

## 2. DEFINITIONS AND PROBLEM STATEMENT

We consider a wireless network with an established network topology, in which  $K \geq 2$  users (point-to-point logical links) share a common wireless spectrum and transmit their independent data concurrently. Let  $\mathcal{K} = \{1, \dots, K\}$  denote the set of all users and let  $\mathbf{p} = (p_1, \dots, p_K) \geq 0$  be the *power vector* or *power allocation*, where  $p_k$  is the transmit power of user  $k \in \mathcal{K}$ . Due to power constraints, we have  $\mathbf{p} \in \mathbf{P}$  where  $\mathbf{P}$  is a compact convex set of all feasible power allocations (called *feasible power region*). For brevity, in this paper,  $\mathbf{P}$

is assumed to be<sup>2</sup>

$$\mathbf{P} = \{\mathbf{p} \in \mathbb{R}_+^K : \forall k \in \mathcal{K} p_k \leq \hat{p}_k\} = \{\mathbf{p} \in \mathbb{R}_+^K : \mathbf{p} \leq \hat{\mathbf{p}}\} \quad (1)$$

for a given vector  $\hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_K)$  with individual power constraints on each link.

The main figure of merit is the SIR at the output of each receiver given by

$$\text{SIR}_k(\mathbf{p}) = \frac{p_k}{I_k(\mathbf{p})}, k \in \mathcal{K}, \quad (2)$$

where  $I_k : \mathbb{R}_+^K \rightarrow \mathbb{R}_{++}$  is a given interference function. The interference functions are assumed to fulfill the following conditions:

- (A.1) For each  $k$ , we have  $I_k(\mathbf{p}) > 0, \mathbf{p} \geq 0$
- (A.2)  $\nabla_{\mathbf{p}} I_k(\mathbf{p}), k \in \mathcal{K}$ , exists and the partial derivatives are continuous functions on  $\mathbb{R}_{++}^K$ . Moreover,  $\partial I_k(\mathbf{p}) / \partial p_l \geq 0$  for each  $k, l \in \mathcal{K}$  and all  $\mathbf{p} > 0$ .
- (A.3)  $\min_{k \in \mathcal{K}} \text{SIR}_k(\mathbf{p})$  attains its maximum on  $\mathbf{P}$ .

Sometimes (see Sect. 4.1) it is also required that the interference functions are convex functions of  $\mathbf{s} = \log(\mathbf{p}), \mathbf{p} \in \mathbb{R}_{++}^K$ . If the Hessian of  $I_k(e^{\mathbf{s}})$  exists and its entries are continuous, then this is equivalent to saying that

- (A.4)  $\nabla_{\mathbf{s}}^2 I_k(e^{\mathbf{s}})$  is positive semidefinite for every  $\mathbf{s} \in \mathbb{R}^K$ .

An important example of interference functions satisfying the above conditions are affine interference functions that are of the form:

$$(A.5) \quad I_k(\mathbf{p}) = (\mathbf{V}\mathbf{p} + \boldsymbol{\sigma})_k = \sum_{l=1}^K v_{k,l} p_l + \sigma_k^2.$$

Here and hereafter,  $\mathbf{V} := (v_{k,l}) \in \mathbb{R}_+^{K \times K}$  is the *gain matrix*,  $v_{k,l} = V_{k,l}/V_{k,k}$  if  $l \neq k$  and 0 if  $l = k$  where  $V_{k,l} \geq 0$  with  $V_{k,k} > 0$  is the attenuation of the power from transmitter  $l$  to receiver  $k$ . The  $k$ th entry of  $\boldsymbol{\sigma} := (\sigma_1^2, \dots, \sigma_K^2)$  is equal to the noise variance at the output of receiver  $k \in \mathcal{K}$  divided by  $V_{k,k} > 0$ .<sup>3</sup> Unless otherwise stated, the interference functions are assumed to be affine and therefore are of the form given by (A.5). Moreover, if the interference functions are affine, then, unless otherwise stated, we assume that

- (A.6) the gain matrix  $\mathbf{V} \geq 0$  is *irreducible*.

In words, (A.6) means that the network is entirely coupled by interference. This is equivalent to saying that under certain optimal conditions, the change of any transmit power has an impact on each SIR.

Now we are in a position to define a max-min SIR power allocation.

<sup>2</sup> $\mathbb{R}_+, \mathbb{R}_{++}$  are nonnegative and positive reals, respectively.  
<sup>3</sup>Thus  $\sigma_k^2$  is a normalized noise variance so that the numerator of (2) contains only the transmit power.

*Definition 1.* Suppose that (A.1)–(A.3) hold. The power vector  $\bar{\mathbf{p}}$  is said to be a *max-min SIR power vector/allocation* if

$$\min_{k \in \mathcal{K}} \text{SIR}_k(\bar{\mathbf{p}}) = \max_{\mathbf{p} \in \mathbf{P}} \min_{k \in \mathcal{K}} \text{SIR}_k(\mathbf{p}). \quad (3)$$

The problem is referred to as the *max-min SIR balancing problem*. A simple examination of the definition shows that  $\bar{\mathbf{p}}$  is a positive vector. Further, it is pointed out that there may be multiple max-min SIR power allocations, which stands in contrast to the max-min fair power allocation defined as follows [7, p. 526].

*Definition 2.*  $\bar{\mathbf{p}}' \in \mathbf{P}$  is said to be a *max-min fair power allocation* if any  $\text{SIR}_k(\bar{\mathbf{p}}')$  cannot be increased without decreasing some  $\text{SIR}_l(\bar{\mathbf{p}}')$ ,  $l \neq k$ , which is smaller than or equal to  $\text{SIR}_k(\bar{\mathbf{p}}')$ .

Let  $\bar{\mathbf{P}} \subset \mathbf{P}$  be the set of all max-min SIR power allocations. The following proposition states a sufficient condition for a max-min SIR power vector to be unique and equal to the max-min fair power vector.

**PROPOSITION 1.** *If (A.6) is true, then  $|\bar{\mathbf{P}}| = 1$  and  $\bar{\mathbf{p}} = \bar{\mathbf{p}}' \in \mathbf{P}_+ := \mathbf{P} \cap \mathbb{R}_{++}^K$ .*

Since  $\bar{\mathbf{p}} > 0$ , we can focus on  $\mathbf{P}_+$ . Considering this and (1),  $\bar{\mathbf{p}}$  defined by (3) can be written as

$$\bar{\mathbf{p}} = \arg \max_{\mathbf{p} > 0} \min_{k \in \mathcal{K}} \text{SIR}_k(\mathbf{p}) \quad \text{s.t.} \quad \max_{k \in \mathcal{K}} g_k(\mathbf{p}) \leq 1, \quad (4)$$

where

$$g_k(\mathbf{p}) := 1/\hat{p}_k \mathbf{c}_k^T \mathbf{p} = p_k/\hat{p}_k \leq 1, \quad k \in \mathcal{K}, \quad (5)$$

and  $\mathbf{c}_k = (0, \dots, 0, 1, 0, \dots, 0) \in \{0, 1\}^K$  is a vector with 1 at the  $k$ th position and zeros elsewhere.

Our approach to the max-min SIR balancing problem (3) is based on a certain saddle point characterization of a weighted sum of utilities of the SIRs (see Section 3 below), why we use  $\phi$  to denote the utility function assigned to each link and assume the following.

(A.7) Let  $\phi: \mathbb{R}_{++} \rightarrow \mathbb{Q} \subseteq \mathbb{R}$  be any continuous and strictly increasing utility function. Moreover,  $\phi$  is a twice continuously differentiable function.

(A.8)  $\phi(e^x)$ ,  $x \in \mathbb{R}$ , is a concave function.

Examples of functions satisfying (A.7) and (A.8) are  $x \mapsto \log(x)$ ,  $x > 0$ , and  $x \mapsto -1/x^n$ ,  $n \geq 1$ ,  $x > 0$ . In order to match the usual derivations in optimization theory, in this paper, we will minimize the objective function with respect to  $\mathbf{p}$ . To this end, we define the function  $\theta: \mathbb{R}_{++} \rightarrow \mathbb{R}$  to be

$$\theta(x) := -\phi(x), \quad x > 0. \quad (6)$$

Obviously, (A.7) implies that  $\theta$  is a strictly decreasing function. By strict decreasingness, we have  $\theta(\max_{k \in \mathcal{K}} \text{SIR}_k(\mathbf{p})) = \max_{k \in \mathcal{K}} \theta(\text{SIR}_k(\mathbf{p}))$  for all  $\mathbf{p} > 0$ . Thus, as  $\bar{\mathbf{p}} \in \mathbf{P}_+$ , we have

$$\bar{\mathbf{p}} = \arg \min_{\mathbf{p} \in \mathbf{P}_+} \max_{k \in \mathcal{K}} \theta(\text{SIR}_k(\mathbf{p})). \quad (7)$$

### 3. SADDLE POINT CHARACTERIZATION

The basic idea for solving (4) is to rewrite the problem as a utility maximization problem. To be more precise, let

$$\Pi_K := \{\mathbf{x} \in \mathbb{R}_+^K : \|\mathbf{x}\|_1 = 1\}$$

and  $\Pi_K^+ = \Pi_K \cap \mathbb{R}_{++}^K$ . We define the objective function of interest  $G: \Pi_K^+ \times \mathbf{P}_+ \rightarrow \mathbb{R}$  to be

$$G(\mathbf{u}, \mathbf{p}) := - \sum_{k \in \mathcal{K}} u_k \phi(\text{SIR}_k(\mathbf{p})) = \sum_{k \in \mathcal{K}} u_k \theta\left(\frac{p_k}{I_k(\mathbf{p})}\right), \quad (8)$$

where  $\mathbf{u} = (u_1, \dots, u_K) \in \Pi_K^+$  is a positive weight vector that has to be determined. We see that any power vector minimizing  $G(\mathbf{u}, \mathbf{p})$  with respect to  $\mathbf{p} \in \mathbf{P}$  for some given weight vector  $\mathbf{u} > 0$  is optimal in the sense of the utility maximization problem considered in [54].

Now if the weight vector is chosen appropriately, then we can solve our problem by minimizing (8) with respect to  $\mathbf{p} \in \mathbf{P}$ . The problem of choosing the weight vector has been solved in [51]. To be precise, let  $\mathbf{B}^{(k)} \in \mathbb{R}_+^{K \times K}$  (for each  $k \in \mathcal{K}$ ) be defined to be

$$\mathbf{B}^{(k)} := \mathbf{V} + \frac{1}{\hat{p}_k} \boldsymbol{\sigma} \mathbf{c}_k^T. \quad (9)$$

Notice that by (A.6),  $\mathbf{B}^{(k)}$  is irreducible for each  $k \in \mathcal{K}$ . As a consequence, the Perron-Frobenius theory [48, 37, 35] implies that the matrix has positive left and right eigenvectors (denoted by  $\mathbf{y}$  and  $\mathbf{x}$ , respectively) associated with the spectral radius  $\rho(\mathbf{B}^{(k)})$  which is a simple eigenvalue of  $\mathbf{B}^{(k)}$ :

$$\rho(\mathbf{B}^{(k)}) \mathbf{x} = \mathbf{B}^{(k)} \mathbf{x}, \mathbf{x} > 0 \quad \rho(\mathbf{B}^{(k)}) \mathbf{y} = (\mathbf{B}^{(k)})^T \mathbf{y}, \mathbf{y} > 0.$$

Moreover, the eigenvectors are unique up to a multiplicative (scalar) factor. They are called principal left and right eigenvectors.

Now it was shown in [51] that the weight vector we are searching for can be computed from  $\mathbf{B}^{(k)}$  for some  $k$  and the index  $k$  must be a member of a set  $\mathcal{K}_0$  which is defined as follows:

$$\mathcal{K}_0 := \left\{ k_0 \in \mathcal{K} : k_0 = \arg \max_{k \in \mathcal{K}} \rho(\mathbf{B}^{(k)}) \right\}. \quad (10)$$

This leads us to the following result.

**PROPOSITION 2.** *Let  $k_0 \in \mathcal{K}_0$  be arbitrary and let*

$$\rho(\mathbf{B}^{(k_0)}) \mathbf{x} = \mathbf{B}^{(k_0)} \mathbf{x} \quad \rho(\mathbf{B}^{(k_0)}) \mathbf{y} = (\mathbf{B}^{(k_0)})^T \mathbf{y}$$

*so that  $\mathbf{x} > 0$  and  $\mathbf{y} > 0$  are principal right and left eigenvectors of the matrix  $\mathbf{B}^{(k_0)}$  with  $\mathbf{y}^T \mathbf{x} = 1$ . Then,*

$$\bar{\mathbf{p}} = \arg \min_{\mathbf{p} \in \mathbf{P}} G(\mathbf{w}, \mathbf{p}) \quad (11)$$

where

$$\mathbf{w} = \mathbf{y} \circ \mathbf{x} > 0 \quad \|\mathbf{w}\| = 1. \quad (12)$$

In words, if the weight vector is chosen to be equal to the Hadamard product of the principal left and right eigenvectors of  $\mathbf{B}^{(k_0)}$  for some  $k_0 \in \mathcal{K}_0$ , then the minimum of  $G$  (as a function of  $\mathbf{p}$ ) on the set  $\mathbf{P}$  is attained for the max-min SIR power allocation. Note that the choice of the weight vector is independent of the choice of the utility function,

provided that (A.7) and (A.8) are satisfied. Furthermore, if  $\mathbf{w}$  is defined by (12), then

$$\begin{aligned} \min_{\mathbf{p} \in \mathbf{P}} G(\mathbf{w}, \mathbf{p}) &= \min_{\mathbf{p} \in \mathbf{P}} \sum_{k \in \mathcal{K}} w_k \theta\left(\frac{p_k}{I_k(\mathbf{p})}\right) \\ &= \min_{\mathbf{p} \in \mathbf{P}} \max_{k \in \mathcal{K}} \theta\left(\frac{p_k}{I_k(\mathbf{p})}\right). \end{aligned} \quad (13)$$

The constraint  $\mathbf{p} \in \mathbf{P}$  can be replaced by  $\max_k g_k(\mathbf{p}) \leq 1$  and in the minimum we have  $g_k(\bar{\mathbf{p}}) = 1, k \in \mathcal{K}_0$ .

Unfortunately the problem in (11) may not be meaningful for an algorithmic approach unless there is a method for an efficient computation of the Hadamard product without knowing the principal eigenvectors in advance (whose computation in a distributed wireless networks is not a trivial task). This is clarified by the following proposition.

**PROPOSITION 3.** *If (A.6) is satisfied, then  $\bar{\mathbf{p}}$  is the unique max-min SIR power allocation if and only if  $\bar{\mathbf{p}}$  is the principle right eigenvector of  $\mathbf{B}^{(k_0)}$  (associated with the spectral radius  $\rho(\mathbf{B}^{(k_0)})$ ) for some  $k_0 \in \mathcal{K}_0$  normalized such that  $g_{k_0}(\bar{\mathbf{p}}) = 1$ :*

$$\rho(\mathbf{B}^{(k_0)})\bar{\mathbf{p}} = \mathbf{B}^{(k_0)}\bar{\mathbf{p}}, \quad g_{k_0}(\bar{\mathbf{p}}) = 1 \quad k_0 \in \mathcal{K}_0.$$

Thus if the principal right eigenvector is known, then there is no problem to solve because this vector provides the solution to our problem. Note that although the cardinality of  $\mathcal{K}_0$  can be larger than one, the principal eigenvectors of  $\mathbf{B}^{(k_0)}$  for different  $k_0 \in \mathcal{K}_0$  are the same due to the assumption of irreducibility of the gain matrix.

Unfortunately, the computation of the weight vector  $\mathbf{w}$  defined by (12) is not amenable to distributed implementation. It is illusive to assume that the gain matrix  $\mathbf{V}$  is known at some nodes.

When developing distributed power control algorithms for achieving the max-min fairness, the characterization of a saddle point of the function  $G : \Pi_K^+ \times \mathbf{P} \rightarrow \mathbb{R}$  turns out to be useful.

**PROPOSITION 4.** *Suppose that (A.6)–(A.7) hold. Then, for each  $k_0 \in \mathcal{K}_0$  defined by (10), we have*

$$\begin{aligned} \theta(1/\rho(\mathbf{B}^{(k_0)})) &= \max_{\mathbf{u} \in \Pi_K} \min_{\mathbf{p} \in \mathbf{P}} G(\mathbf{u}, \mathbf{p}) \\ &= \min_{\mathbf{p} \in \mathbf{P}} \max_{\mathbf{u} \in \Pi_K} G(\mathbf{u}, \mathbf{p}) \end{aligned} \quad (14)$$

and  $(\mathbf{w}, \bar{\mathbf{p}})$  is the unique saddle point in  $\Pi_K \times \mathbf{P}_+$ .

Intuitively, the results implies that  $\mathbf{u} = \mathbf{w}$  with  $\mathbf{w}$  defined by (12) is the worst-case weight vector in the sense that it leads to the worst performance in terms of the aggregate utility function.

Finally we point out that Proposition 4 is substantially related to results of the seminal paper of Friedland and Karlin [20].

## 4. POWER CONTROL ALGORITHM

Now the goal is to use the characterization of Proposition 4 to design an iterative saddle-point algorithm that *simultaneously* maximizes  $G(\mathbf{u}, \mathbf{p})$  with respect to  $\mathbf{u} \in \Pi_K$  and minimizes this function over the feasible power region  $\mathbf{P}$ . Due to Theorem 4, the algorithm will converge to a saddle point  $(\mathbf{u}^*, \mathbf{p}^*)$  of  $G(\mathbf{u}, \mathbf{p})$ , which is a unique point in  $\Pi_K \times \mathbf{P}$ . For brevity, we assume in this section that  $\phi$  (and with it  $\theta$ ) is *continuously differentiable*.

The function  $G(\mathbf{u}, \mathbf{p})$  is in general not convex in  $\mathbf{p} \in \mathbf{P}$  but by Theorem 4, we have  $\mathbf{p}^* > 0$  (and  $\mathbf{u}^* > 0$ ). Therefore, by [54, Section 6.2],

$$G_e(\mathbf{u}, \mathbf{s}) := G(\mathbf{u}, e^{\mathbf{s}}) = \sum_{k \in \mathcal{K}} u_k \theta\left(\frac{e^{s_k}}{I_k(e^{\mathbf{s}})}\right)$$

is well-defined on  $\Pi_K \times \mathbf{S}$  with

$$\mathbf{S} := \{\mathbf{s} \in \mathbb{R}^K : \mathbf{s} = \log(\mathbf{p}), \mathbf{p} \in \mathbf{P}_+\}$$

and convex with respect to  $\mathbf{s} \in \mathbf{S}$ . So,  $G_e(\mathbf{u}, \mathbf{s})$  is a concave-convex function on  $\Pi_K \times \mathbf{S}$ , meaning that it is concave (and convex) with respect to  $\mathbf{u}$  and convex in  $\mathbf{s} \in \mathbf{S}$ . Now since the logarithmic function is a bijective function that map  $\mathbb{R}_{++}$  onto  $\mathbb{R}$ , we can restate Proposition 4 using the logarithmic power vectors. To this end, we define

$$\bar{\mathbf{s}} := \log(\bar{\mathbf{p}}) \in \mathbf{S}. \quad (15)$$

With this definition in hand, it follows directly from (14) that

$$\begin{aligned} \theta(1/\rho(\mathbf{B}^{(k_0)})) &= \max_{\mathbf{u} \in \Pi_K} \min_{\mathbf{s} \in \mathbf{S}} G_e(\mathbf{u}, \mathbf{s}) \\ &= \min_{\mathbf{s} \in \mathbf{S}} \max_{\mathbf{u} \in \Pi_K} G_e(\mathbf{u}, \mathbf{s}) \end{aligned} \quad (16)$$

and  $(\mathbf{w}, \bar{\mathbf{s}})$  is the unique saddle point in  $\Pi_K \times \mathbf{S}$ .

Now using these definitions, the power control problem can be stated as follows:

$$\begin{aligned} \min_{\mathbf{s}} \max_{\mathbf{u}} G_e(\mathbf{u}, \mathbf{s}) &= \max_{\mathbf{u}} \min_{\mathbf{s}} G_e(\mathbf{u}, \mathbf{s}) \\ \text{s.t.} &\begin{cases} \forall k \in \mathcal{K} \quad e^{s_k} - \hat{p}_k \leq 0 \\ \sum_{k \in \mathcal{K}} u_k - 1 = 0 \end{cases} \end{aligned} \quad (17)$$

In order to obtain Newton-like algorithms that are amenable to distributed implementation, we proceed essentially as in [57] to introduce a new interference variable by splitting each SIR, say SIR of link  $k$ , into the power control variables  $e^{s_k}$  and the interference variable  $J_k$  such that  $J_k \leq I_k(e^{\mathbf{s}}), k \in \mathcal{K}$ . Using the telescope variable  $\mathbf{t} = (t_1, \dots, t_K)$  to separate the constraint inequalities for different variables, we obtain

$$\begin{aligned} \min_{\mathbf{s}, \mathbf{t}} \max_{\mathbf{u}, \mathbf{J}} \sum_{k \in \mathcal{K}} u_k \theta\left(\frac{e^{s_k}}{J_k}\right) \\ \text{s.t.} &\begin{cases} \forall k \in \mathcal{K} \quad e^{s_k} - \hat{p}_k \leq 0 \\ \sum_{k \in \mathcal{K}} u_k - 1 = 0 \\ \forall k \in \mathcal{K} \quad J_k - t_k \leq 0 \\ \forall k \in \mathcal{K} \quad I_k(e^{\mathbf{s}}) - t_k = 0 \end{cases} \end{aligned} \quad (18)$$

where  $\mathbf{J} = (J_1, \dots, J_K)$ . It is important to emphasize that in general, (17) and (18) are not equivalent. In fact, even if (A.7)–(A.8) are satisfied, it is easy to see that the objective function in (18) is not necessarily concave in  $J_k, k \in \mathcal{K}$ . For

this reason, it may be not possible to change the order of the min and max operators in (18) and the associated Karush-Kuhn-Tucker conditions are only necessary conditions for the optimum. An immediate consequence of this modification is the lack of important properties that ensure global convergence of primal-dual algorithms with the class of utility functions satisfying Conditions (A.7)–(A.8) [57] (see also Sect. 4.1).

The classical linear Lagrangian function for the modified problem (18) takes therefore the form

$$\begin{aligned}\bar{L}(\mathbf{z}) &= \bar{L}(\mathbf{s}, \mathbf{J}, \mathbf{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}^J, \boldsymbol{\lambda}^I, \mathbf{t}, \omega) \\ &= \sum_{k \in \mathcal{K}} u_k \theta\left(\frac{e^{s_k}}{J_k}\right) \\ &\quad + \sum_{k \in \mathcal{K}} \mu(e^{s_k} - \hat{p}_k) \\ &\quad + \sum_{k \in \mathcal{K}} \lambda_k^J (J_k - t_k) \\ &\quad + \sum_{k \in \mathcal{K}} \lambda_k^I (I_k(e^{\mathbf{s}}) - t_k) \\ &\quad + \omega \left( \sum_{k \in \mathcal{K}} u_k - 1 \right)\end{aligned}\tag{19}$$

with  $\mathbf{z} := (\mathbf{s}, \mathbf{J}, \mathbf{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}^J, \boldsymbol{\lambda}^I, \mathbf{t}, \omega) \in \mathbb{R}^{2K} \times \mathbb{R}_+^{3K} \times \mathbb{R}^{2K} \times \mathbb{R}$ . Note that the primal variable  $\mathbf{u}$  is in this formulation constrained to be nonnegative. We keep this constraint but point out that an additional dual variable can be introduced to deal with this constraint.

To get rid of all the nonnegativity constraints on the dual variables, one can consider the following non-linear Lagrangian function [57] which is a modification of the classical linear Lagrangian defined by (19):

$$\begin{aligned}L(\mathbf{z}) &= L(\mathbf{s}, \mathbf{J}, \mathbf{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}^J, \boldsymbol{\lambda}^I, \mathbf{t}, \omega) \\ &= \sum_{k \in \mathcal{K}} u_k \theta\left(\frac{e^{s_k}}{J_k}\right) \\ &\quad + \sum_{k \in \mathcal{K}} \psi(\mu_k)(e^{s_k} - \hat{p}_k) \\ &\quad + \sum_{k \in \mathcal{K}} \psi(\lambda_k^J)(J_k - t_k) \\ &\quad + \sum_{k \in \mathcal{K}} \lambda_k^I (I_k(e^{\mathbf{s}}) - t_k) \\ &\quad + \omega \left( \sum_{k \in \mathcal{K}} u_k - 1 \right)\end{aligned}\tag{20}$$

$$\mathbf{z} = (\mathbf{s}, \mathbf{J}, \mathbf{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}^J, \boldsymbol{\lambda}^I, \mathbf{t}, \omega) \in \mathbb{R}^{7K+1}$$

where  $\psi(x) = x^2, x \in \mathbb{R}$ , and  $\mathbf{u} \geq 0$ .<sup>4</sup> Note that there are no constraints on the optimization variable  $\mathbf{z}$  except for  $\mathbf{u} \geq 0$ . Thus, if  $\mathbf{u} > 0$  is fixed, then the Lagrangian is unconstrained in which case the formulation falls into the framework of generalized Lagrangian theory [34, 44]. This can be used to significantly improve the convergence rate of the power control algorithms.

<sup>4</sup>The quadratic function is just one possibility for choosing the function  $\psi$ . More general conditions can be found in [57].

## 4.1 An arbitrary fixed weight vector

Suppose for a moment that the weight vector  $\mathbf{u} > 0$  is arbitrary but fixed so that the Lagrangian (20) is not a function of  $\mathbf{u}$  (it is in fact parameterized by the weight vector). If  $\mathbf{u} > 0$  is fixed, then there is no dual variable  $\omega$ , and therefore the modified non-linear Lagrangian for the problem becomes

$$\begin{aligned}L^u(\mathbf{z}) &:= L^u(\mathbf{s}, \mathbf{J}, \boldsymbol{\mu}, \boldsymbol{\lambda}^J, \boldsymbol{\lambda}^I, \mathbf{t}) \\ &= L(\mathbf{s}, \mathbf{J}, \mathbf{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}^J, \boldsymbol{\lambda}^I, \mathbf{t}, \omega), \mathbf{z} \in \mathbb{R}^{6K}.\end{aligned}\tag{21}$$

Since  $\mathbf{u}$  is fixed, it follows from [57] that any Karush-Kuhn point (e.g. a point satisfying Karush-Kuhn-Tucker conditions) can be associated with stationary points of Lagrangian (20). However, the question which immediately arises is whether the stationary points of the Lagrangian correspond to global minima of the function  $G_e(\mathbf{u}, \mathbf{s})$  for some fixed  $\mathbf{u} > 0$ . In other words, the question is whether the stationary points of the Lagrangian (which are Karush-Kuhn points of (18)) are necessary and sufficient for the sought optimum.

As already indicated before, the answer is “no” in general because Conditions (A.7)–(A.8) are not sufficient to ensure concavity with respect to the new interference variables  $J_k, k \in \mathcal{K}$ . However, as shown in [57], the concavity property is ensured whenever the function  $\theta_e(y) := \theta(e^y)$  fulfills the following condition (in addition to (A.7) and (A.4))

$$(A.9) \quad \theta_e''(y) + \theta_e'(y) \leq 0, \theta_e''(y) \geq 0, y \in \mathbb{R}.$$

Two prominent examples of functions fulfilling (A.7) and (A.9) are  $\theta(x) = -\log(x), x > 0$ , and  $\theta(x) = 1/x, x > 0$ . As mentioned in Sect. 2, the requirement of positive semi-definiteness of the interference functions (A.4) are satisfied by affine interference functions.

Now assuming (A.4), (A.7) and (A.9), Lagrangian (20) has the desired convex-concave property<sup>5</sup> and its stationary points are saddle points given by

$$\min_{\mathbf{s}, \boldsymbol{\lambda}^J, \mathbf{t}} \max_{\mathbf{J}, \boldsymbol{\mu}, \boldsymbol{\lambda}^I} L^u(\mathbf{z}) = \max_{\mathbf{J}, \boldsymbol{\mu}, \boldsymbol{\lambda}^I} \min_{\mathbf{s}, \boldsymbol{\lambda}^J, \mathbf{t}} L^u(\mathbf{z}).$$

Moreover, it can be shown that the saddle point is unique up to component signs if (A.6) holds, that is, if the gain matrix is irreducible.

In order to find stationary points of Lagrangian (21), Reference [57] considered a conditional Newton iteration on this Lagrangian of the form:

$$\begin{cases} \begin{pmatrix} \mathbf{s}^{(n+1)} \\ \boldsymbol{\mu}^{(n+1)} \end{pmatrix} = \begin{pmatrix} \mathbf{s}^{(n)} \\ \boldsymbol{\mu}^{(n)} \end{pmatrix} - (\nabla_{(\mathbf{s}, \boldsymbol{\mu})}^2 L^u(\mathbf{z}(n)))^{-1} \nabla_{(\mathbf{s}, \boldsymbol{\mu})} L^u(\mathbf{z}(n)) \\ \nabla_{(\mathbf{J}, \boldsymbol{\lambda}^J, \boldsymbol{\lambda}^I, \mathbf{t})} L^u(\mathbf{z}(n+1)) = 0 \end{cases}\tag{22}$$

For any fixed  $\mathbf{u} > 0$ , it was shown under the assumption of (A.4), (A.6), (A.7) and (A.9), that the algorithm converges to the saddle point of Lagrangian (20) which corresponds to the global optimum of  $G_e(\mathbf{u}, \mathbf{s})$ . This in turn solves the max-min SIR balancing problem (7).

<sup>5</sup>The desired property is that  $L^u$  is a convex-concave function of  $\mathbf{s}$  and  $\mathbf{J}$  for  $\boldsymbol{\lambda}^J \geq 0$  and  $\mathbf{J} - \mathbf{t} \leq 0$ . See [57, Corollary 2]

The nice property of the iteration in (22), which is also called reduced Lagrange-Newton iteration, is that it is amenable to a distributed implementation and it exhibits much higher convergence rates than classical primal-dual iterations based on gradient methods with a constant step size. In fact, [57] proved the quadratic quotient convergence of the iteration. The iteration is called *reduced* because the Newton update is performed under reduced dimensionality (in this case in  $\mathbb{R}^{2K}$  with respect to  $(\mathbf{s}, \boldsymbol{\mu})$ ). Such a dimensionality reduction is required for an efficient distributed implementation because Newton updates involve a linear combination of all gradient components with coefficients being determined by the inverse of the Hessian matrix which in general is notoriously difficult to implement in a distributed manner. In contrast, due to the dimensionality reduction and the variable splitting in (18), the reduced Hessian matrix (in  $\mathbb{R}^{2K \times 2K}$ ) has a relatively simple standard four-block structure with block-diagonals so that the Newton update is well-behaved in terms of distributed implementation.

## 4.2 Computing the optimal weights

The reduced Lagrange-Newton iteration (22) can be performed for an arbitrary but fixed weight vector  $\mathbf{u} > 0$ . Different choices of the weight vector lead in general to different utility maximization problems, and therefore provide different points on some Pareto optimal boundary. In a special case of the max-min SIR balancing problem (4), the weight vector must be of the form given by (12). As aforementioned, this vector is not known in advance and its computation in a distributed environment seems to be notoriously difficult. Moreover, an examination of Proposition 3 reveals that the weight vector highly depends on the solution to the problem at hand. For this reason, such approaches seem to be inappropriate for algorithmic solutions.

A remedy to the problem of computing the weight vector is offered by the saddle-point characterization (14) or (16). The basic idea is to use the characterization of Proposition 4 to design an iterative saddle-point algorithm that maximizes  $G_e(\mathbf{u}, \mathbf{s})$  with respect to  $\mathbf{u} \in \Pi_K$  and *simultaneously* minimizes this function over the feasible power region  $S$ . Due to Theorem 4, the algorithm will converge to a saddle point  $(\bar{\mathbf{w}}, \bar{\mathbf{s}})$  of  $G(\mathbf{u}, \mathbf{s})$ , which is a unique point in  $\Pi_K \times S$ .

In [52], we proposed a saddle-point algorithm operating on a classical linear Lagrangian similar to that in (19). The difference is that in [52] no variable splitting was used and the power constraints were not captured by Lagrange multipliers. Instead, the updates were projected on the feasible set in every iteration. If the associated Lagrangian  $L : \mathbb{R}_+^K \times S \times \mathbb{R} \rightarrow \mathbb{R}$  is defined to be  $L(\mathbf{u}, \mathbf{s}, \lambda) = G_e(\mathbf{u}, \mathbf{s}) + \lambda(\sum_{k \in \mathcal{X}} u_k - 1)$ , the algorithm of [52] takes of the form:

$$\begin{aligned} \mathbf{u}(n+1) &= \max[\mathbf{u}(n) + \delta \nabla_{\mathbf{u}} L(\mathbf{u}(n), \mathbf{s}(n), \lambda(n)), 0] \\ \mathbf{s}(n+1) &= \min[\mathbf{s}(n) - \delta \nabla_{\mathbf{s}} L(\mathbf{u}(n), \mathbf{s}(n), \lambda(n)), \hat{\mathbf{s}}] \\ \lambda(n+1) &= \lambda(n) - \delta \nabla_{\lambda} L(\mathbf{u}(n), \mathbf{s}(n), \lambda(n)) \end{aligned} \quad (23)$$

where  $\hat{\mathbf{s}} = (\hat{s}_1, \dots, \hat{s}_K) := (\log(\hat{p}_1), \dots, \log(\hat{p}_K))$  is the vector of power constraints in  $S$ , the minimum and the maximum are taken component-wise,  $\delta > 0$  is a sufficiently small step size (small enough to ensure convergence but not too small in order to achieve a sufficient convergence speed) and  $\lambda \in \mathbb{R}^K$  is the dual variable associated with the equality

constraint  $\|\mathbf{u}\|_1 = 1$ .

The algorithm solves the problem (17) since it converges to the unique saddle point of  $G_e(\mathbf{u}, \mathbf{s})$ , but the convergence speed and convergence rate are not satisfactory, especially for applications in many wireless networks. One potential approach is to combine the saddle-point characterization (16) with the conditional Newton iteration (or reduced Lagrange-Newton iteration) presented in Sect. 4.1 to obtain a significantly faster algorithm that updates the weight vector  $\mathbf{u}$  in parallel to updates of other primal and dual variables. Due to linearity of  $G_e(\mathbf{u}, \mathbf{s})$  in  $\mathbf{u}$ , we observe that Lagrangian (20) is a convex-concave function and, although we have no formal proof at the time of writing this paper, we conjecture that its stationary points are saddle points given by

$$\min_{\mathbf{s}, \lambda^J, t, \omega} \max_{\mathbf{u}, \mathbf{J}, \mu, \lambda^I} L(\mathbf{z}) = \max_{\mathbf{u}, \mathbf{J}, \mu, \lambda^I} \min_{\mathbf{s}, \lambda^J, t, \omega} L(\mathbf{z})$$

and these stationary points are necessary and sufficient for solving (17). However, if this is even true, it is not clear to us how to incorporate an online computation of weight vector updates into the iteration specified by (22). *The main challenge is to preserve both the quadratic quotient convergence and the simple structure of the Hessian matrix so that the iteration can be efficiently performed in decentralized wireless networks.*

A straightforward idea is to iterate the weight vector with the corresponding dual variable in parallel to the reduced Lagrange-Newton iteration defined by (22) so that the new combined iteration yields

$$\begin{cases} \begin{pmatrix} \mathbf{s}(n+1) \\ \boldsymbol{\mu}(n+1) \end{pmatrix} = \begin{pmatrix} \mathbf{s}(n) \\ \boldsymbol{\mu}(n) \end{pmatrix} - (\nabla_{(\mathbf{s}, \boldsymbol{\mu})}^2 L(\mathbf{z}(n)))^{-1} \nabla_{(\mathbf{s}, \boldsymbol{\mu})} L(\mathbf{z}(n)) \\ \nabla_{(\mathbf{J}, \lambda^J, \lambda^I, t)} L(\mathbf{z}(n+1)) = 0 \end{cases} \\ \begin{cases} \forall_{k \in \mathcal{X}} u_k(n+1) = \max\{u_k(n) + \delta \omega(n) + \theta \left( \frac{e^{s_k(n+1)}}{J_k(n+1)} \right), 0\} \\ \omega(n+1) = \omega(n) - \delta (\sum_{k \in \mathcal{X}} u_k(n) - 1) \end{cases} \end{cases}$$

where  $\delta > 0$  is a sufficiently small step size and  $L$  given by (20) depends on both  $\mathbf{u}$  and  $\omega$ . It is however emphasized that the approach may be not a good idea so that the challenge stated above seems to be an open problem.

## 4.3 Remarks on distributed implementation

Due to the mutual dependence of logical links, the computation of variable updates involves in general coordination and exchange of global information between network nodes. The use of classical flooding protocols to exchange this information (as for instance proposed in [13]) may consume a lot of wireless resources. Therefore, we argue in favor of a scheme based on the use of so-called *adjoint network* to efficiently distribute some locally measurable quantities to all other transmitters. A network is said to be adjoint to a given (primal) network with the gain matrix  $\mathbf{V}$  if it has the same network topology and its gain matrix is  $\mathbf{V}^T$  [53, 54]. Now, instead of each node sending its message separately as in the case of classical flooding protocols, nodes transmit simultaneously over the adjoint network in such a way that each node can estimate its gradient component based on some local measurements. Except for coarse synchronization, no cooperation between the nodes is required. When compared with [13], we expect that the overall signaling overhead can be reduced significantly because, in ad-

dition to a low-rate feedback for each transmitter-receiver pair, only the received powers in the adjoint network must be estimated on each link separately. The problem of noisy estimates can be dealt with using the techniques of stochastic approximation. For more details, the reader is referred to [53, 54].

## 5. JOINT POWER CONTROL AND RECEIVE BEAMFORMING

In this section, we assume that each link is equipped with  $M$  antenna elements at the transmitter side and the receiver side. There is a single-data stream per each link. While the transmit beamformers are arbitrary but fixed, the receive beamformers, which act as linear receivers, are jointly optimized with transmit powers. Since channel properties from one transmitter to distinct receivers are not necessarily the same, each transmitter is in general identified by a set of different spatial covariance matrices, each of which being associated with a distinct receiver. Let  $\mathbf{G}_{k,l} \in \mathbb{C}^{M \times M}$ ,  $k \in \mathcal{K}$ , be the (instantaneous or statistical) spatial covariance matrix from transmitter  $l$  to receiver  $k$  which is fixed positive semi-definite matrix. We use

$$\mathbf{r}_k := \mathbf{r}_k(\mathbf{p}) \in U_k := \{\mathbf{x} \in \mathbb{R}^M : \mathbf{x}^H \mathbf{G}_{k,k} \mathbf{x} = 1\}$$

to represent the beamforming strategy employed by link  $k \in \mathcal{K}$  and point out that the receive beamformers in general depend on  $\mathbf{p}$ . The beamformers  $\mathbf{r}_k \in \mathbb{C}^M$  of all links are collected in the receive beamforming matrix  $\mathbf{R} = (\mathbf{r}_1, \dots, \mathbf{r}_K) \in \mathbb{C}^{M \times K}$ .

With these assumptions in mind, we consider the interference functions under the optimal receive beamformers in the sense that:

$$I_k(\mathbf{p}) = \min_{\mathbf{r}_k \in U_k} (\varphi_k(\mathbf{p}, \mathbf{r}_k)), \quad k \in \mathcal{K} \quad (24)$$

where

$$\varphi_k(\mathbf{p}, \mathbf{r}_k) = \sum_{l \in \mathcal{K}, l \neq k} p_l \mathbf{r}_k^H (\mathbf{G}_{k,l} + \sigma_k^2 \mathbf{I}) \mathbf{r}_k. \quad (25)$$

Under the assumption of (24) in the definition of SIR given by (2), the max-min SIR balancing problem (7) becomes a joint power control and receive beamforming problem. The problem formulation is not novel and the problem was for instance addressed in [9, 16] (see also references therein). In contrast to the setting in this paper, however, note that Reference [9] considered a noiseless channel, which is equivalent to assuming that  $\sigma_k^2 = 0$  for each  $k \in \mathcal{K}$ . This makes the results applicable only to interference-limited systems where the link performance is limited by interference and power constraints can be neglected. Moreover, the authors of [9] were primarily interested in the single-cell downlink channel of a cellular wireless network. As a consequence of this, the algorithm proposed there is for centralized control by a base station. They used the uplink-downlink duality [8] to identify optimal transmit beamformers in the downlink channel.

It may be easily verified that any interference function of the form (24) fulfills Conditions (A.1)–(A.3). In contrast, (A.4) is not satisfied due to the minimum operator. This led the researchers to iterative algorithms that update the power vector and the receive beamforming matrix in an alternating

way. More precisely, the idea is to keep one of the variables  $\mathbf{R}$  and  $\mathbf{p}$  fixed while optimizing with respect to the other. When optimizing with respect to  $\mathbf{p}$ , the gain matrix  $\mathbf{V}(\mathbf{R})$  depends on fixed receive beamformers  $\mathbf{R}$  and is given by

$$(\mathbf{V}(\mathbf{R}))_{k,l} = \begin{cases} \mathbf{r}_k^H \mathbf{G}_{k,l} \mathbf{r}_k & k \neq l \\ 0 & k = l \end{cases}. \quad (26)$$

Further, we write  $G(\mathbf{u}, \mathbf{p}, \mathbf{R})$ ,  $\text{SIR}(\mathbf{p}, \mathbf{R})$  and for each  $k \in \mathcal{K}$   $\mathbf{B}^{(k)}(\mathbf{R})$ ,  $\rho(\mathbf{B}^{(k)}(\mathbf{R}))$ , to emphasize the dependence of the objective function, the SIR, the extended gain matrix and its spectral radius on  $\mathbf{R}$ . In what follows, we also assume that

(A.10)  $\mathbf{V}(\mathbf{R})$  is irreducible for any choice of  $\mathbf{R}$ .

The alternating optimization principle mentioned above was also applied in [9] to find an optimal allocation of transmit powers and receive beamformers. However, as the algorithm of [9] is not amenable to distributed implementation, we limit ourselves to pointing out the following: It seems that the algorithm can be extended to include the background noise and the individual power constraints by considering the principal left and right eigenvectors of  $\mathbf{B}^{(k)}(\bar{\mathbf{R}})$ ,  $k \in \mathcal{K}_0$  instead of  $\mathbf{V}(\bar{\mathbf{R}})$  where  $\bar{\mathbf{R}}$  denotes optimal receive beamformers (in the sense of the max-min SIR balancing problem; see also [9]) and the structure of  $\mathbf{B}^{(k)}(\bar{\mathbf{R}})$  is given by (9) except that  $\mathbf{V}$  is replaced by  $\mathbf{V}(\mathbf{R})$ . Further in the definition of  $\mathcal{K}_0$  given by 10,  $\rho(\mathbf{B}^{(k)})$  should be replaced by  $\rho(\mathbf{B}^{(k)}(\bar{\mathbf{R}}))$ .

In this paper, we propose the following algorithm which appears to be better predestined for distributed wireless environments.

---

**Algorithm 1** Joint power control and receive beamforming achieving the max-min SIR-balancing solution

---

**Input:**  $n = 0$ ,  $\mathbf{r}_k(0) \in U_k$ ,  $k \in \mathcal{K}$ ,  $\mathbf{p}(0) \in \mathcal{P}$ ,  $\mathbf{u}(0) \in \Pi_K^+$   
1: **repeat**  
2:    $n = n + 1$   
3:    $\mathbf{p}(n) = \arg \min_{\mathbf{p} \in \mathcal{P}} \max_{\mathbf{u} \in \Pi_K^+} G(\mathbf{u}, \mathbf{p}, \mathbf{R}(n-1))$   
4:    $\mathbf{r}_k(n) = \arg \min_{\mathbf{r} \in U_k} \varphi_k(\mathbf{p}(n), \mathbf{r}) \quad \forall_k$   
5: **until** termination condition is satisfied

---

Note that in contrast to [16], the optimal weight vector is not computed using (12) but instead it is determined iteratively in parallel to the power vector iteration. The following proposition shows that the algorithm converges to the global optimum of the max-min SIR balancing problem (7) with (24) and (25).

**PROPOSITION 5.** *The proposed algorithm converges to the global optimum of the max-min SIR balancing problem (7) over the joint space of transmit powers and receive beamformers.*

**PROOF.** First let  $n \in \mathbb{N}$  be arbitrary but fixed and let  $\varrho(\mathbf{R}(n)) = \rho(\mathbf{B}^{(k(n))}(\mathbf{R}(n)))$ . Suppose that  $(\mathbf{w}(n), \bar{\mathbf{p}}(n)) \in \Pi_K^+ \times \mathcal{P}_+$  is the saddle point of  $G(\mathbf{u}, \mathbf{p}, \mathbf{R}(n))$  for some given  $\mathbf{R}(n)$  and that  $k(n) \in \mathcal{K}_0$  is the active power constraint in the optimum. By (A.10) and Proposition 4, the saddle point

is unique. By Proposition 2, (13), (14), and (24), Algorithm 1 yields

$$\begin{aligned}
\theta(1/\varrho(\mathbf{R}(n))) &= G(\mathbf{w}(n), \bar{\mathbf{p}}(n), \mathbf{R}(n)) \\
&= \min_{\mathbf{p} \in \mathcal{P}} \max_{k \in \mathcal{K}} \theta(\text{SIR}_k(\bar{\mathbf{p}}(n), \mathbf{R}(n))) \\
&\geq \min_{\mathbf{p} \in \mathcal{P}} \max_{k \in \mathcal{K}} \theta(\text{SIR}_k(\bar{\mathbf{p}}(n), \mathbf{R}(n+1))) \\
&\geq \min_{\mathbf{p} \in \mathcal{P}} \max_{k \in \mathcal{K}} \theta(\text{SIR}_k(\bar{\mathbf{p}}(n+1), \mathbf{R}(n+1))) \\
&= G(\mathbf{w}(n+1), \bar{\mathbf{p}}(n+1), \mathbf{R}(n+1)) \\
&= \theta(1/\varrho(\mathbf{R}(n+1)))
\end{aligned}$$

where the second inequality follows from the fact that the optimal receivers in the sense of (24) maximize each SIR. So since  $\theta$  is strictly decreasing and  $n \in \mathbb{N}$  is arbitrary, we can conclude that  $\{\varrho(\mathbf{R}(n))\}_{n \in \mathbb{N}}$  with is monotonically decreasing. Now we proceed essentially as in the proof of [9, Theorem 1]. First we observe the following: (i) the spectral radius is bounded below, (ii) it is continuous in the matrix entries which are continuous functions of  $\mathbf{R}$  and (iii)  $\mathbf{R}$  is bounded in any norm. This implies that there exist a subsequence  $\{n_l\}$  and a matrix  $\hat{\mathbf{R}}$  such that  $\lim_{n_l \rightarrow \infty} \|\mathbf{R}(n_l) - \hat{\mathbf{R}}\| = 0$  and  $\lim_{n_l \rightarrow \infty} f(\mathbf{R}(n_l)) = f(\hat{\mathbf{R}}) > f_{\min}$  where  $f_{\min}$  is the global minimum. Now suppose that the optimum is not achieved so that  $f(\hat{\mathbf{R}}) = f_{\min} + \delta$  for some  $\delta > 0$ . Thus, by [9, Lemma 5], (24), (25) with (A.10) as well as by the convergence of  $\mathbf{R}(n_l)$  to  $\hat{\mathbf{R}}$ , it follows that there are some  $\epsilon := \epsilon(\delta) > 0$  and  $l_0$  such that  $\epsilon < \varrho(\mathbf{R}(n_l)) - \varrho(\mathbf{R}(n_l+1))$  for all  $l \geq l_0$ . So since  $\varrho(\mathbf{R}(n_{l+1})) \leq \varrho(\mathbf{R}(n_l+1))$ , we have (for all  $l \geq l_0$ )

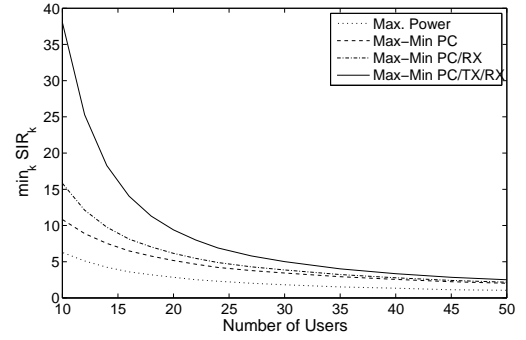
$$\epsilon < \varrho(\mathbf{R}(n_l)) - \varrho(\mathbf{R}(n_l+1)) \leq \varrho(\mathbf{R}(n_l)) - \varrho(\mathbf{R}(n_{l+1})).$$

But the right-hand side converges to zero as  $n_l$  tends to infinity which leads to contradiction.  $\square$

## 6. JOINT POWER AND TRANSCIEVER OPTIMIZATION

In the previous section, the transmit beamformers are assumed to be arbitrary but fixed. The optimization of transmit beamforming however may be crucial for the system performance, because in contrast to receive beamforming it impacts the interference power at all other receivers. Therefore a joint optimization of transmit and receive beamformers together with transmit powers offers huge potential for performance gains.

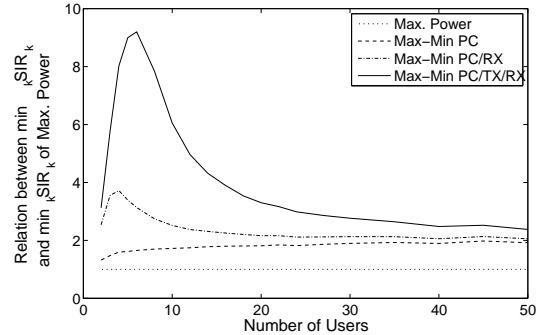
Again the problem of transmit beamforming optimization (separate or joint) has been addressed in many papers, including [43, 10, 47, 16]. In particular, it is shown in [43] that the transceiver optimization problem in networks with general power constraints presents some additional challenges which is mainly due to the lack of the uplink-downlink duality. Nevertheless some joint optimization of power-controlled transmit and receive beamformers is possible. As in [10] the basic idea is to optimize the transmit and receive beamformers in an alternating way. For a better understanding we use the notion of the primal and the reversed network. In contrast to the primal network the reversed network is the network that is obtained by reversing the roles of transmitters and receivers. More precisely, transmitters and transmit beamformers are assumed to be receivers and receive beamformers, respectively. Vice versa receivers and receive



**Figure 1: Balanced SIR as a function of the number of users for different schemes: No power control (users transmit at their maximum transmit powers) with beamformers matched to the local channels (Max.Power), power control with beamformers matched to the local channels (Max-Min PC), power control and receive beamforming (Max-Min PC/RX), transmit and receive beamforming (Max-Min PC/TX/RX).**

beamformers are assumed to be transmitters and transmit beamformers.

Figures 1 and 2 illustrate the performance gain that can be achieved by different resource allocation schemes. We consider a network with  $K$  links,  $M = 4$  transmit and receive antennas and a channel matrix which entries are iid complex Gaussian distributed. The max-min SIR-balanced value averaged over 1000 channel realizations is depicted over the number of users/links in the network for the 4 resource allocation strategies discussed in this paper. All links have the same individual power constraints and operate at SNR = 30dB.



**Figure 2: Relative improvement over the maximum power allocation (Max.Power).**

## 7. REFERENCES

- [1] J. M. Aein. Power balancing in systems employing frequency reuse. *COMSAT Tech. Rev.*, 3(2):277–300, 1973.
- [2] H. Alavi and R. Nettleton. Downstream power control for a spread spectrum cellular mobile radio system. In *Proc. IEEE Global Communications Conference*



- (*GLOBECOM*), pages 84–88, 1982.
- [3] T. Alpcan, T. Basar, R. Srikant, and E. Altman. Cdma uplink power control as a noncooperative game. *Wireless Networks*, 8(6):659–670, Nov. 2002.
  - [4] N. Bambos. Toward power-sensitive network architectures in wireless communications: Concepts, issues, and design aspects. *IEEE Personal Commun. Mag.*, 5:50–59, June 1998.
  - [5] N. Bambos, S. Chen, and G. Pottie. Channel access algorithms with active link protection for wireless communication networks with power control. *IEEE/ACM Trans. Networking*, 8(5):583–597, October 2000.
  - [6] R. Berry, P. Liu, and M. Honig. Design and analysis of downlink utility-based schedulers. In *Proc. 43rd Annual Allerton Conference on Communication, Control and Computing*, Oct. 2002.
  - [7] D. Bertsekas and R. Gallager. *Data Networks*. Prentice-Hall, Englewood Cliffs, 1992.
  - [8] H. Boche and M. Schubert. A general duality theory for uplink and downlink beamforming. In *VTC fall, Vancouver, Canada*, September 2002.
  - [9] H. Boche and M. Schubert. Resource allocation in multiantenna systems - Achieving max-min fairness by optimizing a sum of inverse SIRs. *IEEE Trans. Signal Processing*, 54(6), June 2006.
  - [10] J. Chang, L. Tassiulas, and F. Rashid-Farrokhi. Joint transmitter receiver diversity for efficient space division multiaccess. *IEEE J. Select. Areas Commun.*, 16(8):1437–1449, Oct. 1998.
  - [11] L. Chen, S. Low, M. Chiang, and J. Doyle. Optimal cross-layer congestion control, routing and scheduling design in ad hoc wireless networks. In *Proc. 25th IEEE Conference on Computer Communications (INFOCOM), Barcelona*, April 23-29 2006.
  - [12] M. Chiang. To layer or not to layer: Balancing transport and physical layers in wireless multihop networks. In *Proc. 23rd IEEE Conference on Computer Communications (INFOCOM), Hong Kong*, March 2004.
  - [13] M. Chiang. Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control. *IEEE J. Select. Areas Commun.*, 23(1):104–116, Jan. 2005.
  - [14] R. Cruz and A. Santhanam. Optimal routing, link scheduling and power control in multi-hop wireless networks. In *Proc. 22nd IEEE Conference on Computer Communications (INFOCOM), San Francisco, CA, USA*, March 30-April 3 2003.
  - [15] T. ElBatt and A. Ephremides. Joint scheduling and power control for wireless ad hoc networks. *IEEE Trans. Wireless Commun.*, 3(1):74–85, January 2004.
  - [16] A. Feistel, S. Stanczak, and M. Kaliszan. Resource allocation in multiantenna systems under general power constraints. In *Proc. of International ITG Workshop on Smart Antennas 2009 (WSA)*, Berlin, Germany, Feb. 16–19 2009.
  - [17] A. Feiten and R. Mathar. Optimal power control for multiuser cdma channels. In *Proc. 2005 IEEE International Symposium on Information Theory (ISIT), Adelaide, Australia*, Sept. 4-9 2005.
  - [18] N. Feng, S.-C. Mau, and N. Mandayam. Pricing and power control for joint network-centric and user-centric radio resource management. *IEEE Trans. Commun.*, 52(9):1547–1557, 2004.
  - [19] G. Foschini and Z. Miljanic. A simple distributed autonomous power control algorithm and its convergence. *IEEE Trans. Veh. Technol.*, 42(4):641–646, Nov. 1993.
  - [20] S. Friedland and S. Karlin. Some inequalities for the spectral radius of non-negative matrices and applications. *Duke Math. J.*, 42(3):459–490, 1975.
  - [21] L. Georgiadis, M. Neely, and L. Tassiulas. *Resource Allocation and Cross-Layer Control in Wireless Networks*. Now Publishers Inc., 2006.
  - [22] D. Gerlach and A. Paulraj. Base station transmitting antenna arrays for multipath environments. *Signal Processing (Elsevier Science)*, 54:59–73, 1996.
  - [23] D. Goodman and N. Mandayam. Power control for wireless data. *IEEE Personal Commun. Mag.*, 7:48–54, April 2000.
  - [24] P. Hande, S. Rangan, M. Chiang, and X. Wu. Distributed uplink power control for optimal sir assignment in cellular data networks. *IEEE/ACM Trans. Networking*, 16(6):1420–1433, Dec. 2008.
  - [25] S. Hanly. An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity. *IEEE Journal on Selected Areas in Communications*, 13(7):1332–1340, 1995.
  - [26] B. He, M. Wang, and E. Li. A new distributed power balancing algorithm for CDMA cellular systems. In *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, volume 3, pages 1768 – 1771, 1997.
  - [27] W. Hongyu, H. Aiging, H. Rong, and G. Weikang. Balanced distributed power control. In *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, volume 2, pages 1415 – 1419, 2000.
  - [28] J. Huang, R. Berry, and M. L. Honig. Distributed interference compensation for multi-channel wireless networks. In *Proc. 43rd Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, USA, Sept. 2005.
  - [29] J. Huang, R. Berry, and M. L. Honig. A game theoretic analysis of distributed power control for spread spectrum ad hoc networks. In *Proc. IEEE International Symposium on Information Theory (ISIT), Adelaide, Australia*, Sept. 4-9 2005.
  - [30] J. Huang, R. Berry, and M. L. Honig. Distributed interference compensation for wireless networks. *IEEE J. Select. Areas Commun.*, 24(5):1074–1084, May 2006.
  - [31] J. Huang, V. G. Subramanian, R. Agrawal, and R. Berry. Joint scheduling and resource allocation in uplink ofdm systems for broadband wireless access networks. *IEEE J. Select. Areas Commun.*, 27(2):288–296, Feb. 2009.
  - [32] M. Johansson, L. Xiao, and S. Boyd. Simultaneous routing and resource allocation in cdma wireless data networks. In *Proc. IEEE International Conference on Communications, Anchorage, Alaska*, May 2003.
  - [33] F. Kelly, A. Maulloo, and D. Tan. Rate control for communication networks: Shadow prices, proportional

- fairness and stability. *J. Oper. Res. Soc.*, 49(3):237–252, March 1998.
- [34] O. L. Mangasarian. Unconstrained Lagrangians in Nonlinear Programming. *SIAM Journal on Control and Optimization*, 13(4):772–791, May 1975.
- [35] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, Philadelphia, 2000.
- [36] H. J. Meyerhoff. Method for computing the optimum power balance in multibeam satellites. *COMSAT Tech. Rev.*, 4(1):139–146, 1974.
- [37] H. Minc. *Nonnegative Matrices*. John Wiley & Sons, Inc., 1988.
- [38] J. Mo and J. Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Trans. on Networking*, 8(5):556–567, October 2000.
- [39] G. Montalbano, I. Ghauri, and D. T. M. Slock. Spatio-temporal array processing for CDMA/SDMA downlink transmission. In *Proc. Asilomar Conf. on Signals, Systems and Computers, Monterey, CA, USA*, pages 1337–1341, 1998.
- [40] D. O’Neill, D. Julian, and S. Boyd. Seeking Foschini’s genie: Optimal rates and powers in wireless networks. *IEEE Trans. Veh. Technol.*, 2003 (accepted). Available from <http://www.stanford.edu/boyd/>.
- [41] D. P. Palomar and M. Chiang. A tutorial on decomposition methods for network utility maximization. *IEEE J. Select. Areas Commun.*, 24(8):1439–1451, Aug. 2006.
- [42] J. Price and T. Javidi. Decentralized rate assignments in a multi-sector cdma network. *IEEE Trans. Wireless Commun.*, 5(12):3537–3547, Dec. 2006.
- [43] F. Rashid-Farrokhi, K. J. Liu, and L. Tassiulas. Transmit beamforming and power control for cellular wireless systems. *IEEE J. Select. Areas Commun.*, 16(8):1437–1449, Oct. 1998.
- [44] R. T. Rockafellar. Augmented Lagrange multiplier functions and duality in nonconvex programming. *SIAM Journal on Control and Optimization*, 12:268–285, 1974.
- [45] C. Saraydar, N. Mandayam, and D. Goodman. Pricing and power control in a multicell wireless data network. *IEEE J. Select. Areas Commun.*, 19(10):1883–1892, 2001.
- [46] C. Saraydar, N. Mandayam, and D. Goodman. Efficient power control via pricing in wireless data networks. *IEEE Trans. Commun.*, 50(2):291–303, 2002.
- [47] M. Schubert and H. Boche. QoS-based resource allocation and transceiver optimization. *Foundation and Trends in Communications and Information Theory*, 2(6), 2006.
- [48] E. Seneta. *Non-Negative Matrices and Markov Chains*. Springer, Berlin, 1981.
- [49] S. Stanczak and H. Boche. Information theoretic approach to the Perron root of nonnegative irreducible matrices. In *Proc. 2004 Information Theory Workshop (ITW)*, San Antonio, Texas, USA, Oct. 24–29 2004.
- [50] S. Stanczak, M. Kaliszan, and N. Bambos. A characterization of max-min SIR-balanced power allocation with applications. *Wireless Networks (Springer)*, 2010. To appear. Available at <http://arXiv.org/abs/0901.0824>.
- [51] S. Stanczak, M. Kaliszan, N. Bambos, and M. Wiczanowski. A characterization of max-min SIR-balanced power allocation with applications. In *Proc. IEEE International Symposium on Information Theory (ISIT)*, Seoul, Korea, June 28–July 3 2009.
- [52] S. Stanczak, M. Kaliszan, and M. Goldenbaum. Max-Min Fair Rate Control Based on a Saddle-Point Characterization of Some Perron Roots. In *Accepted for publication in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP ’11)*, Prague, Czech Republic, May 2011.
- [53] S. Stanczak, M. Wiczanowski, and H. Boche. Distributed utility-based power control: Objectives and algorithms. *IEEE Trans. Signal Processing*, 55(10):5058–5068, Oct. 2007.
- [54] S. Stanczak, M. Wiczanowski, and H. Boche. *Fundamentals of Resource Allocation in Wireless Networks*, volume 3 of *Foundations in Signal Processing, Communications and Networking*. Springer, Berlin, 2009.
- [55] A. Subramanian and A. H. Sayed. Joint rate and power control algorithms for wireless networks. *IEEE Trans. Signal Processing*, 53(11):4204–4214, Nov. 2005.
- [56] A. Tang, J. Wang, and S. Low. Is fair allocation always inefficient? In *Proc. 23rd IEEE Conference on Computer Communications (INFOCOM)*, Hong Kong, March 7–11 2004.
- [57] M. Wiczanowski, S. Stanczak, and H. Boche. Providing quadratic convergence of decentralized power control in wireless networks – The method of min-max functions. *IEEE Trans. Signal Processing*, 56(8):4053–4068, Aug. 2008.
- [58] Q. Wu. Optimum transmitter power control in cellular systems with heterogeneous SIR thresholds. *IEEE Trans. Veh. Technol.*, 49(4):1424–1429, July 2000.
- [59] M. Xiao, N. Shroff, and E. Chong. A utility-based power control scheme in wireless cellular systems. *IEEE/ACM Trans. Networking*, 11(2):210–221, April 2003.
- [60] W. Yang and G. Xu. Optimal downlink power assignment for smart antenna systems. In *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Seattle, WA, USA, May 1998.
- [61] R. Yates. A framework for uplink power control in cellular radio systems. *IEEE J. Select. Areas Commun.*, 13(7):1341–1347, September 1995.
- [62] R. Yates and C. Huang. Integrated power control and base station assignment. *IEEE Trans. Veh. Technol.*, 44(3):638–644, August 1995.
- [63] J. Zander. Distributed cochannel interference control in cellular radio systems. *IEEE Trans. Veh. Technol.*, 41:305–311, August 1992.
- [64] J. Zander. Performance of optimum transmitter power control in cellular radio systems. *IEEE Trans. Veh. Technol.*, 41(1):57–62, February 1992.
- [65] J. Zander and S.-L. Kim. *Radio Resource Management for Wireless Networks*. Artech House, Boston, London, 2001.