QUADRATICALLY CONVERGING DECENTRALIZED POWER ALLOCATION ALGORITHM FOR WIRELESS AD-HOC NETWORKS - THE MAX-MIN FRAMEWORK

Marcin Wiczanowski, Slawomir Stanczak†, Holger Boche*†,*

* Heinrich-Hertz Chair, EECS,
University of Technology Berlin,
Einsteinufer 25, 10587 Berlin, Germany
Tel/Fax: +49(0)30-314-28462/-28320
Email: Marcin.Wiczanowski@TU-Berlin.de

† Fraunhofer German-Sino Lab
for Mobile Communications
Einsteinufer 37, 10587 Berlin, Germany
Email: {stanczak,boche}@hhi.fhg.de

ABSTRACT

This work addresses the problem of designing efficient resource allocation algorithms for wireless ad-hoc networks with best-effort traffic. Relying on the framework of generalized Lagrangeans and duality we propose an optimization concept that combines the decentralization with quadratic quotient convergence and unconstrained iteration character.

1. INTRODUCTION

Contemporary wireless ad-hoc networks (802.11,.15,.16) are known to carry a hybrid mixture of traffic flows with different priorities and requirements with respect to QoS (Quality of Service) parameters, like delay or data-rate. The policy of QoS provision for all flows in such multi-hop networks consists of an end-to-end control mechanism (transport layer issue) and node-by-node, or per-link, control policy (MAC layer). The end-to-end policy includes the routing policy and the end-to-end congestion control mechanism consisting in the dynamical adjustment of source rates according to the states of the routes - e.g. by the transmission window control [1]. The node-by-node policy consists of the link scheduling policy, nominating the links for concurrent transmission, and the resource allocation algorithm, allocating transmit power and bandwidth to transmitting nodes according to some objective. It is a known fact that a node-by-node control is an important element contributing to an improved dynamic behavior of the overall QoS provision policy in terms of latency and stability.

In this work we focus on the power allocation algorithm as a part of the node-by-node policy in networks carrying so-called best-effort traffic [1], [2], [3]. We can classify a power allocation algorithm as efficient, if it satisfies three key requirements; fast quotient convergence, decentralized realizability and unconstrained iteration character (in the sense to be described). Due to the specific splitting of optimization variables and the use of interesting local properties of max-min points and saddle points we succeed in constructing a primal-dual algorithm satisfying all the above requirements.

2. SYSTEM MODEL AND PRELIMINARIES

We consider one hop of a multi-hop communication in an ad-hoc network, under some given routing and link scheduling policy. We denote by $\mathcal{K} := \{1, \dots, K\}$ the set of concurrently activated peer-to-peer links. We assume the network to carry best-effort traffic only, which means that there are no stringent QoS values (maximum delay, minimum data-rate) to be ensured for each flow. Instead, the appropriate aim of power allocation is the optimization of some weighted sum of per-link QoS parameters [1], [2], [3], [4]. In wireless networks the value decisive for the link-QoS is the corresponding signal-to-interference ratio (SIR), defined as $\mathrm{SIR}_k(\mathbf{p}) := p_k/J_k(\mathbf{p})$, $k \in \mathcal{K}$ with p_k as the transmit power of k-th transmitter (source of link k), $\mathbf{p} = (p_1, \dots, p_K)$, and $J_k : \mathbb{R}_+^K \to S \subseteq \mathbb{R}_+$ as the interference function describing the interference power at k-th receiver (destination of link k) as a function of $\mathbf{p}$. Hence, we can write the power allocation problem as

$$\min_{\mathbf{p}} \sum_{k \in \mathcal{K}} \alpha_k \phi\left(\frac{p_k}{J_k(\mathbf{p})}\right), \text{ subject to } \begin{cases} -\mathbf{p} \leq 0 \\ \mathbf{p} - \hat{\mathbf{p}} \leq 0, \end{cases} \quad (1)$$

with $\phi : \mathbb{R}_+ \to Q \subseteq \mathbb{R}$ as some decreasing QoS function generating the SIR-dependent values of QoS parameters, $\hat{\mathbf{p}}$ as the vector of per-link transmit power constraints and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_K) > 0$ as the vector of priority factors determined by the traffic types on the links. For instance, we may take $\phi(\mathrm{SIR}) = -\log(1 + \mathrm{SIR})$ when the link data-rate is the QoS parameter of interest, or $\phi(\mathrm{SIR}) = 1/\mathrm{SIR}$, when the interest is in the link reliability expressed by the average bit-error-rate slope. In the remainder, we shall assume a fair statement of the power allocation problem, i.e. the choice of $(\boldsymbol{\alpha}, \phi)$ such that at the optimum of (1) each link $k \in \mathcal{K}$ is allocated nonzero transmit power. Such slight restriction allows us to use an almost always advantageous (see e.g. [4]) bijective variable transformation $\mathbf{x} := \log \mathbf{p}$ ($\mathbf{p} = \exp(\mathbf{x})$). Defining $J_k^e(\mathbf{x}) := J_k(\exp(\mathbf{x}))$, the following condition will be sometimes of interest in the remainder.

1-4244-0469-X/06/$20.00 ©2006 IEEE

Condition C1 *For J^e_k, $k \in \mathcal{K}$ there holds $[\nabla^2 J^e_k(\mathbf{x})]_{jl} = 0$, $j,l \in \mathcal{K}$, $j \neq l$.*
The condition characterizes the class of interference functions which can be written as $J^e_k(\mathbf{x}) = \sum_{l\in\mathcal{K}} j^e_l(x_l) + const$, $j^e_l : \mathbb{R} \to S_l \subseteq \mathbb{R}_+$, $k,l \in \mathcal{K}$. Such class includes the most common case of a linear receiver network, corresponding to

$$J^e_k(\mathbf{x}) = \sum_{l\in\mathcal{K}} V_{kl} e^{x_l} + n_k, \quad k \in \mathcal{K}, \tag{2}$$

with cross-talk factors $V_{kl} \geq 0$ and background noise variances $n_k > 0$ [2], [4].

We can agree on three main criteria, with respect to convergence, realizability and complexity, which identify an online power allocation algorithm solving (1) as efficient.
1) The algorithm has to exhibit sufficiently fast quotient convergence. This allows the power (re-)allocation process to follow the high dynamics of network topology, traffic character and large-scale channel fading with acceptable latency. The quotient convergence is measured by the norm-dependent convergence factor and norm-independent convergence order.
Definition *Let $S_{\mathcal{I}}(\tilde{\mathbf{x}})$ be the set of all iterate sequences generated by some iteration $\mathcal{I}$ and convergent to $\tilde{\mathbf{x}}$.*
i.) The p-th quotient convergence factor ($p \geq 1$) of $\mathcal{I}$ at $\tilde{\mathbf{x}}$ is defined as[1] *$Q_p(\mathcal{I}, \tilde{\mathbf{x}}) = \sup_{\{\mathbf{x}(n)\}\in S_{\mathcal{I}}(\tilde{\mathbf{x}})} q_p(\{\mathbf{x}(n)\}, \tilde{\mathbf{x}})$, with $q_p(\{\mathbf{x}(n)\}, \tilde{\mathbf{x}}) = \limsup_{n\to\infty} \|\mathbf{x}(n+1) - \tilde{\mathbf{x}}\| / \|\mathbf{x}(n) - \tilde{\mathbf{x}}\|^p$.*
ii.) The quotient convergence order of $\mathcal{I}$ at $\tilde{\mathbf{x}}$ is defined as[2] *$O_Q(\mathcal{I}, \tilde{\mathbf{x}}) = \inf_{p\geq 1: Q_p(\mathcal{I},\tilde{\mathbf{x}})=\infty} p$.*
For instance, regardless of the iteration parameters for gradient-based methods we always have $O_Q(\mathcal{I}, \tilde{\mathbf{x}}) \leq 1$ (frequently $O_Q(\mathcal{I}, \tilde{\mathbf{x}}) = 1$, i.e. linear convergence), while for the superior Newton method it may hold $O_Q(\mathcal{I}, \tilde{\mathbf{x}}) = 2$ (quadratic convergence). In some real-world network cases linear convergence may prove to be insufficient, while the quadratic one is regarded as fastest achievable in practical applications.
2) The algorithm has to be realizable in a decentralized manner. An implementation requiring local actions at certain nodes based on their local knowledge conforms to the ad-hoc network nature and makes the existence of some centralized controller superfluous. Hereby, we regard the provision of necessary local knowledge by means of peer-to-peer feedback at each link as maintainable (see the concept in [5]).
3) The iteration has to be of unconstrained nature. By unconstrained nature we mean here that, while the obtained minimizer of (1) is obviously feasible (means: satisfies constraints in (1)), no attention needs to be payed for feasibility of consecutive iterates. This brings complexity advantages and avoids the deterioration of the convergence factor as the consequence of e.g. projection of infeasible iterates on the feasible set. Moreover, in some cases such projection on the feasible set requires global network knowledge, which could make requirement 2) hardly satisfiable as well.

[1] Defined in this way only if $\mathbf{x}(n) = \tilde{\mathbf{x}}$ for finitely many $n \in \mathbb{N}$.
[2] Notice that $Q_p = 0$ for $p \in [1, p_0)$, $Q_p = c < \infty$ for $p = p_0$ and $Q_p = \infty$ for $p \in (p_0, \infty)$.

3. APPROACH WITH MODIFIED LAGRANGEAN

Our experience shows that any traditional approach (e.g. Newton method + barrier functions) applied directly to the form (1) fails at satisfying the requirements 1-3) concurrently. This pushes us to a specific reformulation of (1) of the form

$$\min_{\mathbf{x},\mathbf{I}} \sum_{k\in\mathcal{K}} \alpha_k \phi(\frac{e^{x_k}}{I_k}), \text{ s.t. } \begin{cases} e^{x_k} - \hat{p}_k \leq 0 \\ J^e_k(\mathbf{x}) - I_k = 0 \end{cases}, k \in \mathcal{K}. \tag{3}$$

For this problem we utilize a modified Lagrange function

$$L^\psi(\mathbf{z}) = \sum_{k\in\mathcal{K}} \alpha_k \phi(\frac{e^{x_k}}{I_k}) + \sum_{k\in\mathcal{K}} \psi(\mu_k)(e^{x_k} - \hat{p}_k) + \sum_{k\in\mathcal{K}} \lambda_k (J^e_k(\mathbf{x}) - I_k), \tag{4}$$

with $\mathbf{z} := (\mathbf{x}, \mathbf{I}, \boldsymbol{\mu}, \boldsymbol{\lambda}) \in \mathbb{R}^{4K}$ and function ψ satisfying
Condition C2 *i.) $\psi : \mathbb{R} \to \mathbb{R}_+$, ii.) $\psi(-\mu) = \psi(\mu)$, $\mu \in \mathbb{R}$, iii.) $\psi(\mu) = 0$ iff $\mu = 0$, iv.) $\psi'(\mu) = 0$ iff $\mu = 0$, v.) $\psi''(\mu) > 0$, $\mu \in \mathbb{R}$.*
A simple example of a function satisfying Condition C2 is $\psi(\mu) = \mu^2$. The proof of the following statement is a straightforward consequence of Condition C2.
Lemma 1 *A vector $(\mathbf{x}, \mathbf{I}, \boldsymbol{\nu}, \boldsymbol{\lambda})$ satisfies the Karush-Kuhn-Tucker (KKT) conditions for (3) iff a vector $(\mathbf{x}, \mathbf{I}, \boldsymbol{\mu}, \boldsymbol{\lambda})$, such that $(\mathbf{x}, \mathbf{I})$ is feasible (satisfies constraints in (3)) and $\psi(\pm\mu_k) = \nu_k$, $k \in \mathcal{K}$, is a stationary point of L^ψ.*
In the light of the above Lemma and the fact that dom$L^\psi = \mathbb{R}^{4K}$, the primal-dual search of stationary points of L^ψ is a suitable concept for solving (3) by an unconstrained iteration (requirement 3)). From the classical framework of primal-dual optimization, which can be easily seen to hold for the modified Lagrangean L^ψ in unchanged form [8], it follows further that our interest is in special stationary points

$$\tilde{\mathbf{z}} = \arg \max_{(\boldsymbol{\mu},\boldsymbol{\lambda})\in\mathbb{R}^{2K}} \min_{(\mathbf{x},\mathbf{I})\in S(\tilde{\mathbf{x}},\tilde{\mathbf{I}})} L^\psi(\mathbf{z}), \tag{5}$$

(max-min points) representing the saddle points of L^ψ, i.e. satisfying additionally

$$\tilde{\mathbf{z}} = \arg \min_{(\mathbf{x},\mathbf{I})\in S(\tilde{\mathbf{x}},\tilde{\mathbf{I}})} \sup_{(\boldsymbol{\mu},\boldsymbol{\lambda})\in\mathbb{R}^{2K}} L^\psi(\mathbf{z}), \tag{6}$$

with $S(\mathbf{y})$ as some neighborhood of $\mathbf{y}$. This is because stationary points (6) correspond to local minimizers of (3) [8].

For finding the desired point $\tilde{\mathbf{z}}$, we construct the iteration referred later to as $\mathcal{I}$ and taking the form

$$\begin{cases} \begin{bmatrix} \mathbf{x}(n+1) \\ \boldsymbol{\mu}(n+1) \end{bmatrix} = \begin{bmatrix} \mathbf{x}(n) \\ \boldsymbol{\mu}(n) \end{bmatrix} - \mathbf{H}^{-1}_{\mathbf{x},\boldsymbol{\mu}} L^\psi(\mathbf{z}(n)) \begin{bmatrix} \nabla_{\mathbf{x}} L^\psi(\mathbf{z}(n)) \\ \nabla_{\boldsymbol{\mu}} L^\psi(\mathbf{z}(n)) \end{bmatrix}, \\ (\mathbf{I}(n+1), \boldsymbol{\lambda}(n+1)) \text{ solving } \begin{bmatrix} \nabla_{\mathbf{I}} L^\psi(\mathbf{z}(n+1)) \\ \nabla_{\boldsymbol{\lambda}} L^\psi(\mathbf{z}(n+1)) \end{bmatrix} = 0, \quad n \in \mathbb{N}, \end{cases}$$

with[3]

$$\mathbf{H}_{\mathbf{x},\boldsymbol{\mu}} L^\psi(\mathbf{z}) := \begin{bmatrix} \nabla^2_{\mathbf{x}} L^\psi(\mathbf{z}) & \nabla^2_{\mathbf{x},\boldsymbol{\mu}} L^\psi(\mathbf{z}) \\ \nabla^{2T}_{\mathbf{x},\boldsymbol{\mu}} L^\psi(\mathbf{z}) & \nabla^2_{\boldsymbol{\mu}} L^\psi(\mathbf{z}) \end{bmatrix}. \tag{7}$$

[3] We use the formalism $[\nabla^2_{\mathbf{a},\mathbf{b}}(\cdot)]_{kj} := \delta^2(\cdot)/\delta a_k \delta b_j$, $k,j \in \mathcal{K}$, $\nabla^2_{\mathbf{a}}(\cdot) \equiv \nabla^2_{\mathbf{a},\mathbf{a}}(\cdot)$.

Iteration $\mathcal{I}$ can be classified as a conditional Newton-based search or a Newton-based search under reduced dimensionality. For interference functions satisfying Condition C1, the blocks of the Hessian in (7) have the crucial property of being diagonal. Hence, $\mathbf{H}^{-1}_{\mathbf{x},\boldsymbol{\mu}}L^{\psi}(\mathbf{z})$ consists of blocks (corresponding to those in (7)), which are all diagonal and their k-th diagonal elements are explicitly expressible as functions of k-th diagonal elements of the corresponding blocks in (7). Consequently, the inverse of the Hessian in $\mathcal{I}$ is obtainable by simple elementwise operations. The other crucial property of iteration $\mathcal{I}$ is the explicit solvability of the equation system in $\mathcal{I}$ for $(\mathbf{I}(n+1), \boldsymbol{\lambda}(n+1))$, since with (4) this system can be written as

$$\left\{ \begin{array}{l} I_k(n+1) = J^e_k(\mathbf{x}(n+1)) \\ \lambda_k(n+1) = -\alpha_k \phi'(\frac{e^{x_k(n+1)}}{I_k(n+1)})\frac{e^{x_k(n+1)}}{I^2_k(n+1)} \end{array} \right. , \quad k \in \mathcal{K}. \quad (8)$$

The above properties play a decisive role in the combination of quadratic convergence with distributed realizability of $\mathcal{I}$.

4. LOCAL CONVERGENCE AND DUALITY

Proposition 1 *Assume $\tilde{\mathbf{z}}$ some stationary point of L^{ψ}, such that $\mathbf{H}_{\mathbf{x},\boldsymbol{\mu}}L^{\psi}(\tilde{\mathbf{z}})$ is nonsingular and each of the mappings ψ'', $\delta^2 J^e_k/\delta x^2_l$ and $\phi''_{x,I}(x_k, I_k) := \phi''(\frac{e^{x_k}}{I_k})$, $k \in \mathcal{K}$ is Lipschitz-continuous on some neighborhood of the orthogonal projection of $\tilde{\mathbf{z}}$ on its domain. Then, $\tilde{\mathbf{z}}$ is a point of attraction[4] of $\mathcal{I}$ and $O_Q(\mathcal{I}, \tilde{\mathbf{z}}) = 2$.*
Remark on the proof: The attraction property and $O_Q(\mathcal{I}, \tilde{\mathbf{z}}) \geq 1$ is proven by showing that the conditional iteration $\mathcal{I}$ is equivalent to the unconditional discrete Newton method (on restricted dimension), with the direction matrix being a strongly consistent approximation of $H_{\mathbf{x},\boldsymbol{\mu}}L^{\psi}(\mathbf{z})$ at $\mathbf{z} = \tilde{\mathbf{z}}$ [9]. $O_Q(\mathcal{I}, \tilde{\mathbf{z}}) = 2$ follows then from the Lipschitz conditions.
As already mentioned, additionally to the convergence, we need the point of attraction to be a saddle point satisfying (5-6). Consider therefore the following simple Lemma.
Lemma 2 *A stationary point $\tilde{\mathbf{z}}$ of L^{ψ} is a max-min point (5), such that $(\tilde{\mathbf{x}}, \tilde{\mathbf{I}})$ is feasible, iff it is a saddle point (5-6), iff L^{ψ} is locally strictly convex as a function of $(\mathbf{x}, \mathbf{I})$ in some neighborhood of $\tilde{\mathbf{z}}$, with $(\tilde{\mathbf{x}}, \tilde{\mathbf{I}})$ feasible.*
Remark on the proof: The statement follows from the analysis of the second order condition for the max-min point (5), which is given e.g. in [7].
Lemma 2 characterizes the case of usability of primal-dual optimization methods in general, with $\mathcal{I}$ as only a special case [8]. In fact, the characteristics of (ϕ, J^e_k), $k \in \mathcal{K}$ may prevent L^{ψ} from being locally strictly convex as a function of $(\mathbf{x}, \mathbf{I})$ at some stationary point corresponding to a local minimizer of (3), or even at any feasible stationary point. In such cases of local/ global lack of strong duality, primal-dual methods (and hence also $\mathcal{I}$) are unable to find the desired minimizers.

5. UNIQUENESS OF THE SADDLE POINT

We are interested in the case of a saddle point $\tilde{\mathbf{z}}$ in the sense of (5-6), which represents a unique stationary point of L^{ψ} for all feasible $(\mathbf{x}, \mathbf{I})$. In such case the unique saddle point corresponds to the global minimizer of (3). Recall that by the remark on the proof to Proposition 1, $\mathcal{I}$ is equivalent to a discrete Newton method. Hence, by the (extended) Newton-Kantorovich Theorem [9] one can show that $\mathcal{I}$ converges to $\tilde{\mathbf{z}}$ being a saddle point (5-6) and a unique feasible stationary point of L^{ψ} if i.) strict complementarity holds at $\tilde{\mathbf{z}}$ (see [8]), ii.) $\mathbf{z}(0) \in S$, with some $S \ni \tilde{\mathbf{z}}$, iii.) $\|\mathbf{H}^{-1}_{\mathbf{x},\boldsymbol{\mu}}L^{\psi}(\mathbf{z}(0))\| \leq \beta(S)$, with some S-dependent bound $\beta(S)$, iv.) each of the mappings ψ'', $\delta^2 J^e_k/\delta x^2_l$ and $\phi''_{x,I}(x_k, I_k)$, $k \in \mathcal{K}$ is Lipschitz-continuous on the orthogonal projection of S on its domain. Such kind of convergence behavior is referred to as semi-local[5] (obviously, the property $O_Q(\mathcal{I}, \tilde{\mathbf{z}}) = 2$ still holds).

The problem formulation (3) turns out to be unfavorable in terms of the property of a saddle point (5-6) as a unique feasible stationary point. In particular, a class of functions (ϕ, J^e_k), $k \in \mathcal{K}$ leading to a convex power allocation problem (3), and hence enforcing such property, is small and not useful for QoS considerations. In order to characterize a wider class of functions yielding the desired property, we rewrite problem (3) as a min-max problem over $(\mathbf{x}, \mathbf{I})$, i.e.

$$\min_{\mathbf{x}} \max_{\mathbf{I}} \sum_{k \in \mathcal{K}} \alpha_k \phi(\frac{e^{x_k}}{I_k}), \text{ s. t.} \left\{ \begin{array}{l} e^{x_k} - \hat{p}_k \leq 0 \\ I_k - J^e_k(\mathbf{x}) \leq 0 \end{array} \right. , k \in \mathcal{K}. \quad (9)$$

By separating the constraints for minimization and maximization variables, e.g. according to $I_k - t \leq 0$, $t - J^e_k(\mathbf{x}) = 0$ with some telescope variable $t \in \mathbb{R}$, the modified Lagrangean for (9) takes the form[6]

$$\begin{aligned} L^{\psi}(\mathbf{z}, t) = &\sum_{k \in \mathcal{K}} \alpha_k \phi(\frac{e^{x_k}}{I_k}) + \sum_{k \in \mathcal{K}} \psi(\mu_k)(e^{x_k} - \hat{p}_k) \\ &+ \sum_{k \in \mathcal{K}} \lambda^J_k (t - J^e_k(\mathbf{x})) - \sum_{k \in \mathcal{K}} \psi(\lambda^I_k)(I_k - t), \end{aligned} \quad (10)$$

with $\mathbf{z} := (\mathbf{x}, \mathbf{I}, \boldsymbol{\mu}, \boldsymbol{\lambda}^J, \boldsymbol{\lambda}^I) \in \mathbb{R}^{5K}$. The form of the Newton iteration in $\mathcal{I}$ applied to Lagrangean (10) retains its form. The blocks in $\mathbf{H}_{\mathbf{x},\boldsymbol{\mu}}L^{\psi}(\mathbf{z}, t)$ retain their diagonality under Condition C1. The new equation system in $\mathcal{I}$ consists of equations $\nabla_{\mathbf{I}} L^{\psi}(\mathbf{z}(n+1), t) = 0$, $\nabla_{\boldsymbol{\lambda}^J} L^{\psi}(\mathbf{z}(n+1), t) = 0$, $\nabla_{\boldsymbol{\lambda}^I} L^{\psi}(\mathbf{z}(n+1), t) = 0$ and $\frac{\delta}{\delta t} L^{\psi}(\mathbf{z}(n+1), t) = 0$ and can be shown to be equivalent to the formulation in (8), with

[4] Despite of the convergence, iteration $\mathcal{I}$ is in general not a monotone descent iteration. The descent property retaining quadratic convergence may be however enforced by introducing in $\mathcal{I}$ a damping factor sequence $\{a(n)\}$, such that for some $n_0 \in \mathbb{N}$ holds $a(n) = 1$, $n \geq n_0$ [9].

[5] Notice that global convergence, independent of $\mathbf{z}(0)$, is prevented only by the fact that the construction of L^{ψ} enforces some stationary points corresponding to some infeasible $(\mathbf{x}, \mathbf{I})$, which are potential points of attraction.

[6] Notice that now we have to search for the stationary point of L^{ψ} with respect to the pair $(\mathbf{z}, t)$.

$\boldsymbol{\lambda}(n+1) \equiv -\boldsymbol{\lambda}^J(n+1)$ and $\psi(\lambda_k^I) \equiv -\lambda_k^J$, $k \in \mathcal{K}$. Under the current problem formulation and the required uniqueness of the saddle point $\tilde{\mathbf{z}}$ as a feasible stationary point, the characterization (5-6) translates into

$$\begin{cases} \tilde{\mathbf{z}} = \arg \max_{\mathbf{I}\in\mathbb{R}^K} \max_{(\boldsymbol{\mu},\boldsymbol{\lambda}^J)\in\mathbb{R}^{2K}} \min_{\mathbf{x}\in\mathbb{R}^K} \min_{\boldsymbol{\lambda}^I\in\mathbb{R}^K} L^\psi(\mathbf{z}) \\ \tilde{\mathbf{z}} = \arg \min_{\boldsymbol{\lambda}^I\in\mathbb{R}^K} \min_{\mathbf{x}\in\mathbb{R}^K} \max_{(\boldsymbol{\mu},\boldsymbol{\lambda}^J)\in\mathbb{R}^{2K}} \max_{\mathbf{I}\in\mathbb{R}^K} L^\psi(\mathbf{z}). \end{cases} \tag{11}$$

The sufficient characterization of (ϕ, J_k^e), $k \in \mathcal{K}$ enforcing the uniqueness property of the saddle point, and hence allowing for semi-local convergence of $\mathcal{I}$ to the globally optimal power allocation, can now be given.

Proposition 3 *Under Condition C1, the Lagrangean (10) has a saddle point (11) which is the unique feasible stationary point, if the functions $\phi_e(y) := \phi(e^y)$ and J_k^e, $k \in \mathcal{K}$ satisfy*

$$\begin{cases} \phi_e''(y) + \phi_e'(y) \leq 0 \\ \phi_e''(y) \geq 0 \\ \frac{\delta^2 J_k^e(\mathbf{x})}{\delta x_l^2} \geq 0, \ \ k, l \in \mathcal{K}, \end{cases} \tag{12}$$

with either of the inequalities strict.

Remark on the proof: The result follows from the second order characterization of so-called max-min functions and convex-concave functions [7] applied to (10).

The function class characterized by (12) includes the functions $\phi(y) = -\log(y)$ or $\phi(y) = 1/y$ under the convexity of the interference functions J_k^e, $k \in \mathcal{K}$ (e.g. linear receiver case (2)). Hence, a particular implication of Proposition 3 is, that $\mathcal{I}$ provides semi-local quadratic convergence to the global minimizer in the problem of optimization of weighted sum of link data-rates (in high power regime) and bit-error-rate slopes in linear receiver networks.

6. DISTRIBUTED REALIZATION SCHEME

On the example of problem formulation (3), in this section we present an implementation scheme of iteration $\mathcal{I}$, assuming linear interference functions under no self interference, i.e. taking (2) with $V_{kk} = 0$, $k \in \mathcal{K}$. The main role in the distributed scheme plays the concept of adjoint network feedback from [5], [6]. This concept allows for the provision of the values $\sum_{j\neq\mathcal{K}} V_{jk}a_k$ ($\mathbf{a} \in \mathbb{R}_+^K$) to the corresponding transmitters $k \in \mathcal{K}$ by means of a concurrent transmission of all receivers $k \in \mathcal{K}$ under the use of channel predistortion and additional per-link (peer-to-peer) feedback of certain values (see [5], [6] for details). Denote by $\Delta_{\mathbf{x}}$, $\Delta_{\boldsymbol{\mu}}$ and $\Delta_{\mathbf{x},\boldsymbol{\mu}}$, $\Delta_{\mathbf{x},\boldsymbol{\mu}}^T$ the blocks of $\mathbf{H}_{\mathbf{x},\boldsymbol{\mu}}^{-1} L^\psi(\mathbf{z})$ corresponding to those in (7). The blocks Δ are diagonal and the k-th diagonal element of each block Δ can be easily shown to have the dependence

$$[\Delta]_{kk}(n) = [\Delta]_{kk}(e^{x_k(n)}, I_k(n), \mu_k(n), \textstyle\sum_{j\in\mathcal{K},j\neq k} \lambda_j(n)V_{jk}),$$

$k \in \mathcal{K}$, $n \in \mathbb{N}$. Exactly the same type of dependence holds also for the corresponding gradient components $[\nabla_{\mathbf{x}}]_k(n) := [\nabla_{\mathbf{x}} L^\psi(\mathbf{z}(n))]_k$ and $[\nabla_{\boldsymbol{\mu}}]_k(n) := [\nabla_{\boldsymbol{\mu}} L^\psi(\mathbf{z}(n))]_k$, $k \in \mathcal{K}$, $n \in \mathbb{N}$. The elementwise description of iteration $\mathcal{I}$ is now given by (8) and

$$\begin{cases} x_k(n{+}1) = x_k(n) - [\Delta_{\mathbf{x}}]_{kk}(n)[\nabla_{\mathbf{x}}]_k(n) - [\Delta_{\mathbf{x},\boldsymbol{\mu}}]_{kk}(n)[\nabla_{\boldsymbol{\mu}}]_k(n) \\ \mu_k(n{+}1) = \mu_k(n) - [\Delta_{\mathbf{x},\boldsymbol{\mu}}]_{kk}(n)[\nabla_{\mathbf{x}}]_k(n) - [\Delta_\mu]_{kk}(n)[\nabla_{\boldsymbol{\mu}}]_k(n). \end{cases} \tag{13}$$

With this, a decentralized realization scheme of $\mathcal{I}$ can be described as follows (assumed is the knowledge of $\hat{p}_k$ at the transmitters $k \in \mathcal{K}$ and the knowledge of α_k and ϕ at the transmitters and receivers $k \in \mathcal{K}$).

1. Concurrent transmission with $\mathbf{p} = \exp(\mathbf{x}(n))$;
   $\rightarrow$ $e^{x_k(n)}$ and $J_k^e(\mathbf{x}(n))$ obtainable at receivers $k \in \mathcal{K}$
   $\rightarrow \lambda_k(n)$ and $I_k(n)$ computable at receivers $k \in \mathcal{K}$ from (8).
2. Per-link feedback of values $I_k(n)$ to transmitters $k \in \mathcal{K}$ and adjoint network feedback with signal powers $\lambda_k(n)$, $k \in \mathcal{K}$ [5];
   $\rightarrow \sum_{j\in\mathcal{K}\,,\,j\neq k} \lambda_j(n)V_{jk}$ computable at transmitters $k \in \mathcal{K}$.
3. Computation of component iterations (13) at transmitters $k \in \mathcal{K} \rightarrow n := n + 1$.

7. REFERENCES

[1] J. Mo, J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Transactions on Networking*, No. 8(5), pp. 556-567, Oct 2000

[2] M. Chiang, J. Bell, "Balancing Supply and Demand of Bandwidth in Wireless Cellular Networks: Utility Maximization over Powers and Rates," *IEEE INFOCOM'04*, Mar 2004

[3] F. P Kelly, "Charging and rate control for elastic traffic (corrected version)," *European Transactions on Telecommunications*, Vol. 8, No. 1, pp. 33-37, Jan 1997

[4] H. Boche, M. Wiczanowski, S. Stanczak, "Characterization of Optimal Resource Allocation in Cellular Networks," *IEEE Workshop on Signal Processing Advances in Wireless Communications*, Jul 2004

[5] S. Stanczak, M. Wiczanowski, H. Boche, "Distributed Power Control for Optimizing a Weighted Sum of QoS Values," *IEEE Global Telecommun. Conf. (Globecom 2005)*, Nov 2005

[6] M. Wiczanowski, S. Stanczak, H. Boche, "Distributed Optimization and Duality in QoS Control for Wireless Best-Effort Traffic," *Asilomar Conference on Signals, Systems and Computers*, Oct 2005

[7] Yu. G. Evtushenko, "Some Local Properties of Minimax Problems," *USSR Computational Mathematics and Mathematical Physics*, Vol. 14, No. 3, 1974

[8] D. G. Luenberger, "Optimization by Vector Space Methods," *John Wiley & Sons, New York*, 1969

[9] J. M. Ortega, W.C. Rheinboldt, "Iterative Solution of Nonlinear Equations in Several Variables," *Academic Press, New York*, 1970