14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

AUTONOMOUS INTERFERENCE CONTROL FOR WIRELESS MESH AND AD-HOC NETWORKS - THE GENERALIZED LAGRANGIAN APPROACH

Marcin Wiczanowski, Slawomir Stanczak† and 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

Invited Paper

ABSTRACT

We consider the combined problem of performance optimization and interference control in wireless mesh and ad-hoc networks. Relying on the specific construction of the generalized Lagrangian function we propose a simple primal-dual unconstrained iteration providing convergence to a (local) optimum under arbitrary performance objectives. We present a decentralized implementation of such routine in linear networks.

1. INTRODUCTION

Over the last decade we have observed a lively evolution of wireless ad-hoc networking, established within the IEEE 802.11, 802.15 and 802.16 standards. A related network form, regarded sometimes as a special case of an ad-hoc network and sometimes as its generalization, evolved recently and is referred to as a mesh(ed) network [1]. Wireless mesh networks may soon turn out to be a disruptive technology for home/ neighborhood and enterprise networking. The potential of mesh networks already pushed forward the activities in reinvention and enhancement of existing MAC and routing concepts for ad-hoc networks, aiming especially at the improvement of scalability achieved in mesh networks. A wireless mesh network consists of a mesh of wireless access areas of *mesh clients*, connected (in general) by the mixed wired-wireless backbone of *mesh routers* and having a backhaul access to some wired network. In the case of so called *infrastructure* or *backbone meshing*, mesh clients access the mesh routers, e.g. for backhaul internet access, or communicate peer-to-peer across different wireless access areas through the mesh routers as in the multihop ad-hoc network. In the case of so-called *hybrid meshing* additionally a direct peer-to-peer communication within one access area is possible. In this way, mesh networks allow for the integration and inter-communication between different wireless standards, e.g. WLANs, cellular standards, sensor networks, etc.

A well-known key problem in ad-hoc and mesh networks is interference of links activated concurrently in different, spatially separated areas (clusters). In mesh networks such clusters correspond naturally to different wireless access areas, e.g. concurrently activated WLAN and sensor network for building surveillance. In a multihop ad-hoc network such cluster structure is determined implicitly by link scheduling. I.e., the clusters can be identified with spatially separated link ensembles, which are activated concurrently within a multi-hop transmission policy. It is therefore intuitive that the overall multihop performance of the mesh/ ad-hoc network is improved if some appropriate performance objective in each cluster is optimized, while the intercluster interference is kept at some appropriately small level. In energy-constrained networks we can think of a clusterwise objective in form of the transmit sum-energy, to be minimized subject to certain data-rate constraints per link. Similarly, under elastic traffic the objective is likely to be some weighted sum of per-link QoS parameters [2], [3], [4], [5], to be optimized in each cluster. Due to the above there is some interest in the development of efficient algorithms optimizing the clusterwise performance and controling the intercluster interference. Elementary requirements on candidate algorithms are clearly a fast convergence rate, low computational complexity (in particular, an unconstrained character with no need for paying attention to constraints may be of advantage) and a decentralized implementation, consisting in decoupled actions of nodes in the cluster.

In this work we present a performance optimization and interference control algorithm satisfying those requirements (Section 4.1). The iteration relies on a specific construction of the Lagrangian function (Section 3), which ensures positive definiteness on certain stationary points. The proposed algorithm corresponds to an unconstrained search of a saddle-point of such Lagrangian in the primal-dual domain. Without the special Lagrangian construction no iteration of comparable simplicity and convergence rate could be designed. We prove linear convergence of the algorithm (Section 4.2) and propose a decentralized implementation scheme for networks with linear receivers (Section 5).

2. PROBLEM STATEMENT AND MOTIVATION

2.1 Performance in mesh and ad-hoc networks

We consider a cluster of a multihop ad-hoc network (as explained in the introduction) or an access area to a single mesh router in the infrastructure or hybrid mesh network. In what follows we refer to both structures as clusters. We assume the set $\mathscr{A}$ to be the set of concurrently transmitting nodes in the cluster, i.e. $k \in \mathscr{A}$ are concurrently transmitting (peer-to-peer) nodes in the ad-hoc network, or nodes concurrently accessing the mesh router in the mesh network. The transmit powers are grouped in the (column) vector $\mathbf{p} = (p_1, \dots, p_{|\mathscr{A}|})$. The vector of maximal allowed per-node transmit powers is denoted by $\hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_{|\mathscr{A}|})$. The signal-to-interference-and-noise-ratio (SINR) of a link is the decisive parameter in terms of per-link quality-of-service (QoS) or performance. For instance, the data-rate or bit-error-rate as the QoS parameters of common use are monotone functions of the corresponding link-SINR. We denote by $\gamma_k(\mathbf{p})$, $k \in \mathscr{A}$ the SINRs at the receivers of peer-to-peer links, or at the mesh router. We denote by F, with $\mathbf{p} \mapsto F(\mathbf{p})$, the real-valued function expressing the efficiency of concurrent transmission of nodes $k \in \mathscr{A}$. Throughout the work we assume F to be twice Frechet-differentiable. We put no further assumptions on F, so that F is allowed to have an arbitrary number of stationary points. For instance, with real-valued function ϕ, $\gamma \mapsto \phi(\gamma)$, as the SINR-dependent QoS/ performance function, we can think here of the ratio of sum-QoS to transmit energy $F(\mathbf{p}) = \sum_{k\in\mathscr{A}} \phi(\gamma_k(\mathbf{p}))/\sum_{k\in\mathscr{A}} p_k T$ (T as time of operation) as an appropriate objective for clusters with sensor nodes. Further, the weighted sum of link-QoS functions $F(\mathbf{p}) = \sum_{k\in\mathscr{A}} \alpha_k \phi(\gamma_k(\mathbf{p}))$, with traffic-dependent weights $\alpha_k \geq 0$, is an established objective mirroring the efficiency of wireless and wired communication under so-called elastic or best-effort traffic, see e.g. [2], [3], [4], [5]. With increasing ϕ and $\phi_k^{\min}$ as the minimum acceptable QoS value of k-th link we can also think of the smallest ratio of perceived and required QoS, i.e. $F(\mathbf{p}) = \min_{k\in\mathscr{A}} \phi(\gamma_k(\mathbf{p}))/\phi_k^{\min}$, as the objective suitable for clusters with minimum service traffic, like e.g. delay-constrained multimedia streams (here we do not have Frechet-differentiability).

2.2 Interference control and problem formulation

We denote by $\mathscr{B} \ni k$ the set of certain nodes located near the boundary of the cluster and referred to as *(interference-) critical* nodes. Due to their location, the strength of all link signals measured at such nodes allows for the prediction of the interference caused by the cluster signals within the neighbor clusters. By setting suitable local constraints $\hat{g}_k$ on certain real-valued functions g_k, $\mathbf{p} \mapsto g_k(\mathbf{p})$, $k \in \mathscr{B}$, expressing the signal power measured at critical nodes, the interference to neighbor clusters can be kept at an acceptable level. In particular, under the use of linear receivers the suitable function g_k at node $k \in \mathscr{B}$ is the directly received signal power, i.e. $g_k(\mathbf{p}) = \sum_{j\in\mathscr{A}} V_{jk} p_j + n_k$, with n_k as the background noise variance and V_{jk} as the squared magnitude of the channel coefficient to the j-th transmitter (if $k \in \mathscr{A} \cap \mathscr{B}$ we set $V_{kk} = 1$ since the transmission and measurement locations coincide). Summarizing the above, the joint problem of performance optimization and inter-cluster interference control in mesh/ ad-hoc network can be formulated for each cluster as

$$\min_{\mathbf{p}} F(\mathbf{p}), \quad \text{subject to} \begin{cases} -\mathbf{p} \leq 0, \quad \mathbf{p} - \hat{\mathbf{p}} \leq 0 \\ g_k(\mathbf{p}) - \hat{g}_k \leq 0, k \in \mathscr{B}, \end{cases} \tag{1}$$

where we assumed that the global minimum of F represents the optimum performance. In the remainder we restrict the class of objectives F very slightly by concentrating on those, which can be referred to as *fair*.

Definition 1 *Assuming $\tilde{\mathbf{p}}$ to be a minimizer in problem (1), we refer to F as to a fair objective iff $\tilde{\mathbf{p}} > 0$.*

Clearly, fairness of F is in particular implied by the property $\lim_{n\to\infty} F(\mathbf{p}^{(n)}) = \infty$, with $\lim_{n\to\infty} \mathbf{p}^{(n)} = \tilde{\mathbf{p}}$, such that there exists $k \in \mathscr{A}$, such that $\tilde{p}_k = 0$. In particular, from [3] is known that a weighted sum of data-rates and a weighted sum of bit-error-rates are fair objectives. Fairness of the objective allows for a bijective transformation of the optimization domain of the form $\mathbf{x} = \log(\mathbf{p})$ (it is assumed throughout that $\log(\mathbf{a}) = (\log a_1, \dots \log a_{|\mathscr{A}|})$ and $\exp(\mathbf{a}) = (e^{a_1}, \dots, e^{a_{|\mathscr{A}|}})$, $\mathbf{a} \in \mathbb{R}^{|\mathscr{A}|}$). This allows for the simplification of (1) by dropping the nonnegativity constraint, which results in the equivalent formulation of (1) in the form

$$\min_{\mathbf{x}} F(\exp(\mathbf{x})), \text{ s.t.} \begin{cases} \exp(\mathbf{x}) - \hat{\mathbf{p}} \leq 0 \\ g_k(\exp(\mathbf{x})) - \hat{g}_k \leq 0, k \in \mathscr{B}. \end{cases} \tag{2}$$

Notice, that the first constraint is convex. Hence, convexity of the problem, or at least the existence of only global minimizers, is determined purely by the properties of F and g_k, $k \in \mathscr{B}$.

2.3 Requirements on the algorithmic solution

It is easy to agree on three essential requirements ensuring efficient implementation of a candidate algorithmic solution to (2) in mesh and ad-hoc networks.

1.) The iteration should exhibit fast local convergence to the (local) minimizer of (2) under a maintainable complexity. Clearly, if F and g_k, $k \in \mathscr{B}$ imply the existence of only global minimizers, the convergence is to the global minimizer. A linear or superlinear quotient convergence (see Section 4.1) appears to us to be sufficient.

2.) The iteration should be of unconstrained nature. This means that, while the obtained minimizer is obviously *feasible* (satisfies all constraints in (2)), no attention needs to be payed for feasibility of consecutive iterates. This brings complexity advantages and prevents the deterioration of the convergence speed due to projecting of infeasible iterates onto the optimization domain, e.g. by the gradient projection [8]. Furthermore, unconstrained iteration is in general a necessary condition for decentralized implementation, since in general ensuring feasibility of the iterates requires global knowledge at each iterating node.

3.) A decentralized implementation, requiring local actions at certain nodes based on their local knowledge, should be possible. The provision of necessary local knowledge by means of peer-to-peer feedback at each link is maintainable. The algorithm proposed in Section 4.1 satisfies all three requirements.

3. GENERALIZED LAGRANGIAN

The fulfillment of the above requirements needs a specific construction of a generalized Lagrangian as a basis for the proposed iteration. We begin with some basic notions from optimization theory. For notational simplicity in this and the next section we use a uniform formulation of all constraints in (2) as $h_k(\mathbf{x}) \leq 0$, $k \in \mathscr{K}$, with set $\mathscr{K}$ such that $|\mathscr{K}| = |\mathscr{A}| + |\mathscr{B}|$.

3.1 Some optimization-theoretic notions

It is well-known that the common (i.e. linear) form of the Lagrangian function of the studied problem is now

$$L_0(\mathbf{x}, \lambda) = F(\mathbf{x}) + \sum_{k\in\mathscr{K}} \lambda_k h_k(\mathbf{x}), \tag{3}$$

with $\lambda = (\mu_1, \dots, \mu_{|\mathscr{K}|}) \geq 0$ as the vector of dual variables. We let $\mathscr{T}(\mathbf{x}) := \{k \in \mathscr{K} : h_k(\mathbf{x}) = 0\}$ be the set of tight (i.e. satisfied with equality) constraints at $\mathbf{x}$.

Definition 2 *Strict complementarity is said to hold at* $(\mathbf{x},\lambda)$ *iff* $\lambda_k \neq 0$, $k \in \mathscr{T}(\mathbf{x})$.

Strict complementarity at $\mathbf{x}$ can be interpreted as the *actual* tightness of the constraints, in the sense that loosening of any constraint at $\mathbf{x}$ implies the existence of some feasible $\tilde{\mathbf{x}}$, such that $F(\tilde{\mathbf{x}}) < F(\mathbf{x})$.

Definition 3 *Constraint qualification is said to hold at* $\mathbf{x}$ *iff* $\nabla h_k(\mathbf{x})$, $k \in \mathscr{T}(\mathbf{x})$ *are linearly independent.*

In all statements of the paper it is implicitly assumed that the constraint qualification holds.

Definition 4 *The second order sufficiency conditions (SOSC) are said to be satisfied at the stationary point* $(\mathbf{x},\lambda)$ *of the Lagrangian (3) of problem (2) iff*
i.) $(\mathbf{x},\lambda)$ *satisfies the Karush-Kuhn-Tucker (KKT) conditions,*
ii.) $\mathbf{x}^T\nabla^2_{\mathbf{x}}L_0(\mathbf{x},\lambda)\mathbf{x} > 0$ *for* $\mathbf{x} \neq 0$, *satisfying* $\nabla^T h_k(\mathbf{x})\mathbf{x} = 0$ *for* $k \in \mathscr{T}(\mathbf{x})$ *with* $\lambda_k > 0$ *and satisfying* $\nabla^T h_k(\mathbf{x})\mathbf{x} \leq 0$ *for* $k \in \mathscr{T}(\mathbf{x})$ *with* $\lambda_k = 0$.

SOSC are of immense importance in the development and analysis of locally convergent algorithms for nonconvex optimization problems, since they distinguish the local minimizers of the problem from other stationary points of the Lagrangian. Note that under strict complementarity $k \in \mathscr{T}(\mathbf{x})$ implies $\lambda_k \neq 0$.

3.2 Lagrangian Construction

One of the aims of our Lagrangian construction is the property of positive definiteness at points satisfying SOSC. This is the necessary condition for convergence of the proposed unconstrained iteration and is not guaranteed by the common Lagrangian (3) in the very most cases. The first Lagrangian construction with such property, applicable to equality constrained problems, has been proposed by Hestenes [10] (see also [11]). In [12] and [13] related sophisticated constructions have been presented. We propose a Lagrangian different from the ones in [12], [13].

Definition 5 *We define the Lagrangian*

$$L(\mathbf{x},\mu,c) = F(\mathbf{x}) + \sum_{k\in\mathscr{K}} \psi(\phi(\mu_k)h_k(\mathbf{x}) + c), \qquad (4)$$

with some constant $c \geq 0$, *twice differentiable function* $\psi : \mathbb{R} \to I$, $I \subseteq \mathbb{R}$, *satisfying*
Conditions 1:
i.) $\psi'(y) > 0$, $y \in \mathbb{R}$ *(increasingness),*
ii.) $\psi''(y) > 0$, $y \in \mathbb{R}$ *(strict convexity),*
iii.) $\psi''(y)$ *is nondecreasing,*
and function $\phi : \mathbb{R} \to \mathbb{R}_+$ *satisfying*
Conditions 2:
i.) $\phi(y) = \phi(-y)$, $y \in \mathbb{R}$ *(evenness),*
ii.) $\phi(y) \geq 0$, $y \in \mathbb{R}$ *(nonnegativity),*
iii.) ($\phi(y) = 0$ *iff* $y = 0$*) and (*$\phi'(y) = 0$ *iff* $y = 0$*) and* $\phi''(0) > 0$ *(unique minimum with value* 0 *at* 0*).*

Numerous functions satisfy Conditions 1 and 2. Prominent examples are $\psi(y) = e^y$ and $\phi(y) = y^{2n}$, $n \in \mathbb{N}_+$. For the arguments in (4) we should sometimes write shortly $\mathbf{z} := (\mathbf{x},\mu)$.

The following Proposition characterizes a one-to-one connection between points satisfying the KKT conditions and stationary points of $L(\mathbf{z},c)$.

Proposition 1 *Let* $\lambda = (\lambda_1,\ldots,\lambda_{|\mathscr{K}|}) \geq 0$, *and for any* $\mu \in \mathbb{R}^{|\mathscr{K}|}$ *let* $\nu(\mu) := \nu = (\nu_1,\ldots,\nu_{|\mathscr{K}|})$ *be any of the vectors satisfying* $\nu_k = \pm\mu_k$ *and* $\lambda_k = \psi'(c)\phi(\nu_k)$, $k \in \mathscr{K}$. *Point* $(\mathbf{x},\lambda)$, *with* $\mathbf{x}$ *feasible, satisfies the KKT conditions of problem (2) if and only if each point* $(\mathbf{x},\nu)$ *is a stationary point of Lagrangian (4).*

Sketch of the proof: Assume first $(\mathbf{x},\lambda)$ satisfies the KKT conditions. Then, due to complementary slackness conditions and $\nabla_{\mathbf{x}}L_0(\mathbf{x},\lambda) = 0$ we have

$$\nabla_{\mathbf{x}}F(\mathbf{x}) + \sum_{k\in\mathscr{T}(\mathbf{x})} \lambda_k \nabla_{\mathbf{x}}h_k(\mathbf{x}) = 0. \qquad (5)$$

With the assumptions and Conditions 1, 2 follows further

$$\begin{cases} \lambda_k = 0 \text{ iff } \mu_k = 0 \text{ iff } \phi(\mu_k) = 0 \\ \lambda_k > 0 \text{ iff } \pm\mu_k \neq 0 \text{ iff } \phi(\pm\mu_k) > 0. \end{cases} \qquad (6)$$

The assumptions with respect to λ and ν, complementary slackness conditions, (5) and (6) yield now together

$$\nabla_{\mathbf{x}}L(\mathbf{x},\nu,c) = \nabla_{\mathbf{x}}F(\mathbf{x}) + \sum_{k\in\mathscr{T}(\mathbf{x})} \phi(\nu_k)\psi'(c)\nabla_{\mathbf{x}}h_k(\mathbf{x}) = 0.$$

For any $(\mathbf{x},\mu)$ we have further $[\nabla_\mu L(\mathbf{x},\mu,c)]_k = \psi'(\phi(\mu_k)h_k(\mathbf{x}) + c)\phi'(\mu_k)h_k(\mathbf{x})$, $k \in \mathscr{K}$. We get $[\nabla_\mu L(\mathbf{x},\nu,c)]_k = 0$ for $k \in \mathscr{T}(\mathbf{x})$ immediately due to constraint tightness and the same for $k \notin \mathscr{T}(\mathbf{x})$ due to complementary slackness and (6). Conversely, $h_k(\mathbf{x}) \leq 0$ holds already by assumption and $\lambda \geq 0$ by assumptions and Conditions 1, 2. Starting now with $\nabla L(\mathbf{x},\nu,c) = 0$, implies

$$\begin{cases} \nabla_{\mathbf{x}}F(\mathbf{x}) + \sum_{k\in\mathscr{K}} \phi(\nu_k)\psi'(\phi(\nu_k)h_k(\mathbf{x}) + c)\nabla_{\mathbf{x}}h_k(\mathbf{x}) = 0 \\ \psi'(\phi(\nu_k)h_k(\mathbf{x}) + c)\phi'(\nu_k)h_k(\mathbf{x}) = 0 \quad k \in \mathscr{K}. \end{cases} \qquad (7)$$

The complementary slackness conditions and $\nabla_{\mathbf{x}}L_0(\mathbf{x},\lambda) = 0$ can be now obtained in the similar manner as above by applying to (7) Conditions 1 and 2, (6) and the assumptions with respect to λ and ν.

It can be observed that due to evenness of ϕ a single KKT point of problem (2) corresponds to $2^{|\mathscr{K}|}$ stationary points of $L(\mathbf{z},c)$, which are equivalent in terms of the objective value. In terms of iteration convergence it is therefore not of interest which of such points is the point of attraction of the iteration. The next statement concerns the announced positive semidefiniteness of $L(\mathbf{z},c)$ at any SOSC point.

Proposition 2 *If the point* $\tilde{\mathbf{z}} = (\tilde{\mathbf{x}},\tilde{\mu})$ *is a stationary point of* $L(\mathbf{z},c)$, *such that* $(\tilde{\mathbf{x}},\tilde{\lambda})$ *related to* $\tilde{\mathbf{z}}$ *as in Proposition 1 satisfies SOSC and strict complementarity conditions, then for all* $c \geq c_0$, *with some finite* c_0 *holds* $\nabla^2_{\mathbf{x}}L(\tilde{\mathbf{z}},c) \succ 0$.

Sketch of the proof: We have in general

$$\begin{aligned} &\nabla^2_{\mathbf{x}}L(\mathbf{z},c) = \nabla^2_{\mathbf{x}}F(\mathbf{x}) + \\ &\sum_{k\in\mathscr{K}} (\phi^2(\mu_k)\psi''(\phi(\mu_k)h_k(\mathbf{x}) + c)\nabla_{\mathbf{x}}h_k(\mathbf{x})\nabla^T_{\mathbf{x}}h_k(\mathbf{x}) + \\ &\phi(\mu_k)\psi'(\phi(\mu_k)h_k(\mathbf{x}) + c)\nabla^2_{\mathbf{x}}h_k(\mathbf{x})). \end{aligned} \qquad (8)$$

If $\mathscr{T}(\tilde{\mathbf{x}}) = \varnothing$ then $\tilde{\mu} = 0$ due to $\nabla L(\tilde{\mathbf{z}},c) = 0$ (see (7)). Hence, by Conditions 1 and 2 and SOSC follows $\nabla^2_{\mathbf{x}}L(\tilde{\mathbf{z}},c) =$

$\nabla^2_{\mathbf{x}} F(\tilde{\mathbf{x}}) \succ 0$, i.e. $\tilde{\mathbf{x}}$ is a local isolated minimum of $F(\mathbf{x})$. If $\mathscr{T}(\tilde{\mathbf{x}}) \neq \emptyset$, then again by Proposition 1 and Conditions 1 and 2 $\tilde{\mu}_k = 0$, $k \notin \mathscr{T}(\tilde{\mathbf{x}})$. Hence, taking $\mathbf{z} = \tilde{\mathbf{z}}$ in (8) we get with the assumptions on $\tilde{\lambda}$ from Proposition 1

$$\nabla^2_{\mathbf{x}} L(\tilde{\mathbf{z}}, c) = \nabla^2_{\mathbf{x}} L_0(\tilde{\mathbf{x}}, \tilde{\lambda}) + \sum_{k \in \mathscr{T}(\tilde{\mathbf{x}})} \phi^2(\tilde{\mu}_k)\psi''(c)\nabla_{\mathbf{x}} h_k(\tilde{\mathbf{x}})\nabla^T_{\mathbf{x}} h_k(\tilde{\mathbf{x}}). \quad (9)$$

From strict complementarity, Proposition 1 and Conditions 1, 2 follows $\phi^2(\tilde{\mu}_k)\psi''(c) > 0$, $k \in \mathscr{T}(\tilde{\mathbf{x}})$. Due to SOSC ii.) we can now apply to (9) Debreu's Theorem [9] and conclude that there exists some finite $b_0 := \phi^2(\tilde{\mu}_k)\psi''(c_0)$, such that for all $b = b(c) := \phi^2(\tilde{\mu}_k)\psi''(c)$, $b \geq b_0$ holds $\nabla^2_{\mathbf{x}} L(\tilde{\mathbf{z}}, c) \succ 0$. Finally, with nondecreasingness of ψ'', for each $b \geq b_0$ we can find a corresponding $c \geq c_0$.

With our Lagrangian definition it is now a matter of elementary calculations to yield the following property.

Lemma 1 *If $\tilde{\mathbf{z}}$ is a stationary point of $L(\mathbf{z}, c)$, with $\tilde{\mathbf{x}}$ feasible, then $\nabla^2_{\mu} L(\tilde{\mathbf{z}}, c)$ is diagonal with $[\nabla^2_{\mu} L(\tilde{\mathbf{z}}, c)]_{kk} = 0$ for $k \in \mathscr{T}(\tilde{\mathbf{x}})$ and $[\nabla^2_{\mu} L(\tilde{\mathbf{z}}, c)]_{kk} = \psi'(c)\phi''(0)h_k(\tilde{\mathbf{x}})$ for $k \notin \mathscr{T}(\tilde{\mathbf{x}})$.*

An important corollary arises now from Proposition 2 and Lemma 1.

Corollary 1 *At any stationary point $\tilde{\mathbf{z}}$ of $L(\mathbf{z}, c)$, such that $(\tilde{\mathbf{x}}, \tilde{\lambda})$ related to $\tilde{\mathbf{z}}$ as in Proposition 1 satisfies SOSC, we have $\nabla^2_{\mu} L(\tilde{\mathbf{z}}, c) \preceq 0$ and there exists finite c such that $\nabla^2_{\mathbf{x}} L(\tilde{\mathbf{z}}, c) \succ 0$. Hence, $\tilde{\mathbf{z}}$ is a saddle point of $L(\mathbf{z}, c)$ isolated over $\mathbf{x}$, i.e. there exists some neighborhood $S(\tilde{\mathbf{x}}) \ni \tilde{\mathbf{x}}$, such that for $\mu \in \mathbb{R}^{|\mathscr{K}|}$, $\mathbf{x} \in S(\tilde{\mathbf{x}})$ holds $L(\tilde{\mathbf{x}}, \mu, c) \leq L(\tilde{\mathbf{z}}, c) < L(\mathbf{x}, \tilde{\mu}, c)$.*

Notice again that such saddle point property does not hold for the Lagrangian (3). In fact, a stationary point of $L_0(\mathbf{z})$ may satisfy $\max_{\lambda \geq 0} \min_{\mathbf{x} \text{ s. t.(2)}} L(\mathbf{x}, \lambda)$ (i.e. is a max-min point), but it is in general not a saddle-point.

4. ALGORITHM CONSTRUCTION

4.1 Optimization principle and iteration form

From the last section follows that the problem of performance optimization and interference control (2) corresponds to finding of a saddle point of the generalized Lagrangian (4). If problem (2) has only global minimizers, then any saddle point from Corollary 1 corresponds to the same objective value and we can speak of an efficiently solvable problem. Our algorithm corresponds precisely to the primal-dual search of a saddle point of $L(\mathbf{z}, c)$, based on the simple gradient method. With step-size $s > 0$ the n-th iteration step can be formulated as

$$\mathbf{z}(n+1) = \mathbf{z}(n) + s\ diag(-\mathbf{I}_{|\mathscr{A}|}, \mathbf{I}_{|\mathscr{K}|})\nabla L(\mathbf{z}(n), c) \quad (10)$$

($\mathbf{I}_n$ is identity matrix of size n). This implies that primal steps (over the logarithmic transmit powers $\mathbf{x}$) and dual steps (over the vector of dual variables μ) are conducted concurrently. By the Lagrangian construction there are no constraints on the dual variables μ. Moreover, iteration (10) provides local convergence to a local minimum of (2) without requiring feasibility of $\mathbf{x}(n+1)$ after each iteration $n \in \mathbb{N}$. This is shown in the next Proposition and implies that the algorithm requires no steps or actions in addition to (10) at all.

Proposition 3 *Assume that $\tilde{\mathbf{x}}$ is a (local) minimizer of (2), such that strict complementarity condition holds at $(\tilde{\mathbf{x}}, \tilde{\lambda})$. Then, under the step-size choice*

$$s < \min_{1 \leq k \leq |\mathscr{A}| + |\mathscr{K}|} \frac{2 Re\,\sigma_k(diag(-\mathbf{I}_{|\mathscr{A}|}, \mathbf{I}_{|\mathscr{K}|})\nabla^2 L(\tilde{\mathbf{z}}, c))}{|\sigma_k(diag(-\mathbf{I}_{|\mathscr{A}|}, \mathbf{I}_{|\mathscr{K}|})\nabla^2 L(\tilde{\mathbf{z}}, c))|^2}, \quad (11)$$

with $\sigma_k(\cdot)$ as the eigenvalues, $\tilde{\mathbf{z}}$ is a point of attraction of iteration (10) for $c \geq c_0$ with some finite c_0.

Sketch of the proof: By Proposition 2, if $\tilde{\mathbf{x}}$ minimizes $F(\mathbf{x})$ locally, then $\tilde{\mathbf{z}}$ is a stationary point of $L(\mathbf{z}, c)$, with $L(\tilde{\mathbf{z}}, c) \succ 0$ for some $c \geq c_0$. Since (10) is a gradient-based method, any stationary point of $L(\mathbf{z}, c)$ is an equilibrium point of mapping $G(\mathbf{z}) := \mathbf{z} + s\ diag(-\mathbf{I}_{|\mathscr{A}|}, \mathbf{I}_{|\mathscr{K}|})\nabla L(\mathbf{z}, c)$ [14], with gradient $\nabla G(\mathbf{z}) = \mathbf{I} + s\nabla^2_{(-)} L(\mathbf{z}, c)$, $\nabla^2_{(-)} L(\mathbf{z}, c) := diag(-\mathbf{I}_{|\mathscr{A}|}, \mathbf{I}_{|\mathscr{K}|})\nabla^2 L(\mathbf{z}, c)$. Taking $\mathbf{z} = \tilde{\mathbf{z}}$, we have

$$\nabla^2_{(-)} L(\tilde{\mathbf{z}}, c) = \begin{bmatrix} -\nabla^2_{\mathbf{x}} L(\tilde{\mathbf{z}}, c) & -\nabla^2_{\mathbf{x}\mu} L(\tilde{\mathbf{z}}, c) \\ \nabla^{2T}_{\mathbf{x}\mu} L(\tilde{\mathbf{z}}, c) & \nabla^2_{\mu} L(\tilde{\mathbf{z}}, c) \end{bmatrix}, \quad (12)$$

with $\nabla^2_{\mathbf{x}} L(\tilde{\mathbf{z}}, c)$ given by (9), $\nabla^2_{\mu} L(\tilde{\mathbf{z}}, c)$ given by Lemma 1 and $\nabla^{2T}_{\mathbf{x}\mu} L(\tilde{\mathbf{z}}, c) = \nabla^2_{\mu\mathbf{x}} L(\tilde{\mathbf{z}}, c)$. For the k-th row of $\nabla^2_{\mu\mathbf{x}} L(\tilde{\mathbf{z}}, c)$ we yield from a simple calculation

$$\begin{cases} [\nabla^2_{\mu\mathbf{x}} L(\tilde{\mathbf{z}}, c)]_{k\cdot} = \psi'(c)\phi'(\tilde{\mu}_k)\nabla^T_{\mathbf{x}} h_k(\tilde{\mathbf{x}}), k \in \mathscr{T}(\tilde{\mathbf{x}}) \\ [\nabla^2_{\mu\mathbf{x}} L(\tilde{\mathbf{z}}, c)]_{k\cdot} = 0, k \notin \mathscr{T}(\tilde{\mathbf{x}}). \end{cases} \quad (13)$$

From Ostrowski's Theorem is known that $\tilde{\mathbf{z}}$ is a point of attraction of (10) if the spectral radius of $\nabla G(\tilde{\mathbf{z}})$ satisfies $\rho(\nabla G(\tilde{\mathbf{z}})) < 1$. By an elementary calculation one can show that for $s > 0$ this holds if and only if $\max_{1 \leq k \leq |\mathscr{A}| + |\mathscr{K}|} Re\,\sigma_k(\nabla^2_{(-)} L(\tilde{\mathbf{z}}, c)) < 0$ and (11) hold. By block-skew-symmetry of $\nabla^2_{(-)} L(\mathbf{z}, c)$ we have $Re\,\sigma_k(\nabla^2_{(-)} L(\tilde{\mathbf{z}}, c)) = Re(-\mathbf{v}^H_k \nabla^2_{\mathbf{x}} L(\tilde{\mathbf{z}}, c)\mathbf{v}_k + \mathbf{w}^H_k \nabla^2_{\mu} L(\tilde{\mathbf{z}}, c)\mathbf{w}_k)$, $1 \leq k \leq |\mathscr{A}| + |\mathscr{K}|$, with $\mathbf{u}_k := (\mathbf{v}_k, \mathbf{w}_k) \in \mathbb{R}^{|\mathscr{A}| + |\mathscr{K}|}$ as the k-th eigenvector of $\nabla^2_{(-)} L(\tilde{\mathbf{z}}, c)$. By Corollary 1 follows now that $\max_{1 \leq k \leq |\mathscr{A}| + |\mathscr{K}|} Re\,\sigma_k(\nabla^2_{(-)} L(\tilde{\mathbf{z}}, c)) \leq 0$. To show strict inequality assume by contradiction $Re\,\sigma_k(\nabla^2_{(-)} L(\tilde{\mathbf{z}}, c)) = 0$ for some k, which implies $Re(\nabla^2_{(-)} L(\tilde{\mathbf{z}}, c)\mathbf{u}_k) = \mathbf{0}$. By Lemma 1 and since $\nabla^2_{\mathbf{x}} L(\tilde{\mathbf{z}}, c) \succ 0$ for some $c \geq c_0$, we must then have $\mathbf{v}_k = 0$ and $\mathscr{T}(\tilde{\mathbf{x}}) \neq \emptyset$. Since $\mathbf{u}_k \neq 0$, by (13) it follows then that $\sum_{i \in \mathscr{T}(\tilde{\mathbf{x}})} a_i \nabla_{\mathbf{x}} h_i(\tilde{\mathbf{x}}) = 0$, with at least one $a_i := \psi'(c)\phi'(\tilde{\mu}_k) w^{(i)}_k > 0$. But this contradicts constraint qualifications.

4.2 Convergence behavior

For the iteration (10), written shortly $\mathbf{z}(n+1) = G(\mathbf{z}(n))$, consider first briefly the convergence in absolute errors. Such one is expressible by the *root convergence factor (of 1-st order)* defined as

$$R(\mathscr{I}, \tilde{\mathbf{z}}) = \sup_{\{\mathbf{z}(n)\}_{n} \in \mathscr{I}} \limsup_{n \to \infty} \|\mathbf{z}(n) - \tilde{\mathbf{z}}\|^{\frac{1}{n}}, \quad (14)$$

with $\mathscr{I}$ as the set of all sequences of iterates (10) convergent to a point of attraction $\tilde{\mathbf{z}}$. It is an immediate consequence from Ostrowski's Theorem [14] and the proof of Proposition 3, that for the presented algorithm holds

$$R(\mathscr{I},\tilde{\mathbf{z}}) = \rho(G(\tilde{\mathbf{z}})) < 1, \qquad (15)$$

which is referred to as linear root convergence. More interesting is the convergence of consecutive error ratios and the related value of the *quotient convergence factor (of p-th order, $p \geq 1$)*

$$Q_p(\mathscr{I},\tilde{\mathbf{z}}) = \sup_{\{\mathbf{z}(n)\}_n \in \mathscr{I}} \limsup_{n\to\infty} \frac{\|\mathbf{z}(n+1)-\tilde{\mathbf{z}}\|}{\|\mathbf{z}(n)-\tilde{\mathbf{z}}\|^p}, \qquad (16)$$

defined if $\mathbf{z}(n) \neq \tilde{\mathbf{z}}$ for all but finitely many $n \in \mathbb{N}$.

Lemma 2 *For any $\varepsilon > 0$ there exists a norm $\|\cdot\|$ on $\mathbb{R}^{|\mathscr{A}|+|\mathscr{K}|}$, such that for the iteration (10) holds $Q_1(\mathscr{I},\tilde{\mathbf{z}}) \leq \rho(G(\tilde{\mathbf{z}}))+\varepsilon$.*

Sketch of the proof: With $\{\mathbf{z}(n)\}_n \in \mathscr{I}$, for any neighborhood $S(\tilde{\mathbf{z}})$ there exists $n_0 \in \mathbb{N}$, such that $\mathbf{z}(n) \in S(\tilde{\mathbf{z}})$, $n \geq n_0$. Frechet-differentiability at $\tilde{\mathbf{z}}$ implies for any $\varepsilon > 0$ the existence of some neighborhood $S(\tilde{\mathbf{z}})$ with the corresponding n_0, such that for any $n \geq n_0$ and any norm $\|G(\mathbf{z}(n)) - G(\tilde{\mathbf{z}}) - \nabla G(\tilde{\mathbf{z}})(\mathbf{z}(n)-\tilde{\mathbf{z}})\| \leq \frac{\varepsilon}{2}\|\mathbf{z}(n)-\tilde{\mathbf{z}}\|$ [8]. Further we know that for any $\varepsilon > 0$ there exists a norm satisfying $\|\mathbf{A}\| \leq \rho(\mathbf{A}) + \frac{\varepsilon}{2}$ [8]. With $\tilde{\mathbf{z}} = G(\tilde{\mathbf{z}})$ (Proposition 3) these arguments yield for $\mathbf{z}(n) \in S(\tilde{\mathbf{z}})$

$$\begin{aligned}&\|\mathbf{z}(n+1)-\tilde{\mathbf{z}}\| = \|G(\mathbf{z}(n)) - G(\tilde{\mathbf{z}}) - \nabla G(\tilde{\mathbf{z}})(\mathbf{z}(n)-\tilde{\mathbf{z}})\\ &+\nabla G(\tilde{\mathbf{z}})(\mathbf{z}(n)-\tilde{\mathbf{z}})\| \leq \|G(\mathbf{z}(n)) - G(\tilde{\mathbf{z}}) - \nabla G(\tilde{\mathbf{z}})(\mathbf{z}(n)-\tilde{\mathbf{z}})\|\\ &+\|\nabla G(\mathbf{z})\|\|\mathbf{z}(n)-\mathbf{z}\| \leq (\rho(G(\mathbf{z}))+\varepsilon)\|\mathbf{z}(n)-\mathbf{z}\|.\end{aligned}$$

Hence, $Q_1(\mathscr{I},\tilde{\mathbf{z}}) \leq \rho(G(\tilde{\mathbf{z}}))+\varepsilon$ follows immediately.

Hence, similarly to the root convergence, due to $\rho(\nabla G(\tilde{\mathbf{z}})) < 1$ we have a linear quotient convergence of (10). With Proposition 3, result (15) and Lemma 2 we arrive at the following key result, verified in simulation in Section 6.

Proposition 4 *Under any choice of s satisfying (11), the convergence of iteration (10) to a local minimum of (2) is linear in roots and quotients.*

5. DECENTRALIZED IMPLEMENTATION

Distributed implementation of an algorithm for mesh/ ad-hoc networks is of central importance for its practical usefulness. By distributed implementation we understand the lack of provision of global knowledge of optimization parameters to all nodes, e.g. by means of the flooding protocol. Merely the use of local per-link feedback from link-destinations to corresponding link-sources appears to be maintainable. We are fortunate that routine (10) allows for decentralized implementation for certain optimization approaches. As a representative approach we take the performance optimization with interference control for best-effort traffic under the use of linear single-user receivers. With no intercluster interference control such approach has been studied widely in the framework of wired and wireless networks [2], [3], [4], [5] and is especially applicable to networks carrying data-traffic with no strict minimum-service requirements, like e.g. high-rate file exchange. As announced in Section 2, in such case the most common objective for (2) is $F(\exp(\mathbf{x})) = \sum_{k\in\mathscr{A}} \alpha_k \phi(\gamma_k(\exp(\mathbf{x})))$, with each weight $\alpha_k \geq 0$ determined by the traffic-type priority on link $k \in \mathscr{A}$. An appropriate signal measurement function at critical nodes $k \in \mathscr{B}$ is simply the received power, i.e. $g_k(\exp(\mathbf{x})) := g_k(\mathbf{x}) = \sum_{j\in\mathscr{A}} V_{jk}e^{x_j} + n_k$. Due to receiver linearity we have also for the SINR of link $k \in \mathscr{A}$

$$\gamma_k(\exp(\mathbf{x})) := \gamma_k(\mathbf{x}) = \frac{e^{x_k}}{\sum_{j\in\mathscr{A}\,,\, j\neq k} V_{jk}e^{x_j} + n_k}.$$

Let now η_k, $k \in \mathscr{B}$ be the dual variables in (4) associated with interference constraints $g_k(\mathbf{x}) - \hat{g}_k \leq 0$ and μ_k, $k \in \mathscr{A}$ be the duals associated with power constraints $e^{x_k} - \hat{p}_k \leq 0$. In the above setting the decentralized implementation of iteration (10) is possible mainly due to the concept of *adjoint network feedback* proposed in [6]. Such feedback transmission corresponds to a specific scheme of concurrent per-link feedbacks (from receivers corresponding to $k \in \mathscr{A}$), which was shown to deliver the knowledge of $[\nabla_{\mathbf{x}}F(\mathbf{x})]_k$ to each transmitter $k \in \mathscr{A}$. Fortunately, by the same feedback principle from nodes $k \in \mathscr{B}$ it is possible to make the terms $[\nabla_{\mathbf{x}} \sum_{k\in\mathscr{A}} \psi(\phi(\eta_k)(g_k(\mathbf{x}) - \hat{g}_k) + c)]_k$ and consequently the terms $[\nabla_{\mathbf{x}}L(\mathbf{z},c)]_k$ available at nodes $k \in \mathscr{A}$. Due to observability of terms $e^{x_k} - \hat{p}_k$ at nodes $k \in \mathscr{A}$ and $g_k(\mathbf{x}) - \hat{g}_k$ at nodes $k \in \mathscr{B}$, the gradient iterations over μ_k are computable locally at transmitter nodes $k \in \mathscr{A}$ and the iterations over η_k at the critical nodes $k \in \mathscr{B}$. The resulting implementation scheme can be written as follows (assumed is the knowledge of ϕ, ψ, c and s at all nodes). Due to space constraints, we must refer here to [6] for details on the adjoint network feedback concept.

Implementation of n-th iteration

1. Concurrent transmission with powers $e^{x_k(n)}$ by nodes $k \in \mathscr{A}$

   $\Rightarrow$ knowledge of $g_k(\mathbf{x}(n)), \gamma_k(\mathbf{x}(n))$ and computability of $[\nabla_\eta L(\mathbf{z}(n))]_k$ at nodes $k \in \mathscr{B}$.
2. a.) Per-link feedback of $\gamma_k(\mathbf{x}(n))$ by nodes $k \in \mathscr{A}$.
   b.) *Adjoint network feedback* with signals
   $s_k = \alpha_k\phi'(\gamma_k(\mathbf{x}(n)))\gamma_k^2(\mathbf{x}(n))/e^{x_k(n)}$
   by nodes $k \in \mathscr{A}$

   $\Rightarrow$ computability of $[\nabla_{\mathbf{x}}F(\mathbf{x}(n))]_k$ at nodes $k \in \mathscr{A}$.
3. a.) Per-link feedback of $\psi'(\phi(\eta_k(n))(g_k(\mathbf{x}(n)) - \hat{g}_k) + c)$ by nodes $k \in \mathscr{A} \cap \mathscr{B}$.
   b.) *Adjoint network feedback* with signals
   $r_k = \psi'(\phi(\eta_k(n))(g_k(\mathbf{x}(n)) - \hat{g}_k) + c)$
   by nodes $k \in \mathscr{B}$

   $\Rightarrow$computability of $[\nabla_{\mathbf{x}}L(\mathbf{z}(n),c)]_k$ at nodes $k \in \mathscr{A}$.
4. $x_k(n+1) = x_k(n) - s[\nabla_{\mathbf{x}}L(\mathbf{z}(n),c)]_k$ at nodes $k \in \mathscr{A}$,
   $\mu_k(n+1) = \mu_k(n) + s[\nabla_\mu L(\mathbf{z}(n),c)]_k$ at nodes $k \in \mathscr{A}$,
   $\eta_k(n+1) = \eta_k(n) + s[\nabla_\eta L(\mathbf{z}(n),c)]_k$ at nodes $k \in \mathscr{B}$.

6. SIMULATION RESULTS

We analyzed the convergence properties of the proposed algorithm in simulation for a mesh/ ad-hoc cluster with $|\mathscr{A}| = 10$ transmitting nodes and the set $\mathscr{B}$ of 4 critical nodes. As the optimization problem statement we took the performance optimization under best-effort traffic and linear receivers, as described in Section 5. As the QoS function of interest we nominated the data rate in high power regime, i.e. we set $\phi(\gamma) = -\log(\gamma)$. For such function ϕ it is known from [3]

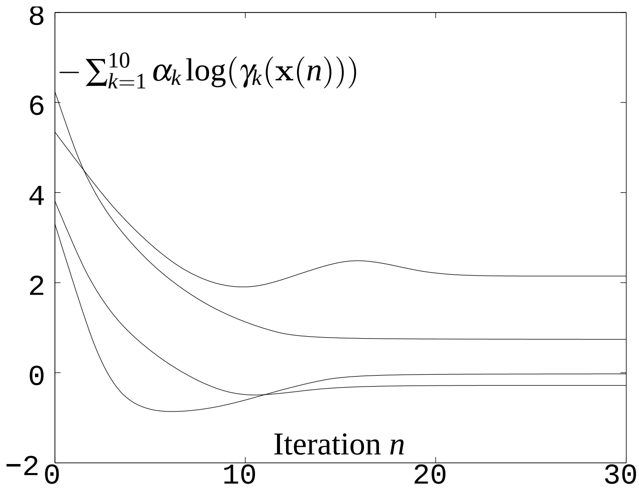

Figure 1: Convergence of the algorithm (10) for $K = 10$ and the optimization problem statement from Section 5.

that the problem has only global minimizers. In Figure 1 the convergence of the algorithm for different starting points $\mathbf{z}(0)$ and different settings of parameters V_{jk}, α_k, n_k, $\hat{g}_k$ is illustrated. For the Lagrangian (4) we chose the functions $\psi(y) = e^y$ and $\phi(y) = y^2$. In our setting, the constant c in the order of 10 turned out to be sufficiently large to ensure positive definiteness and hence local convergence.

The observable slight oscillation results directly from the unconstrained nature of iteration (10), since this allows the converging iterate sequence to be temporary in the infeasible region of variables $\mathbf{x}$.

7. CONCLUDING REMARKS

We were concerned with the problem of performance optimization with interference control in wireless mesh and ad-hoc networks. The studied problem consisted in clusterwise optimization of some arbitrary QoS/ performance measure, subject to power constraints on transmitter nodes and interference constraints on the cluster boundary, in order to control the intercluster interference. We constructed a specific nonlinear Lagrangian function with certain needed properties (positive definiteness on some stationary points), which allowed for the design of a simple iteration optimizing the problem of interest and requiring no feasibility of consecutive iterates. Linear root and quotient convergence of the proposed iteration was shown. Finally, we showed that a slightly extended implementation scheme from [6] allows for decentralized implementation of the proposed routine in linear networks under best-effort traffic.

The generalized Lagrangian framework allows for more efficient iteration designs. For instance, under combination of a different generalized Lagrangian construction with a specific approach of variable splitting, a different iteration for the optimization of weighted sum of QoS functions is proposed in [7]. The conditional Newton algorithm in [7] is unconstrained and distributedly implementable as well, while its convergence is quadratic.

REFERENCES

[1] I. F. Akyildiz, X. Wang, W. Wang, "Wireless mesh networks: a survey," *Computer Networks, to appear*

[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] 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

[4] 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

[5] 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

[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] M. Wiczanowski, S. Stanczak, H. Boche, "Quadratically Converging Decentralized Power Allocation Algorithm for Wireless Ad-Hoc Networks - The Max-Min Framework," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2006)*, May 2006

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

[9] G. Debreu, "Definite and semidefinite quadratic forms," *Econometrica*, Vol. 20, pp. 295-300, 1952

[10] M. R. Hestenes, "Multiplier and Gradient Methods," *Journal of Optimization Theory and Applications*, Vol. 4, pp. 303-320, 1969

[11] D. P. Bertsekas, "Combined Primal-Dual and Penalty Methods for Constrained Minimization," *SIAM Journal on Control and Optimization*, Vol. 13, No. 3, pp. 521-544 May 1975

[12] O. L. Mangasarian, "Unconstrained Lagrangians in Nonlinear Programming," *SIAM Journal on Control and Optimization*, Vol. 13, No. 4, pp. 772-791 May 1975

[13] R. T. Rockafellar, "Augmented Lagrange multiplier functions and duality in nonconvex programming," *SIAM Journal on Control and Optimization*, Vol. 12, pp. 268-285, 1974

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