Joint Routing and Power Control in Rayleigh-Faded Wireless Networks with ARQ Protocols

Meng Zheng
1.Key Lab of Industrial Informatics
Shenyang Institute of Automation
Chinese Academy of Sciences
Shenyang, 110016 China.
2.Graduate School of Chinese Academy of Sciences, Beijing, 100039 China
Email: zhengmeng_6@sia.cn

Slawomir Stanczak
1.Heinrich-Hertz-Lehrstuhl für Informationstheorie und theoretische Informationstechnik
Technische Universität Berlin
Einsteinufer 27, Berlin, 10587 Germany
2. Fraunhofer HHI, Berlin, 10587 Germany
Email: slawomir.stanczak@hhi.fraunhofer.de

Haibin Yu
Key Lab of Industrial Informatics
Shenyang Institute of Automation
Chinese Academy of Sciences
Shenyang, 110016 China
Email: yhb@sia.cn

Abstract—In this paper we formulate a goodput-oriented utility optimization problem for routing and power control in Rayleigh faded wireless networks with Automatic Retransmission reQuest (ARQ) protocols. This work proposes two heuristic approaches to estimate the goodput capacity in such wireless networks. The resulting approximated capacities are non-convex functions of power variables. As a result, the utility optimization problem is nonconvex, and for this we address the problem by solving a sequence of convex approximation problems. If the initial convex approximation is feasible, the sequence is shown to converge to a Karush-Kuhn-Tucker(KKT) point of the original utility optimization problem. The convex approximation problems are solved recursively by means of primal-dual methods that are shown to be amenable to distributed implementation by adjoint network. The seamless interaction between the successive convex approximation and the primal-dual algorithm constitutes the proposed successive primal-dual convex approximation (SPDCA) algorithm.

I. INTRODUCTION

Routing and power control are vital mechanisms for resource allocation in wireless networks [1]-[3]. In this paper, we consider a Rayleigh-faded wireless network with an established network topology and per-flow queuing in which the packets are passed from node to node to their destinations according to node-based routing algorithms. Detections of outage will trigger ARQ to combat the Rayleigh fading channel. In this case transmission rate is not any more equal to the reception rate. A new quantity named goodput is usually introduced to describe the actual rate of correctly transmitted bits [4]. Our focus is then on the problem of a joint optimization of transmit powers and flow rates to maximize network utility functions of goodput.

Former work [1]-[3] mainly focus on the resource allocation problem in wireless networks with error-free communications. However, error-free communication is impossible in fast fading environment since perfect channel knowledge is not available at both the transmitter and receiver due to the instantaneous feedback requirement for fast fading channel. References [5]-[6] capture the effects of erroneous nature of wireless channel and propose the rate-reliability tradeoff in wireless networks. In [7] and [8], the rate-outage constraint per link is introduced to cope with the fast fading(Rayleigh) channel. By doing so, the feedback requirement of channel state is unnecessary, but the network still suffers a little amount of fading-introduced outage.

ARQ protocol is adopted in this paper to compensate the outage in Rayleigh fading channel. The problem of joint routing and power control in wireless networks is in general difficult to solve because of the lack of convexity and separability due to the presence of interference. The introduction of Rayleigh fading channel and ARQ protocols further aggravates the difficulty in solving the optimization problem.

The first challenge is how to compute the goodput capacity for each link. The goodput capacity, also known as the maximum goodput [4], is defined as the maximal product of the physical layer transmission rate and successful transmission probability(STP) over a feasible set of transmission rates called the feasible rate region. For Rayleigh fading channel, STP has a closed form solution expressed in terms of transmission rate and power allocation. However, the power variables and rate variable are highly coupled, and thus an optimal transmission rate is difficult to calculate. We adopt two heuristic approaches to compute the goodput capacity. One is that we uniformly choose a rate from the feasible set. The other is that we find a lowerbound expression for STP, which has been proven to be tight in [8]. The adaptive transmission rate, Lambert W function [9] of 'average SIR', is also achieved correspondingly for the relaxed lowerbound of goodput capacity.

The second challenge is that the resulting goodput capacities are still inseparable and nonconcave with respect to power variables no matter which heuristic is adopted. This paper addresses the problem by combining a successive convex relaxation/approximation with primal-dual methods. More precisely, the proposed non-convex routing and power control problem is first approximated by a convex problem that is solved using primal-dual methods. The optimal solution is used to improve the approximation by stating and solving a new convex approximation problem. This procedure is repeated and the resulting sequence of optimal solutions of the convex approximation problems is shown to converge to
a KKT point of the original problem, provided that the initial
convex problem is feasible.

The third challenge results from the need for distributed
implementation. As shown, once update in the primal-dual
algorithm needs the whole information among the interfering
links. Of course, excessive message passing mechanism, such
as the "flooding" in [2], is not welcomed in practice. With the
use of so-called adjoint network [1], we can implement the
primal-dual algorithm in a distributed manner.

II. MODELING AND PROBLEM STATEMENT

A. Network and Communication Model

Consider a wireless network with an established network
topology, in which all links share a common channel. We use
\( \mathcal{N} := \{1, \ldots, N\} \) to denote the set of nodes. A number of flows
compete for access to the wireless links without scheduling.
One hop flows associated with logical links are called users.
We use \( \mathcal{W} := \{1, \ldots, K\} \), \( \mathcal{I}(n) \), and \( \mathcal{W}(n) \) to denote the set
of all logical links, the set of logical links incoming to node
\( n \in \mathcal{N} \), and the set of logical links originating at node \( n \in \mathcal{N} \),
respectively. The logical links are labeled in such a way that,
for any two nonempty sets \( \mathcal{I}(n), \mathcal{W}(m) \subset \mathcal{W} \) with \( 1 \leq n < m \leq N \)
and all \( a \in \mathcal{I}(n), b \in \mathcal{W}(m) \), there holds \( a < b \). Finally,
we assume that all nodes are synchronized and the time is
divided in slots.

Suppose that \( \mathcal{W} \) is a collection of communication sessions
and each session is identified by its unique source-destination
node pair. For any session \( w \in \mathcal{W} \), let \( O(w) \) and \( D(w) \) denote
its origin and destination nodes, respectively. The average rate
of all sessions, say \( w \), is fixed and denoted by \( \lambda(w) \). Let the
goodput flow variable of session \( w \) on link \( k \) be denoted by
\( f_k(w) \) and let \( f = (f_k(w))_{w \in \mathcal{W}, k \in \mathcal{W}} \) be the link goodput vector.
Then, by the flow conservation law, we have
\[
\forall w \in \mathcal{W} \forall n \in \mathcal{N} / \{D(w)\}, \lambda(w) I_n(w) = \sum_{k \in \mathcal{I}(n)} f_k(w) - \sum_{k \in \mathcal{W}(n)} f_k(w)
\]
where \( I_n(w) := \mathbb{I}(n = O(w)) \) and \( \mathbb{I}(\cdot) \) is the indicator function.
In addition, the total goodput cannot exceed the corresponding

goodput capacity \( c_k(p), k \in \mathcal{W} \). The goodput capacity
constraint for link \( k \) yields
\[
\sum_{w \in \mathcal{W}} f_k(w) \leq c_k(p), \quad k \in \mathcal{W}
\]
where \( p_k \geq 0 \) denotes a transmit power of user \( k \in \mathcal{W} \) and
\( p = (p_1, \ldots, p_k) \geq 0 \) is a vector of all transmit powers, referred
to as the power vector. The goodput capacity function of link
\( k, c_k(p) \), will be discussed in Section II. B in detail.

Each node, say node \( n \in \mathcal{N} \), is constrained on total power
\( P_n > 0 \). This means that
\[
p = \mathcal{P}, \mathcal{P} := \mathcal{P}_1 \times \ldots \times \mathcal{P}_N
\]
where \( \mathcal{P}_n := \{ x \in \mathbb{R}^{|\mathcal{I}(n)|} : \sum_{k = 1}^{\mathcal{I}(n)} x_k \leq P_n \} \) and \( \mathcal{I} \times \mathcal{P}_m \)
is the Cartesian product of \( \mathcal{I} \) and \( \mathcal{P}_m \). In addition, there
may be some hardware or regulatory limitations that impose
individual constraints on link transmit powers and flow rates.

These constraints will be captured by maximizing our objective
function over the following set \( \mathcal{P} := \{ (p, f) : 0 \leq p_k \leq
p_{\text{max}}(w), f_k(w) \leq f_k(w) \leq \lambda_k(w), \quad k \in \mathcal{W}, w \in \mathcal{W} \} \).

B. Goodput Capacity

We assume a Rayleigh-Rayleigh fading environment where both
desired signals and interference signals at the receivers
are subject to Rayleigh fading. In the Rayleigh-Rayleigh
fading environment, the goodput capacities \( c_k(p), k \in \mathcal{W} \)
are dictated by its transmission rate \( \mu_k \) and
\[
\text{SIR}_k(p) := \frac{V_{k,j}f_{k,j}p_k}{\sum_{l \neq k} V_{k,l}f_{k,l}p_l + \sigma_k^2}.
\]
The above notion is defined as follows: \( \sigma_k^2 > 0 \) is the variance
of an additive zero-mean Gaussian noise. \( V_{k,j} \geq 0 \) represent the path gain (not including fading) from transmitter
\( l \in \mathcal{W} \) to receiver \( k \in \mathcal{W} \). In the same way, \( F_k \geq 0 \) represent
model Rayleigh fading. They are assumed to be independent
exponentially distributed random variables with unit mean.

Given \( p \geq 0 \), the goodput capacity of \( c_k(p) \) of link \( k \) is assumed to be \( c_k(p) = \max_{\mu_k \in \mathbb{R}^+} \mu_k \Pr(\text{SIR}_k(p) \geq \gamma_k) \), where \( \Pr(\text{SIR}_k(p) \geq \gamma_k) \) with \( \gamma_k = e^{p_k} - 1 \) denotes the STP of link \( k \) when transmission rate is \( \mu_k \) and \( \mathcal{W} \) denotes the set of all
possible transmission rates. Obviously, \( c_k(p) \) depends on the
statistics of wireless fading channel, power and rate control.
Following we propose two methods for estimating \( c_k(p) \).

1) Fixed Transmission Rate: In this case, each link \( k, l \in \mathcal{W} \)
transmits according to a fixed rate \( \mu_k > 0 \) that is uniformly
chosen from \( \mathcal{W} \). In Rayleigh-Rayleigh fading environment,
the power received at the receiver of link \( k \) from the transmitter
of link \( l \) is also an exponentially distributed random variable
with mean value. If the transmission rate is \( \mu_k \geq 0 \), then
the goodput capacity function of link \( k \) is given by
\[
c_k(p) = \mu_k e^{-\frac{\gamma_k^2}{V_{k,k}p_k}} \prod_{l \neq k} (1 + \frac{\gamma_k V_{k,l}p_l}{V_{k,k}p_k})^{-1}
\]
where STP takes the form [7]: \( \Pr(\text{SIR}_k(p) \geq \gamma_k) =
\frac{\gamma_k^2}{V_{k,k}p_k} \prod_{l \neq k} \left(1 + \frac{\gamma_k V_{k,l}p_l}{V_{k,k}p_k}\right)^{-1} \).

2) Adaptive Transmission Rate: In this case, we consider a
lower bound on STP given by [8] to approximate it
\[
\Pr(\text{SIR}_k(p) \geq \gamma_k) \geq \exp(-\gamma_k / \overline{\text{SIR}}_k(p))
\]
where \( \overline{\text{SIR}}_k(p) = \frac{V_{k,k}p_k}{\sum_{l \neq k} V_{k,l}p_l + \sigma_k^2} \) is called 'average SIR' in [8].
Then the goodput capacity is underestimated as
\[
c_k(p) = \max_{\mu_k \in \mathbb{R}^+} g_k(\mu_k, p), \quad g_k(\mu_k, p) = \mu_k \exp\left(-\frac{\gamma_k}{\overline{\text{SIR}}_k(p)}\right)
\]
Let the partial derivative of \( g_k(\mu_k, p) \) with respect to \( \mu_k \)
be zero, namely, \( \frac{\partial g_k(\mu_k, p)}{\partial \mu_k} = \exp\left(-\frac{\gamma_k}{\overline{\text{SIR}}_k(p)}\right) \left(1 - \frac{\mu_k e^{p_k}}{\overline{\text{SIR}}_k(p)}\right) = 0 \).
From this we obtain the optimal implicit solution of (6),
\( \mu_k = W(\overline{\text{SIR}}_k(p)) \), where \( W(\cdot) \) is the Lambert W function1.

1The Lambert W function, \( W(x) \), is defined to be the multivalued inverse of the
function \( f(x) = xe^x \) [9]. \( W(x) \) is positive for \( x > 0 \) and is strictly concave.
With this in hand, the optimal goodput capacity yields
\[
    c_k(p) = W(\mathcal{STR}_k(p)) \exp \left( \frac{1}{\mathcal{STR}_k(p)} - \frac{1}{W(\mathcal{STR}_k(p))} \right).
\]

(7)

\section{Optimization Problem}

In this paper, we take into account the Rayleigh fading channel and ARQ protocols by incorporating the goodput capacity into the traditional NUM framework to study the resource allocation problem over a joint space of transmit powers and routes:

\[
    U = \max_{(p,f) \in \mathcal{A}} U(f) \quad \text{s.t.} \quad (1), (2), (3), (4) \text{ or } (7) \quad (8)
\]

where \( U : \mathcal{A} \rightarrow R \) is a utility function defined to be

\[
    U(f) = \sum_{k \in \mathcal{K}} U_k(f_k) = \sum_{k \in \mathcal{K}} \sum_{w \in \mathcal{W}} a_k^{(w)} \Phi(f_k^{(w)}).
\]

(9)

Here and hereafter, \( U_k \) is the utility function associated with link \( k \in \mathcal{K} \). The utility function of rates \( \Phi : R \rightarrow R \) is assumed to be differentiable, increasing and strictly concave. The weight vector \( \omega = (\omega_1^{(w)}, \ldots, \omega_K^{(w)}), w \in \mathcal{W} \), is any fixed positive vector, which can for instance be chosen according to the back-pressure policy [3]. We assume that \( \Phi(x) \rightarrow -\infty \) and \( \Phi'(x) \rightarrow +\infty \) as \( x \rightarrow 0 \) which implies that if \((p^*, f^*)\) is a solution to (8), then \((p^*, f^*)\) is a positive solution. Since an optimal power vector is positive, we can reformulate the problem using \( s := \log(p), p \in \mathcal{P} \cap \mathcal{P}_+ \), where the logarithm is taken elementwise.

\section{Successive Convex Approximation Method}

\subsection{Utility Maximization with Fixed Transmission Rate}

In this section, we will start with the utility maximization problem (8) with fixed transmission rate (4). Taking the logarithm on all sides in (2) yields

\[
    d_k(s,f) := \log(\sum_{w \in \mathcal{W}} f_k^{(w)} + \sum_{l \neq k} \log(1 + \frac{\nu_l V_{l,k}}{V_{k,k}} e^{n-s_l})) \quad (10)
\]

All terms in (10) are convex except the first term \( \log(\sum_{w \in \mathcal{W}} f_k^{(w)}) \), while in contrast, \( \log(\sum_{w \in \mathcal{W}} f_k^{(w)}) \) is concave in \( f_k = (f_k^{(w)})_{w \in \mathcal{W}} \). As a result, the problem is difficult to solve and we approximate it by a sequence of convex approximations indexed by \( t \in \{0,1,2,\cdots\} \). These approximations are based on the fact that any concave function can be tightly upper bounded by an affine function [10, page 70]. More precisely, for the \( t \)th approximation, we choose some feasible rates \( f_k^{(w)}(t) \neq 0 \) \((k \in \mathcal{K}, w \in \mathcal{W} \)), and bound from above the first term in (10) as follows

\[
    \log(\sum_{w \in \mathcal{W}} f_k^{(w)}) \leq a_k^t(t) f_k + b_k(t) \quad (11)
\]

where \( a_k^t(t) = 1/(\sum_{w \in \mathcal{W}} f_k^{(w)}(t)/(1,\cdots,1) \) and \( b_k(t) = \log(\sum_{w \in \mathcal{W}} f_k^{(w)}(t)) - 1 \). Now, using the upperbound

(11) in (10), the \( r \)th convex relaxation of (8) with fixed transmission rate takes the form

\[
    \tilde{U}(t) := \max_{s \in \mathcal{B}(t)} U(f) \quad \text{s.t.} \quad (1), (3) \quad (12)
\]

\[
    \tilde{d}_k(s,f) := a_k^t(t) f_k + b_k(t) + \frac{\nu_l V_{l,k}}{V_{k,k}} e^{n-s_l} + \sum_{l \neq k} \log(1 + \frac{\nu_l V_{l,k}}{V_{k,k}} e^{n-s_l}) - \log(c_k) \leq 0, \quad k \in \mathcal{K} \quad (13)
\]

where \( \mathcal{B}(t) = \{(s,f) : e_k \leq p_k^{\max}, f_k^{(w)} \leq f_k^{(w)} \leq f_k^{(w)}(0), \quad k \in \mathcal{K}, \quad w \in \mathcal{W} \} \). We start the sequence of convex approximations with \( t = 0 \) and compute \( d_k^0(0), b_k(0) \) for any feasible rates \( f_k^{(w)}(0) \neq 0 \) \((k \in \mathcal{K}, w \in \mathcal{W}) \). Once the \( t \)th convex problem is solved, we use the resulting rates \( \tilde{f}(t) \) to update

\[
    a_k^t(t+1) = 1/\sum_{w \in \mathcal{W}} f_k^{(w)}(t)(1,\cdots,1), \quad b_k(t+1) = \log(\sum_{w \in \mathcal{W}} \tilde{f}^{(w)}(t)) - 1 \quad (14)
\]

and solve the \( (t+1) \)th problem (12)–(13). We repeat this process until some convergence criteria are satisfied.

Theorem 1: If the initial convex approximation is feasible, then the sequences \( \{U(t)\} \) and \( \{\tilde{s}(t), \tilde{f}(t)\} \) converge to some \( \tilde{U}^* \leq U \) and \( (s^*, f^*) \), respectively, where \( (s^*, f^*) \) is a KKT point of original problem (8).

\textbf{Proof:} Due to the lack of space, we only give the sketch of this proof. We will show that the sequence \( \{U(t)\} \) is non-decreasing and upper bounded. Further, we can easily verify that \( d_k(s,f) \) and \( \tilde{d}_k(s,f) \) satisfy the Theorem 1 of [11], which means \( \{\tilde{s}(t), \tilde{f}(t)\} \) is a KKT point of original problem (8).}

\textbf{B. Utility Maximization with Adaptive Transmission Rate}

In this section, we will proceed with the utility maximization problem (8) with adaptive transmission rate (7). Taking the logarithm on all sides in (2) yields

\[
    \log(\sum_{w \in \mathcal{W}} f_k^{(w)}) - \frac{1}{\mathcal{STR}_k(e^w)} + G_k(s) < 0. \quad (15)
\]

where \( G_k(s) = \frac{1}{W(\mathcal{STR}(e^w))} \left[ \log \frac{1}{W(\mathcal{STR}(e^w))} \right] \). According to Lemma 1 in [1], the logarithm of inverse SIR

\[
    h_k(s) := \log(1/(\mathcal{STR}(e^w))), 1 \leq k \leq K \quad (16)
\]

is a convex function of \( s \in \mathcal{R}_K \). An immediate conclusion is that \( \frac{1}{\mathcal{STR}(e^w)} \) is a convex function since log-convexity implies convexity. Thus, the first two items in (15) are concave. In contrast, the third term \( G_k(s) \) is a convex function of \( s \) as shown by the following theorem.

\textbf{Theorem 2:} \( G_k(s) \) is convex on \( \mathcal{R}_K \), i.e., \( G_k(s(\mu)) \leq (1-\mu)G_k(s) + \mu G_k(\hat{s}) \) \( \forall \hat{s}, \hat{s} \in \mathcal{R}_K \) and \( \mu \in [0,1] \).

\textbf{Proof:} According to the definition of Lambert W function, \( x = W(x)e^{W(x)} \), we differentiate both sides of it to obtain

\[
    dW(x)/dx = 1/(e^{W(x)}(1+W(x))). \quad (18)
\]
Further, based on (18), it is easy to prove that both \( \log \frac{1}{W(e^\mu)} \) and \( \frac{1}{W(e^\mu)} \) are strictly decreasing and convex with respect to \( x \).

We define the logarithm of SIR as \( g_k(x) = -\delta_k(x) \) which is a concave function of \( x \). Let \( \delta, \delta' \in \mathbb{R}^K \) with \( \delta \neq \delta' \) be arbitrary. For all \( \mu \in [0, 1] \), we have

\[
G_k(\mu) = \frac{1}{W(\Sigma e^{1-\mu} \delta e^{\mu})} + \log \frac{1}{W(\Sigma e^{1-\mu} \delta e^{\mu})} \\
\leq \frac{a}{W(\Sigma e^{1-\mu} \delta e^{\mu})} + \log \frac{1}{W(\Sigma e^{1-\mu} \delta e^{\mu})} \\
= \frac{1}{W(e^{1-\mu} \delta)} + \log \frac{1}{W(e^{1-\mu} \delta)} \\
\leq b (1 - \mu) \left( \frac{1}{W(e^\delta)} + \log \frac{1}{W(e^\delta)} \right) \\
+ \mu \left( \frac{1}{W(e^\delta)} + \log \frac{1}{W(e^\delta)} \right) \\
= (1 - \mu) G_k(\delta) + \mu G_k(\delta)
\]

where \( \leq_a \) follows from Lemma 1 in [1] and the monotonicity of \( G_k(\delta) \) and \( \leq_b \) follows from the convexity of functions \( \log \frac{1}{W(e^\mu)} \) and \( \frac{1}{W(e^\mu)} \).

In the following, we still use a sequence of convex approximations to approximate the problem (8). Similar to (11), for the \( r \)th approximation, we use an affine function to bound from above the convex function as follows: for some feasible solutions \( f_k^{(w)}(t) \neq 0, s \in \mathbb{R}^K(k \in \mathcal{K}, w \in \mathcal{W}) \),

\[
\log \sum_{w \in \mathcal{W}} f_k^{(w)} = \frac{1}{\mathcal{R}(e^\mu)} \leq \frac{a_k(t)}{\hat{a}_k(t)} \left[ \begin{array}{c} f_k \\ s \end{array} \right] + b_k(t) + \hat{b}_k(t)
\]

where \( a_k(t)/\hat{a}_k(t) = \left( -\frac{\partial \log w_k}{\partial \log w_k}, \ldots, -\frac{\partial \log w_k}{\partial \log w_k} \right) \) and \( \hat{b}_k(t) = -\frac{1}{\mathcal{R}(e^\delta)} - \hat{a}_k(t)(s,t) \). Then, using the above convex upperbound in (15), the \( r \)th convex relaxation of (8) with adaptive transmission rate takes the form

\[
\hat{U} := \max_{(s,f) \in \hat{B}} U(f) \quad \text{s.t.} \quad (1), (3), (19)
\]

Analogous to last subsection, we repeat the successive convex approximation process until convergence is achieved.

Theorem 3: Suppose that \( \hat{U}(t), \hat{f}(t) \) solves (19)-(20). If the initial convex approximation is feasible, then the sequences \( \{\hat{U}(t)\} \) and \( \{\hat{f}(t)\} \) converge to some \( \hat{U}^* \leq U \) and a KKT point of original problem (8), \( (s^*, f^*) \), respectively.

### IV. DISTRIBUTED IMPLEMENTATION

Due to the limited space, we just devise a primal-dual algorithm to solve problem (19)-(20). Let \( \mathcal{V} = (v^{(w)}_{a})_{w \in \mathcal{W}, a \in \mathcal{A}}, e = (e_{a})_{a \in \mathcal{A}}, \) and \( \theta = (\theta_{a})_{a \in \mathcal{A}} \) be Lagrange multipliers for constraints (1), (3), and (20), respectively. The physical constraints \( \mathcal{B} \) are satisfied implicitly. For brevity, let \( \Gamma \) and \( \Lambda \) denote the set of original variables \( (s,f) \) and the set of dual variables \( (v, e, \theta) \), respectively. We define the feasible set \( \Omega = \mathcal{V}^{(v)} \times \mathcal{E}^{(e)} \times \mathcal{S}^{(\theta)} \) of \( \Lambda \). The Lagrangian function associated with problem (19) is given for some \( t \in \{0, 1, 2, \ldots\} \)

\[
L(\Gamma, \Lambda) = U(f) - \sum_{n \in \mathcal{N}} \sum_{k \in \mathcal{N}} \frac{1}{e_k} \gamma_{k,n}^{(w)}(\lambda^{(w)}(\nu_{k,n}^{(w)}) - \sum_{k \in \mathcal{N}} f_k^{(w)} + \sum_{k \in \mathcal{N}} f_k^{(w)} - \theta_{k,n}^{(w)}(\nu_{k,n}^{(w)} + G_k(s)) \quad (21)
\]

It is easy to show that the Lagrangian function (21) is a concave-convex function and strong duality holds. These observations imply that the KKT conditions are necessary and sufficient conditions for optimality. Thus, we can find the maximum in (19) by solving the KKT conditions, which is equivalent to finding a stationary point \( (\Gamma, \Lambda) \) of the Lagrangian function (21). We apply the primal-dual algorithm to find a stationary point of \( L(\Gamma, \Lambda) \). In the \( (m+1) \)th iteration, we have

\[
\{ \Gamma(m+1) = \Pi \mathcal{B}[\Gamma(m) + \delta k \mathcal{L}(\Gamma(m), \Lambda(m))] \}
\]

\[
\{ \Lambda(m+1) = \Pi \mathcal{I}[\Lambda(m) - \delta k \mathcal{L}(\Gamma(m), \Lambda(m))] \},
\]

where \( \delta \) is a sufficiently small stepsize, \( \Pi \mathcal{B} [/Omega] \) and \( \Pi \mathcal{I} [/Omega] \) are the projections on \( \Omega \) and \( \mathcal{B} [/Omega] \), respectively. From (21), we compute partial derivatives of \( L(\Gamma, \Lambda) \) with respect to \( (\Gamma, \Lambda) \),

\[
\frac{\partial L}{\partial f_k^{(w)}} = \omega_k^{(w)} \phi' (f_k^{(w)}) - \frac{\theta_{k,n}^{(w)}}{\sum_{k \in \mathcal{N}} f_k^{(w)}(t) + \nu_{k,n}^{(w)}} - \delta_{k,n}^{(w)} \nu_{k,n}^{(w)}
\]

\[
\frac{\partial L}{\partial \theta_{k,n}^{(w)}} = -\epsilon_{k,n}^{(w)} \epsilon_{k,n} + \frac{1}{\theta_{k,n}^{(w)}}
\]

and

\[
\frac{\partial L}{\partial e_k} = \sum_{k \in \mathcal{N}} f_k^{(w)}(t) - \sum_{k \in \mathcal{N}} f_k^{(w)} - \lambda(w) \end{w} \]

\[
\frac{\partial L}{\partial \theta_{k,n}^{(w)}} = -\frac{\theta_{k,n}^{(w)}}{\sum_{k \in \mathcal{N}} f_k^{(w)}(t) + \nu_{k,n}^{(w)}} - \delta_{k,n}^{(w)} \nu_{k,n}^{(w)}
\]

and

\[
\frac{\partial L}{\partial \theta_{k,n}^{(w)}} = -\frac{\theta_{k,n}^{(w)}}{\sum_{k \in \mathcal{N}} f_k^{(w)}(t) + \nu_{k,n}^{(w)}} - \delta_{k,n}^{(w)} \nu_{k,n}^{(w)}
\]

where \( \gamma(t) \) and \( \nu(t) \) stand for the transmitter node and the receiver node of link \( k \), respectively.

We observe that all gradient components in (23) and (24) are local except for the terms \( \frac{\partial L}{\partial \theta_{k,n}^{(w)}} \) and \( \frac{\partial L}{\partial e_k} \). In Algorithm 1, we use the adjoint network [1] to facilitate the distributed implementation. In addition, a non fading channel with a narrow bandwidth is assumed. We denote the Rayleigh fading channel and the non fading channel as Channel 1 and Channel 2, respectively.

If \( \sigma > 0 \) is sufficiently small, then it can be shown that SPDCA converges to some \( (s^*, f^*) \) which is a KKT point of (8). For any given \( t \), the inner loop in Algorithm 1 converges to a stationary point of the associated Lagrange function (21). This immediately follows from the fact that simultaneous application of gradient methods converges to a saddle point of the associated Lagrange function [12, pp. 125–126]. Once the inner loop has converged, Theorem 1 ensures that the outer
Algorithm 1 SPDCA

Require: $\omega > 0$, $\sigma > 0$, $t = 0$, $m = 0$, $\Gamma(0)$, $\Lambda(0)$, and $\delta > 0$. 
Ensure: $(p, f)$
1: repeat 
2: Each transmitter updates affine parameters $(a_k(t), b_k(t), \bar{a}_k(t), \bar{b}_k(t))$ of the $t$th approximation.
3: repeat 
4: Over Channel 1: Concurrent transmission of one data packet at transmit power $e^{\xi(k)}$, $k \in X$, with the value of $s_k(m)$ appended in the packet.
5: Over Channel 2: Concurrent transmission of pilot sequences at powers $e^{\xi(k)}$, $k \in X$, with receiver-side estimation of $\overline{SIR}_k(e^{\xi(m)})$.
6: All receivers feed necessary estimates and variables including $\overline{SIR}_k(e^{\xi(m)})$ and $v^{(w)}(m)$ back to the corresponding transmitters via per-link control channel.
7: Over Channel 2: Concurrent transmission in the adjoint network with transmitter-side estimation of the received power. Receivers with negative $s_k(m)$ transmit first and then receivers with non-negative $s_k(m)$ transmit concurrently. The variance of the zero-mean input symbols is $|e^{\xi(i)}s_k(m)|$.
8: Transmitter side computation of $\Gamma(m)$ and $\Lambda(m)$.
9: $m = m + 1$.
10: until $L(\Gamma(m), \Lambda(m)) - L(\Gamma(m-1), \Lambda(m-1)) < \sigma$.
11: $(\tilde{s}(t), \tilde{f}(t)) = (s(m), f(m))$.
12: $t = t + 1$.
13: until $\tilde{U}(t) - U(t-1) < \sigma$.
14: $(p^*, f^*) = (\tilde{e}(\tilde{s}(t), \tilde{f}(t))$.

loop converges to some KKT point $(s^*, f^*)$ of the original problem (8) as $t \to \infty$.

V. SIMULATION

Consider a wireless network as in Fig.1. There are six logical links, four nodes and two sessions($\lambda(1 \to 4) = 0.3, \lambda(2 \to 4) = 0.3$). The path gain $V_{k,l}(k, l \in X)$ were randomly chosen from $[0, 0.2]$. Moreover, we assume $\omega = (1)_{2 \times 6}$. $P_r = 3$, $\sigma_k^2 = 0.001$, and $\Phi(x) = \log(x)$.

The convergence of Algorithm 1 is validated by Fig.2. Though there are no theoretical results regarding to the speed of convergence of successive convex approximation method, we can see that the optimal MAC-layer sequence $\{\tilde{U}(t)\}$ in SPDCA converges within 10 times approximation from Fig.2.

ACKNOWLEDGMENT

This work was supported by the China National Funds for Distinguished Young Scientists(Grant No. 60725312), The Special Program for Key Basic Research Funded by MOST(Grant No. 2010CB334705), The Important National Science and Technology Specific Project(Grant No. 2010ZX03006-005-01), the Liaoning Provincial Natural Science Foundation of China (Grant No. 20092083), and by the German Research Foundation (DFG) under grant STA864/3-1.

REFERENCES

[1] S. Stanczak, M. Wiczanowski, and H. Boche, “Distributed utility-based power control: objectives and algorithms,” IEEE Trans. Signal Process., 2007, vol.55, no.10, pp 5058–5068.
[2] M. Chiang, "Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control," IEEE J. Sel. AreasCommun., 2005, vol.23, no.1, pp 104–116.
[3] M. Neely, "Dynamic power allocation and routing for time-varying wireless networks," IEEE J. Sel. AreasCommun., 2005, vol.23, no.1, pp 89–103.
[4] I. Betjens and S. Shamai, "Optimal power and rate control for minimal average delay: the single user case," IEEE Trans. Inf. Theory, 2006, vol. 52, no. 9, pp 4115–4141.
[5] Q. Gao, J. Zhang, and S.Y. Hanly, "Cross-layer rate control in wireless network with lossy links: leak-pipe flow, effective network utility maximization and hop-by-hop algorithms," IEEE J. Sel. AreasCommun., 2009, vol. 24, no. 5, pp 3068–3076.
[6] D.O. Neill, B.S. Thian, A. Goldsmith, and S. Boyd, "Wireless NUM: rate and reliability tradeoff in random environments," IEEE WCNC, 2009.
[7] J. Papandriopoulos, J. Evans, and S. Dey, "Optimal power control for Rayleigh-faded multilayer systems with outage constraints," IEEE Trans. Wireless Commun., 2005, vol. 4, pp 2705–2715.
[8] J. Papandriopoulos, S. Dey, and J. Evans, "Optimal and distributed protocols for cross-layer design of physical and transport layers in MANETs," IEEE/ACM Trans. Netw., 2008, vol. 16, no. 6, pp 1392–1405.
[9] R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey, and D.E. Knuth, “On the Lambert W Function,” Advances Comp. Mathematics, 1995, vol. 5, pp 329–359.
[10] S. Boyd and L. Vandenberghe, Convex optimization. Cambridge, UK: Cambridge University Press, 2003.
[11] B.R. Marks and G.P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," Operations Research, 1978, vol. 26, no. 4, pp 681–683.
[12] R. T. Rockafellar, “Saddle points and convex analysis,” Differential Games and Related Topics. The Netherlands: North-Holland, 1971, pp 109–127.