Centralized Solution of Joint Source and Relay Power Allocation for AF Relay based Network

Rukhsana Ruby, *Student Member, IEEE*

I. System Model and Problem Formulation

In this section, we consider a system where there exists an access point, one relay node and N number of source nodes which need help from the relay to get their packets transmitted to the access point. Edge nodes essentially act as servers which have some special applications and destination node needs to get the content of that application on time. For example, application could be some video which needs to be displayed on the destination. Relay node amplifies the

R. Ruby is with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC, V6T 1Z4 Canada e-mail: rukhsana@ece.ubc.ca.

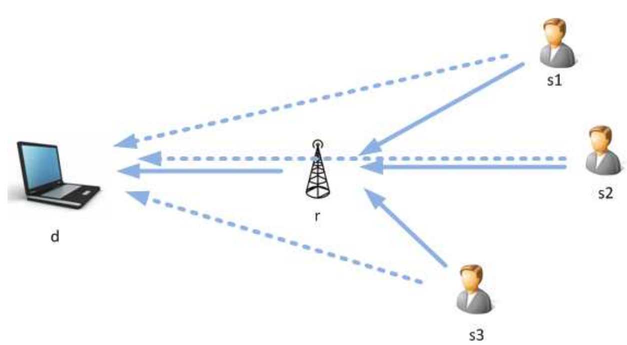


Fig. 1. System model.

received signal from the source nodes and then forwards towards the destination as depicted in Figure 1. Although the system has one destination, proposed solution can easily be adapted to a multi-destination network. Transmit channels of the sources and relay are orthogonal, either in time or frequency domain. We assume a block-fading (or quasi-static) model: the channels remain invariant over a time slot whose duration is less than the coherence time of the channels. Denote the channel gain between source s_i and destination d is $G_{s_i,d}$; the channel gain between source s_i and relay r as $G_{s_i,r}$ and the channel gain between relay r and destination d is $G_{r,d}$.

The entire transmission operation using AF relay consists of two phases (i.e., time slots). At each phase, the sources or relay use orthogonal frequency channel for multiple transmissions. At the first phase, source s_i broadcasts its information to both destination d and relay node r. The received signals $y_{s_i,d}$ and $y_{s_i,r}$ at destination d and relay r can be expressed as

$$y_{s_i,d} = \sqrt{E_{s_i} G_{s_i,d}} x_{s_i} + \eta_{s_i,d} \text{ and } y_{s_i,r} = \sqrt{E_{s_i} G_{s_i,r}} x_{s_i} + \eta_{s_i,r}, \tag{1}$$

where E_{s_i} represents the transmit power at node s_i, x_{s_i} is the broadcast information symbol with unit energy from source s_i to nodes d and r. $\eta_{s_i,d}$ and $\eta_{s_i,r}$ are the additive noises received at destination d and relay r respectively. In the second step, the relay amplifies its received signal and forwards it to destination d. Denote the power the relay uses to help source s_i is E_{r_i}. The signal received at destination d for source s_i can be shown

$$y_{r_i,d} = \frac{\sqrt{E_{r_i} G_{r,d}} \left(\sqrt{E_{s_i} G_{s_i,r}} x_{s_i} + \eta_{s_i,r} \right)}{\sqrt{E_{s_i} G_{s_i,r} + \sigma^2}} + \eta_{r_i,d}. \tag{2}$$

$\eta_{r_i,d}$ is the received noise from relay r to destination d (for source s_i). Without loss of generality, we assume that the noise power is the same additive white gaussian noise for all links, denoted by σ^2. After maximum ratio combining of both the direct and relay path, the effective received SNR for source s_i's transmission can be given by

$$\Gamma_{s_i,r,d} = \frac{E_{s_i} G_{s_i,d}}{\sigma^2} + \frac{E_{s_i} G_{s_i,r} E_{r_i} G_{r,d}}{\sigma^2 \left(E_{s_i} G_{s_i,r} + E_{r_i} G_{r,d} + \sigma^2\right)}. \tag{3}$$

If the set consisting of the source nodes is $L_s = \{s_1, s_2, \cdots, s_N\}$, total capacity achieved by the system can be given by

$$R_{s,r,d} = \gamma_L W \sum_{s_i \in L_s} log_2 \left(1 + \Gamma_{s_i,r,d}\right). \tag{4}$$

Because of the orthogonal transmissions $\gamma_L = 1/(2N)$ and W is the aggregate bandwidth in the system. Since W and γ_L are constants, we skip these terms in the subsequent discussion.

Our goal is to allocate power optimally among the sources and relay so that system capacity is maximized. Likewise traditional network resource optimization problems, there are constraints on the sources and relay power. Moreover, in order to mitigate the interference imposed on another network due to the transmission operations in this network, there is a total power constraint, meaning total power allocated to the sources and relay node cannot exceed E^{max}. For the sake of simplicity, we have converted the maximization problem into the minimization one by introducing minus sign in front of the objective function, i.e, $R_{s,r,d}$.

$$\min \prod_{s_i \in L_s} \frac{\sigma^2(\sigma^2 + E_{s_i} G_{s_i,r} + E_{r_i} G_{r,d})}{(\sigma^2 + E_{s_i} G_{s_i,d})(\sigma^2 + E_{s_i} G_{s_i,r} + E_{r_i} G_{r,d}) + E_{s_i} G_{s_i,r} E_{r_i} G_{r,d}} \tag{5}$$

$$\text{where} \quad E_{s_i} \leq E_s^{max}, \quad s_i \in L_s,$$

$$\sum_{s_i \in L_s} E_{r_i} \leq E_r^{max},$$

$$\sum_{s_i \in L_s} E_{s_i} + \sum_{s_i \in L_s} E_{r_i} \leq E^{max},$$

$$\{E_{s_i}\}_{s_i \in L_s} \geq 0, \; \{E_{r_i}\}_{s_i \in L_s} \geq 0.$$

The aforementioned optimization problem is valid if and only if $\sum_{s_i \in L_s} E_{s_i} + E_r^{max} > E^{max}$ and $\sum_{s_i \in L_s} E_{s_i}^{max} > E^{max}$.

II. CENTRALIZED SOLUTION

The problem in Equation 5 is not convex due to the non-convexity property of the objective function. This statement can be proved very easily by the help of special type of convex optimization formulation, i.e., GP [1], [2]. A GP is a type of mathematical optimization problem characterized by the objective and constraint functions that have a special form. It focuses on monomial and posynomial functions. A monomial is a function, $h : R^n \rightarrow R$, where the domain contains all real vectors with non-negative components, $h(x) = c{x_1}^{a_1}{x_2}^{a_2}\cdots{x_n}^{a_n}$. A posynomial is a sum of monomials, $f(x) = \sum_k c_k{x_1}^{a_{1k}}{x_2}^{a_{2k}}\cdots{x_n}^{a_{nk}}$. GP is an optimization problem with the form

$$minimize \quad f_0(x) \quad s.t. \quad f_i(x) \leq 1, \; h_j(x) = 1,$$

where f_0 and f_i are posynomials and h_j are monomials. This problem in the above form is not convex. However, with a change of variables: $y_i = log x_i$ and $b_{ik} = log c_{ik}$, we can transform it into convex form given the assumption that the logarithm of a sum of exponentials is a convex function.

As mentioned, the objective function is the ratio of two posynomials which cannot be solved by GP. There are ways to transform such type of problem to GP form, i.e., single condensation method, double condensation method [2]. We have used single condensation method which requires to approximate the denominator of the objective function by some monomial term. We denote the denominator by $F\left(\{E_{s_i}\}_{s_i \in L_s}, \{E_{r_i}\}_{s_i \in L_s}\right)$ and the monomial is given by

$$\prod_{s_i \in L_s} (\sigma^2 + E_{s_i}G_{s_i,d})(\sigma^2 + E_{s_i}G_{s_i,r} + E_{r_i}G_{r,d}) + E_{s_i}G_{s_i,r}E_{r_i}G_{r,d}$$
$$\approx \lambda \prod_{s_i \in L_s} E_{s_i}^{a_i}E_{r_i}^{b_i},$$

where $a_i = \frac{E_{s_i}}{F\left(\{E_{s_i}\}_{s_i \in L_s}, \{E_{r_i}\}_{s_i \in L_s}\right)} \frac{\partial F\left(\{E_{s_i}\}_{s_i \in L_s}, \{E_{r_i}\}_{s_i \in L_s}\right)}{\partial E_{s_i}}$, $b_i = \frac{E_{r_i}}{F\left(\{E_{s_i}\}_{s_i \in L_s}, \{E_{r_i}\}_{s_i \in L_s}\right)} \frac{\partial F\left(\{E_{s_i}\}_{s_i \in L_s}, \{E_{r_i}\}_{s_i \in L_s}\right)}{\partial E_{r_i}}$, and $\lambda = \frac{F\left(\{E_{s_i}\}_{s_i \in L_s}, \{E_{r_i}\}_{s_i \in L_s}\right)}{\prod_{s_i \in L_s} E_{s_i}^{a_i}E_{r_i}^{b_i}}$.

Derivations are below

$$\frac{\partial F\big(\{E_{s_i}\}_{s_i\in L_s},\{E_{r_i}\}_{s_i\in L_s}\big)}{\partial E_{s_i}} =$$
$$[G_{s_i,d}(\sigma^2+E_{s_i}G_{s_i,r}+E_{r_i}G_{r,d})+(\sigma^2+E_{s_i}G_{s_i,d})G_{s_i,r}+G_{s_i,r}E_{r_i}G_{r,d}]$$
$$\prod_{s_j\in L_s, s_j\neq s_i}\big[(\sigma^2+E_{s_j}G_{s_j,d})(\sigma^2+E_{s_j}G_{s_j,r}+E_{r_j}G_{r,d})+E_{s_j}G_{s_j,r}E_{r_j}G_{r,d}\big],$$
$$\frac{\partial F\big(\{E_{s_i}\}_{s_i\in L_s},\{E_{r_i}\}_{s_i\in L_s}\big)}{\partial E_{r_i}} =$$
$$[G_{r,d}(\sigma^2+E_{s_i}G_{s_i,d})+E_{s_i}G_{s_i,r}G_{r,d}]$$
$$\prod_{s_j\in L_s, s_j\neq s_i}\big[(\sigma^2+E_{s_j}G_{s_j,d})(\sigma^2+E_{s_j}G_{s_j,r}+E_{r_j}G_{r,d})+E_{s_j}G_{s_j,r}E_{r_j}G_{r,d}\big].$$

Finally, the overall procedures for joint source and relay power allocation is given as follows.

1) Set the initial value of power $\mathbf{E}^{(0)} := \left[E_{s_1}^{(0)},\cdots,E_{s_N}^{(0)},E_{r_1}^{(0)},\cdots,E_{r_N}^{(0)}\right]$, $n:=1$.
2) Determine $\left[a_1^{(n)},\cdots,a_N^{(n)}\right]$, $\left[b_1^{(n)},\cdots,b_N^{(n)}\right]$ and $\lambda^{(n)}$.
3) Solve the optimization problem with the help of GP.
4) Denote the optimal power allocation in the nth round as $\mathbf{E}^{(n)}$.
5) If $||\mathbf{E}^{(n)}-\mathbf{E}^{(n-1)}||\leq\epsilon$, where ϵ is a pre-defined threshold, the enumerations stop; otherwise, $n:=n+1$ and reiterate from step 2 to 5.

Notice that the above procedure updates $2N$ principle variables in every iteration. And, each iteration needs to update $2N+1$ number of intermediate variable to assist updating principle variables. In order to simplify this procedure, we can consider, the system has only one source (s^*). s^* is the representative of all sources. Gain between s^* and r is weighted average of the gains between the sources and relay. In the similar manner, gain between s^* and d is determined. After this transformation, the objective function is still the ratio of two posynomials. In order to cast it to GP, we can approximate the denominator (by denoting, $H(E_{s^*},E_r)$) of it by monomial

$$(\sigma^2+E_{s^*}G_{s^*,d})(\sigma^2+E_{s^*}G_{s^*,r}+E_rG_{r,d})+E_{s^*}G_{s^*,r}E_rG_{r,d}$$
$$\approx\mu E_{s^*}^cE_r^d,$$

where $c=\frac{E_{s^*}}{H(E_{s^*},E_r)}\frac{\partial H(E_{s^*},E_r)}{\partial E_{s^*}}$, $d=\frac{E_r}{H(E_{s^*},E_r)}\frac{\partial H(E_{s^*},E_r)}{\partial E_r}$, and $\mu=\frac{H(E_{s^*},E_r)}{E_{s^*}^cE_r^d}$.

Furthermore,

$$\frac{\partial H}{\partial E_{s^*}} = \quad G_{s^*,d}(\sigma^2+E_{s^*}G_{s^*,r}+E_rG_{r,d})+(\sigma^2+E_{s^*}G_{s^*,d})G_{s^*,r}+G_{s^*,r}E_rG_{r,d},$$

$\frac{\partial H}{\partial E_{s^*}} = (\sigma^2 + E_{s^*} G_{s^*,d}) G_{r,d} + E_{s^*} G_{s^*,r} G_{r,d}$.

Iterative procedure in order to obtain optimal E_{s^*} and E_r follows the same procedure mentioned above. However, it requires to update 2 principle variables and 3 auxiliary variables in each iteration in order to achieve convergence. Following two subsections are for distributing power $E_{s^*}^*$ among all sources and third subsection is for disseminating relay power E_r^* among all sources.

A. Source Power Allocation (Greedy Solution)

From the optimal solution, we have observed that the sources with better channel condition obtain more power compared to others. Since each source has individual power constraint and this power is moderately lower than the total allowable power for all sources, we can propose a greedy power allocation for the source nodes given the total allowable for them is $E_{s^*}^*$. If the direct link SNR of a source is better than its relayed link SNR, it is likely, that source obtains 0 relay power. Therefore, it is rational to distribute $E_{s^*}^*$ among the sources taking direct link SNR into account. We sort $G_{s_i,d}, s_i \in L_s$ in decreasing order and allocate maximum individual power to each sorted source until there is no left over power.

The above approach for distributing power among the sources is greedy. This is similar to the MaxCIR technique [3] which assigns subcarrier among the users in OFDM based networks according to their channel condition. The drawback of this approach is, the sources with worse channel may starve and may never get chance to transmit as they are assigned zero power. This reminds us one important issue which is called fairness. In order to tackle fairness, we have proposed an algorithm which considers both the instantaneous channel condition and fairness.

B. Source Power Allocation (Fair Algorithm)

At scheduling time instant t, we denote the gains between the sources and relay as a vector $\boldsymbol{\alpha}_t = \{\alpha_t(1), \alpha_t(2), \cdots, \alpha_t(N)\}$; the gains between the sources and destination as a vector

$\boldsymbol{\gamma}_t = \{\gamma_t(1), \gamma_t(2), \cdots, \gamma_t(N)\}$; the gain between the relay and destination as β_t. Moreover, individual sources' average rate as a vector $\bar{\boldsymbol{\zeta}}_t = \{\bar{\zeta}_t(1), \bar{\zeta}_t(2), \cdots, \bar{\zeta}_t(N)\}$. We initialize all elements of this vector as 0 at time $t = 0$. Fair algorithm is presented in *Algorithm 1*. Steps in the algorithm are followed in the channel coherent time of each time instant t.

Algorithm 1 Fair algorithm for subdividing power among the sources.

Get $\boldsymbol{\alpha}_t$, $\boldsymbol{\gamma}_t$ and β_t.
Sort $\boldsymbol{\gamma}_t$ in descending order and get sorted index set **I** $:= \{I_1, I_2, \cdots, I_N\}$.
Set χ := $E^*_{s^*}$.
for $\forall i \in$**I** **do**
    if $\bar{\zeta}_{t-1}(i) = 0$ **then**
        $\eta(i) := min\,(E_s^{max}, \chi)$.
        $\chi := \chi - \eta(i)$.
    end if
end for
if $\chi > 0$ **then**
    Let the last unallocated index $j := i$.
    for $\forall j \in$**I** **do**
        Set $M(j) := \frac{\gamma_t(j)}{\bar{\zeta}_{t-1}(j)}$.
    end for
    Sort index vector **I** according to descending order of vector **M**.
    for $\forall j \in$**I** **do**
        Set $\eta(j) := min\left(\frac{\chi * M(j)}{\sum_{j=i}^{N} M(j)}, E_s^{max}\right)$.
        $\chi := \chi - \eta(j)$.
    end for
end if

C. Relay Power Allocation

In order to subdivide relay power E_r^* among all sources, we have adopted water filling approach and the resultant formulated problem is given by

$$\underset{\sum_{s_i \in L_s} E_{r_i} = E_r^*}{\arg\max} \sum_{s_i \in L_s} log_2 \left(1 + \frac{E_{s_i} G_{s_i,r} E_{r_i} G_{r,d}}{\sigma^2 (E_{s_i} G_{s_i,r} + E_{r_i} G_{r,d} + \sigma^2)}\right). \tag{6}$$

By invoking the lagrange multiplier μ for the total relay power constraint of the problem in Equation 6, we obtain the lagrangian $\sum_{s_i \in L_s} log_2 \left(1 + \frac{E_{s_i} G_{s_i,r} E_{r_i} G_{r,d}}{\sigma^2 (E_{s_i} G_{s_i,r} + E_{r_i} G_{r,d} + \sigma^2)}\right) + \mu(E_r^* - \sum_{s_i \in L_s} E_{r_i})$. Following the K.K.T condition, we take the differentiation of the lagrangian with respect to E_{r_i} and we obtain the following equation

$$\mu\sigma^2(\sigma^2 + E_{s_i} G_{s_i,r} + E_{r_i} G_{r,d})^2 + \mu E_{s_i} G_{s_i,r} E_{r_i}^2 G_{r,d}^2 + \\ \mu(\sigma^2 + E_{s_i} G_{s_i,r}) E_{s_i} G_{s_i,r} E_{r_i} G_{r,d} - (\sigma^2 + E_{s_i} G_{s_i,r}) E_{s_i} G_{s_i,r} G_{r,d} = 0. \tag{7}$$

After simplifying, resultant E_{r_i} is $-\frac{E_{s_i} G_{s_i,r} + 2\sigma^2}{2G_{r,d}} + \frac{1}{2G_{r,d}} \sqrt{E_{s_i}^2 G_{s_i,r}^2 + \frac{4E_{s_i} G_{s_i,r} G_{r,d}}{\mu}}$. Substituting $E_{r_i}, s_i \in L_s$ back into the equation $\sum_{s_i \in L_s} E_{r_i} - E_r^* = 0$, we obtain the upper bound of μ which is $\frac{\sum_{s_i \in L_s} \sqrt{4E_{s_i} G_{s_i,r} G_{r,d}}}{2G_{r,d} E_r^* + \sum_{s_i \in L_s} (2\sigma^2 + E_{s_i} G_{s_i,r})}$. As the lower bound of μ is 0, we apply a bisection search between these two bounds in order to obtain optimal μ. Replacing optimal μ in E_{r_i}, finally we obtain optimal E_{r_i}, i.e., $E_{r_i}^*, s_i \in L_s$.

D. Performance Evaluation

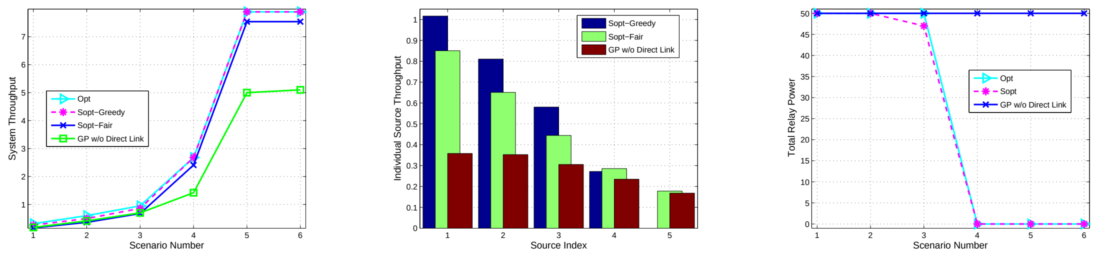
In this subsection, we will evaluate the performance of our proposed solutions. Section II-D1 is for the methodology we have adopted to evaluate the performance and the following one presents the results while comparing with the approach proposed in [4] and [5].

1) Simulation Methodology: We presume a simple network where there are 5 source nodes, one relay and one destination node. Maximum power of individual source is $E_s^{max} = 30$ mW and that of relay node is $E_r^{max} = 50$ mW, total available power in the system is $E^{max} = 120$ mW. Noise variance σ^2 has been set as 1. Channel between two nodes suffers from the shadowing and Rayleigh fading effects. We take the same channel model and the similar values of its parameters as mentioned in [6]. Moreover, we assume, each channel has a unit capacity. One of the major assumptions of the works [4], [5] is channel condition between the source and destination is always worse than that between the source and relay. However, that not necessarily happens in practice and for the counter scenario, their model fails to provide optimal solution. In order to fix the model up, we have considered the SNR due to the direct link in our formulation and the resultant solution is optimal which is able to give better performance even when the direct link's channel is better than that of relayed link. In order to evaluate the performance of our solution, we have placed all nodes in the following coordinates.

- Destination: (0,12).
- Relay: (0,6).
- Sources: X-coordinates are fixed at $\{-1, -2, -3, -4, -5\}$, Y-coordinates are varied from 2 to 14 with the interval of 2 except 12.

There are total 6 different positions of the sources we have experimented. In the evaluation part, we have denoted each position by Scenario Number. All the results we have presented here is the average of 100 simulation runs.

2) Simulation Results: Figure 2(a) presents the total average system throughput with respect to 6 different positions of the source nodes. Notice that, positions of the relay and destination are fixed, we are varying 5 sources' position towards the destination. As the sources move to the destination, resultant channel gain becomes better for them, hence gradually their throughput get improved. For the 5th and 6th scenarios, absolute distance between the sources and destination is very close, however they are on the different sides of the destination. Since their absolute



(a) System throughput.

(b) Individual source throughput.

(c) Total relay power.

Fig. 2. Comparison between our centralized solutions and others.

distances are very close, resultant throughput are very close for these two scenarios. Now if we intuitively compare all approaches, for scenarios 1 and 2, direct link's channel condition is worse than the relayed link, resultant outcome proposed by [4], [5] does not deviate much from our optimal solution. For scenario 3, direct link's channel quality is very close to the relayed link and from this scenario, procedure without considering the direct link SNR starts to differ from our optimal approach. And, for the scenarios 4, 5 and 6, relayed link's channel is worse than that of direct link. For these scenarios, Figure 2(c) shows that allocated power for the relay is 0 and total allowable power is distributed among the sources considering their channel gain towards the destination. However, the technique without considering the direct link SNR always assigns full power to the relay no matter the relayed link is worse or better than the direct link. Since our suboptimal approach for allocating power to the source and relay is based on somewhat weighted averaging of all gains, for scenario 3, relayed power by this approach is little less than the optimal one. For rest of the other scenarios, suboptimal approach confers to the optimal one. We have noticed that, once we obtain total allowable power for the sources, we can distribute this power among the sources by 2 techniques, i.e., greedy and fair algorithms. From Figure 2(a), greedy one has very close performance to the optimal one. Greedy algorithm gives privilege to the sources with better channel condition and makes starvation for others. For being fair to the sources with worse channel condition, resultant system throughput by the fair algorithm deteriorates compared to the other one. Even for scenarios 1, 2, and 3, achieved throughput by this algorithm is worse compared to the GP solution without considering the direct link SNR, however for the other scenarios, it outperforms.

Figure 2(a) has sources at different distance from the relay as well as from the destination. In order to have detailed performance comparison of these four approaches, for input index 4, we have shown each individual source's throughput contribution towards the performance of the system in Figure 2(b). In the X-axis, we put source node index and in the Y-axis, corresponding node's throughput contribution has been projected. As discussed in the previous paragraph, the relay power is 0 for this position. Therefore, the technique without considering direct link SNR has worse performance except some fluctuation comparing with the fair one. Greedy algorithm first assigns full allowable power to the source with the best channel, and this process goes on for all sources with better channel quality until total allowable power runs out. Because of this nature of power subdivision, source 5 obtains 0 power since its channel condition is the worst compared to the rest others. Fair algorithm assigns some power to the sources with worse channel over the time and hence, those sources contribute some throughput towards overall performance. Because of giving some privilege to this type of sources, the sources with the best channel quality obtain less amount of power compared to that given by the greedy one. Therefore, with the fair algorithm, the source node with the best channel has worse performance compared to the greedy one.

References

[1] M. Chiang, "Geometric programming for communication systems," *Commun. Inf. The.*, vol. 2, no. 1/2, pp. 1–154, Jul. 2005. [Online]. Available: http://dx.doi.org/10.1516/0100000005

[2] S. Boyd, S.-J. Kim, L. Vandenberghe, and A. Hassibi, "A tutorial on geometric programming," *Optimization and Engineering*, vol. 8, no. 1, pp. 67–127, 2007.

[3] S. Borst, "User-level performance of channel-aware scheduling algorithms in wireless data networks," *IEEE/ACM Trans. Netw.*, vol. 13, pp. 636–647, June 2005.

[4] K. Phan, T. Le-Ngoc, S. Vorobyov, and C. Tellambura, "Power allocation in wireless multi-user relay networks," *IEEE Trans. Wirel. Commun.*, vol. 8, no. 5, pp. 2535–2545, 2009.

[5] W. H. Fang, M. J. Deng, and Y. T. Chen, "Joint source and relay power allocation in amplify-and-forward relay networks: a unified geometric programming framework," *IET Commun.*, vol. 5, no. 16, pp. 2301–2309, 2011.

[6] R. Ruby and V. Leung, “Uplink scheduling solution for enhancing throughput and fairness in relayed LTE networks,” *Accepted for IET Commun.*, 2013.