Article publicat / *Published paper*:

Rubio, J., Muñoz, O., Pascual Iserte, A., Vidal, J. Stochastic resource allocation with a backhaul constraint for the uplink. A: IEEE International Symposium on Personal, Indoor and Mobile Radio Communications. "Proceedings PIMRC 2016". Valencia: Institute of Electrical and Electronics Engineers (IEEE), 2016, p. 1-6. DOI 10.1109/PIMRC.2016.7794672

Stochastic Resource Allocation with a Backhaul Constraint for the Uplink

Javier Rubio, Olga Muñoz-Medina, Antonio Pascual-Iserte, Josep Vidal
Dept. Signal Theory and Communications - Universitat Politècnica de Catalunya (UPC), Barcelona, Spain
Emails:{javier.rubio.lopez, olga.munoz, antonio.pascual, josep.vidal}@upc.edu

***Abstract*—We propose a novel stochastic radio resource allocation strategy for the uplink that achieves long-term fairness in terms of similar bitrates considering backhaul and air-interface capacity limitations. We focus on a single cell scenario based on WCDMA technology. We propose to use a maximin criteria to introduce fairness among the different users' throughputs. An stochastic approximation is implemented to obtain an online algorithm where the Lagrange multipliers are estimated at each scheduling period. Our results show that the proposed scheme achieves higher fairness among the users and, in some cases, a higher sum-rate compared with the well-known proportional fair scheduler.**

I. Introduction

We consider in this paper an uplink (UL) radio resource allocation strategy for a system with limited backhaul capacity. Although backhaul availability has been taken for granted in conventional systems, backhaul is, in general, a limited resource. This is the case, for instance, in rural deployments, and in particular of the deployment planned in the European project TUCAN3G (http://www.ict-tucan3g.eu). This project aims to provide mobile telephony and data services in isolated rural areas of developing countries (in particular, in rural remote locations in Perú). This is achieved with an access network of 3G femtocells empowered by solar panels of limited size and connected to the core network through a limited capacity WiFi-LD backhaul. Such limited WiFi-LD connections already exist and are used basically to provide remote health services. The final *social* objective of the project is to contribute to the economical development of such areas through the provision of communication services to the general users in addition to the current limited health services.

The backhaul capacity limitation can be introduced in the resource allocation problem by imposing a maximum instantaneous aggregate traffic rate constraint [1], [2], [3]. However, in real deployments, the backhaul capacity can only be measured in average terms. In addition, it is not clear whether limiting the instantaneous sum-rate at the air interface for each scheduling period will hamper the performance of the system in terms of the achievable long-term rates. If the radio channel is time variant, it seems less limiting to use high data rates in the access network whenever the channel conditions allow (even using greater instantaneous values than the average constraint imposed by the backhaul) provided that the average backhaul rate constraint is met when averaging the traffic served in several scheduling periods. Note that the backhaul constraint in terms of average traffic is suitable if we assume that queues are implemented between the access network and the backhaul.

In this paper we propose a long term fairness scheduler that considers a long-term backhaul constraint. When there is no reason for treating flexible service rate users differently, the maximin criterion is a meaningful scheduling approach [4]. This approach maximizes, at each scheduling period, the minimum of the throughputs of the users. Essentially our goal is to provide a balanced long-term rate to a set of users. In addition, the scheduler will take advantage in an opportunistic way of the instantaneous good wireless channel conditions. The authors of this paper have a journal article considering a similar resource allocation strategy but for the downlink setup [5]. Although we resort to stochastic optimization tools as we did in [5] the problem now is different as in the UL we need to deal with individual constraints per user (for power and also for codes assigned) instead of a global constraint for the BS. Additionally, in [5] the BS is powered with solar panels, which is not the case here.

The rest of this paper is organized as follows. In Section II we present the system model. Section III presents the overall problem formulation for the resource allocation strategy. Section IV presents some numerical results and, finally, conclusions are drawn in Section V.

II. System Model

Let us consider an UL system composed of a single BS providing service to several users. The system is based on WCDMA technology and two different types of users coexist: voice users and data users. Let us denote the set of voice and data users by $\mathcal{K}_V$ and $\mathcal{K}_D$, respectively. We assume that voice users request a fixed service rate whereas data users request a flexible service rate.

Users in WCDMA are multiplexed using spreading codes [6]. Let us assume that the network operator has already reserved a set of codes for the voice users and the remaining codes are to be allocated among the data users. Thus, the amount of available codes in each set is known and fixed at the BS. We also consider that there is a maximum backhaul rate constraint for the uplink connections in terms of the average throughput.

Let us collect all the channel gains, h_k that includes the antenna gains, the path loss, and the fading, in $\mathbf{h} = \{h_k, \forall k \in \mathcal{K}_V \cup \mathcal{K}_D\}$. Generally, the wireless channels depend on the specific scheduling period, $\mathbf{h}(t)$, as they vary over time but for simplicity in the notation, we will just refer to them as $\mathbf{h}$

The research leading to these results has received funding from the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications NEWCOM# (Grant agreement no. 318306) and the project TUCAN3G (Grant agreement no. ICT-2011-601102), from the Spanish Ministry of Economy and Competitiveness (Ministerio de Economía y Competitividad) through the project TEC2011-29006-C03-02 (GRE3N-LINK-MAC), project TEC2013-41315-R (DISNET), and FPI grant BES-2012-052850, and from the Catalan Government (AGAUR) through the grant 2014 SGR 60.

$$r_k(\mathbf{h}) \leq n_k(\mathbf{h}) \frac{W}{M_D} \log_2 \left(1 + \frac{M_D p_k(\mathbf{h}) h_k}{n_k \left(\sum_{\ell \in \mathcal{K}_D} p_\ell(\mathbf{h}) h_\ell - \frac{p_k(\mathbf{h}) h_k}{n_k(\mathbf{h})} + \sum_{m \in \mathcal{K}_V} \check{p}_m(\mathbf{h}) + \sigma^2 \right)} \right), \quad (2)$$

throughout the paper. Let $\check{p}_j(\mathbf{h})$ and $p_k(\mathbf{h})$ be the instantaneous powers corresponding to the transmission toward the j-th and k-th voice and data user, respectively and $n_k(\mathbf{h})$ be the number of codes assigned to the k-th data user.

The set of voice users request a fixed data rate and we assume that just one WCDMA code is assigned to them. This is translated into a minimum signal to interference and noise ratio (SINR) requirement, Γ_k, at the output of the de-spreader at the BS as follows:

$$\frac{M_V \check{p}_k(\mathbf{h}) h_k}{\sum_{\ell \in \mathcal{K}_D} p_\ell(\mathbf{h}) h_\ell + \sum_{\substack{m \in \mathcal{K}_V \\ m \neq k}} \check{p}_m(\mathbf{h}) + \sigma^2} \geq \Gamma_k, \, \forall k \in \mathcal{K}_V, \quad (1)$$

where M_V is the spreading factor for voice codes and σ^2 is the noise power. It is important to emphasize that the target SINR has to be achieved with equality (any other solution that fulfills the SNR constraint with strict inequality implies a power spending higher than necessary and, consequently, higher levels of interference). Note also that the orthogonality factor among codes usually considered in the DL of a WCDMA system [6] does not appear in the previous expression since in UL we assume that users are not synchronized in time.

On the other hand, the set of data users request a flexible service rate. The instantaneous throughput in the wireless access channel achieved during one particular scheduling period by the k-th user, $r_k(\mathbf{h})$, is upper bounded by the maximum achievable rate that the access network is able to provide, which is given in (2), where M_D is the spreading factor for data codes and W is the chip rate. Throughout the paper will approximate the denominator of (2) by

$$\sum_{\ell \in \mathcal{K}_D} p_\ell(\mathbf{h}) h_\ell - \frac{p_k(\mathbf{h}) h_k}{n_k(\mathbf{h})} + \sum_{m \in \mathcal{K}_V} \check{p}_m(\mathbf{h}) + \sigma^2 \approx$$
$$\sum_{\ell \in \mathcal{K}_D} p_\ell(\mathbf{h}) h_\ell + \sum_{m \in \mathcal{K}_V} \check{p}_m(\mathbf{h}) + \sigma^2, \quad (3)$$

which is a reasonable assumption if the number of users is relatively high, and, in any case, when using the approximation, the obtained rate corresponds to a lower bound of the achievable rate.

The number of data codes assigned to data users has to fulfill the following condition:

$$n_k(\mathbf{h}) \leq N_{\max}^{(k)}, \quad \forall k \in \mathcal{K}_D. \quad (4)$$

Notice that each user has an independent constraint in terms of maximum number of data codes to be used denoted by $N_max^(k)$. This is so, as each user is allocated a different scrambling code. By following the approximation in (3), the right-hand side in (2) is an increasing function of $n_k(\mathbf{h})$. This implies that, it is optimum to use all data codes available for all data users:

$$n_k^\star(\mathbf{h}) = N_{\max}^{(k)}, \quad \forall k \in \mathcal{K}_D. \quad (5)$$

Given the previous result, only the powers and the data rates for the voice and data users remains to be allocated.

III. Problem Formulation and Resolution

Let us introduce the following set of definitions: $\mathbf{r} \triangleq \{r_k(\mathbf{h}), \, \forall k \in \mathcal{K}_D\}$, $\check{\mathbf{p}} \triangleq \{\check{p}_k(\mathbf{h}), \, \forall k \in \mathcal{K}_V\}$, and $\mathbf{p} \triangleq \{p_k(\mathbf{h}), \, \forall k \in \mathcal{K}_D\}$. We formulate an optimization problem for the resource allocation strategy with backhaul constraints to be executed at the beginning of each particular scheduling period, which involves finding the optimum resource allocation variables, $\mathbf{r}$, $\check{\mathbf{p}}$, and $\mathbf{p}$ that maximize the minimum of the expected throughputs:

$$\underset{\mathbf{r}, \check{\mathbf{p}}, \mathbf{p}}{\text{maximize}} \quad \min_{k \in \mathcal{K}_D} \; \mathbb{E}_{\mathbf{h}}[r_k(\mathbf{h})] \quad (6)$$
$$\begin{aligned} \text{subject to} \quad & C1 : \mathbb{E}_{\mathbf{h}}[r_k(\mathbf{h})] \leq \frac{R_{BH} - \check{R}_{BH}(|\mathcal{K}_V|)}{\xi |\mathcal{K}_D|}, \; \forall k \in \mathcal{K}_D \\ & C2 : p_k(\mathbf{h}) \leq P_T^{(k)}, \quad \forall k \in \mathcal{K}_V \cup \mathcal{K}_D \\ & C3 : r_k(\mathbf{h}) \geq 0, \, p_k(\mathbf{h}) \geq 0, \quad \forall k \in \mathcal{K}_V \\ & C4 : \check{p}_j(\mathbf{h}) \geq 0, \quad \forall j \in \mathcal{K}_V \\ & C5 : (1) \\ & C6 : (2), \end{aligned}$$

where ξ, ($\xi > 1$), is an overhead considered for the data transmissions to be sent through the backhaul, $\check{R}_{BH}(|\mathcal{K}_V|)$ is the backhaul capacity used by the voice users[1], being $|\mathcal{K}_V|$ the number of voice users, R_{BH} is the overall backhaul capacity, Γ is the target SINR for the voice users, and $P_T^{(k)}$ is the maximum radiated power for the uplink connections.

It is important to realize that problem (6) may not be feasible due to constraint $C5$ as it may happen that there could not be enough power to satisfy all the target SINR constraints simultaneously. Constraint $C1$ states that the average throughput that a user is experiencing in the access network should not exceed the maximum backhaul rate assigned to such user (every user has been already assigned a portion of the backhaul, as commented before). If we would like to optimize the amount of backhaul assigned to each particular user, constraint $C1$ could be rewritten as $\sum_{k \in \mathcal{K}_D} \mathbb{E}_{\mathbf{h}}[r_k(\mathbf{h})] \leq \frac{R_{BH} - \check{R}_{BH}(|\mathcal{K}_V|)}{\xi}$. In any case, notice that the instantaneous rates allocated to one user in the access network can be higher in some scheduling periods than the maximum backhaul per-user rate $\left(\frac{R_{BH} - \check{R}_{BH}(|\mathcal{K}_V|)}{\xi |\mathcal{K}_D|}\right)$ thanks to the fact that queues are considered at the output of the access network, i.e., the queues are considered before the traffic is injected into the backhaul. The average rate constraint $C1$ assures that the queues will be stable.

It is easy to realize that the problem is separable into voice and data users without loss of optimality as voice users do not affect the objective function and each user has its own power budget constraint. For this reason, we start by analyzing the voice users.

[1] The capacity required for a set of voice calls to be sent through the backhaul generally depends on the current number of voice users being served. In real deployments, voice users can be jointly encoded and, thus, the overall overhead for voice users may be reduced as the number of voice users increases. Anyway, in the problem formulation and the for the sake of generality, we just use the notation $\check{R}_{BH}(|\mathcal{K}_V|)$.

A. Resource Allocation for the Voice Users

Let us consider that all voice users request the same bit rate, i.e., they request the same SINR constraint, $\Gamma_k = \Gamma$, $\forall k \in \mathcal{K}_V$. Let us define the following variable that takes into account the noise in addition to the received power corresponding to the data connections:

$$\sigma_{int}^2 = \sum_{\ell \in \mathcal{K}_D} p_\ell(\mathbf{h}) h_\ell + \sigma^2. \tag{7}$$

According to this, the set of equations presented in (1) can be written in matrix form as follows (each row corresponds to each of the voice users that are assumed to be numbered with the following order: $k = 1, 2, ..., |\mathcal{K}_V|$):

$$\begin{bmatrix} M_V & -\Gamma & -\Gamma & \dots & -\Gamma \\ -\Gamma & M_V & -\Gamma & \dots & -\Gamma \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\Gamma & -\Gamma & -\Gamma & \dots & M_V \end{bmatrix} \times \begin{bmatrix} \check{p}_1(\mathbf{h}) h_1 \\ \check{p}_2(\mathbf{h}) h_2 \\ \vdots \\ \check{p}_{|\mathcal{K}_V|}(\mathbf{h}) h_{|\mathcal{K}_V|} \end{bmatrix} = \sigma_{int}^2 \Gamma \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Notice that all the previous equations are completely symmetric with respect to users. This implies that the power allocated to each user is inversely proportional to the user channels, i.e., the powers received at the BS from all voice users must be equal:

$$\check{p}_k(\mathbf{h}) = \frac{\alpha}{h_k}, \quad \forall k \in \mathcal{K}_V, \tag{8}$$

where

$$\alpha = \frac{\sigma_{int}^2 \Gamma}{M_V - \Gamma(|\mathcal{K}_V| - 1)}. \tag{9}$$

In the UL scenario, the transmit power constraints are individual, i.e., on a per-user basis. That means that if $\check{p}_k(\mathbf{h}) = \frac{\alpha}{h_k} > P_T^{(k)}$, then the SINR constraints cannot be fulfilled and some users should be dropped off from the system. Notice that if we want to assure that all voice users achieve its minimum SINR, we have to impose a constraint on the interference that the data users generate to the voice users as follows:

$$\begin{aligned} \alpha &= \frac{\sigma_{int}^2 \Gamma}{M_V - \Gamma(|\mathcal{K}_V| - 1)} && (10) \\ &= \left(\sum_{\ell \in \mathcal{K}_D} p_\ell(\mathbf{h}) h_\ell + \sigma^2 \right) \frac{\Gamma}{M_V - \Gamma(|\mathcal{K}_V| - 1)} && (11) \\ &\leq \min_{k \in \mathcal{K}_V} P_T^{(k)} h_k, && (12) \end{aligned}$$

or equivalently

$$\sum_{\ell \in \mathcal{K}_D} p_\ell(\mathbf{h}) h_\ell \leq \left(\min_{k \in \mathcal{K}_V} P_T^{(k)} h_k \right) \frac{M_V - \Gamma(|\mathcal{K}_V| - 1)}{\Gamma} - \sigma^2. \tag{13}$$

Note that (13) is a constraint on the maximum power radiated by all data users simultaneously, i.e., sum-power constraint. As a result, (13) must be added in the resource allocation for data users if we want to assure that all voice users receive the service they demand.

B. Resource Allocation for the Data Users

Now, we can proceed to obtain the optimum power allocation for the data users. Before presenting the optimization problem, let us formulate the achievable rate in terms of the previous found results. The new expression is shown in (15) (in the next page), where c_V is defined as

$$c_V = 1 + \frac{\Gamma |\mathcal{K}_V|}{M_V - \Gamma(|\mathcal{K}_V| - 1)}. \tag{16}$$

Given the previous definitions, we can formulate the resource allocation strategy for the data users for the UL connections as

$$\begin{aligned} \underset{\mathbf{r}, \mathbf{p}, s}{\text{maximize}} \quad & U(s) && (17) \\ \text{subject to} \quad & C1: s \leq \mathbb{E}_{\mathbf{h}}[r_k(\mathbf{h})], \quad \forall k \in \mathcal{K}_D \\ & C2: \mathbb{E}_{\mathbf{h}}[r_k(\mathbf{h})] \leq \frac{R_{BH} - \check{R}_{BH}(|\mathcal{K}_V|)}{\xi |\mathcal{K}_D|}, \ \forall k \in \mathcal{K}_D \\ & C3: \sum_{\ell \in \mathcal{K}_D} p_\ell(\mathbf{h}) h_\ell \leq \left(\min_{k \in \mathcal{K}_V} P_T^{(k)} h_k \right) \frac{|\mathcal{K}_V|}{c_V - 1} - \sigma^2 \\ & C4: p_k \leq P_T^{(k)}, \quad \forall k \in \mathcal{K}_D \\ & C5: (15), \end{aligned}$$

where we have introduced a general differentiable monotonically increasing cost function $U(\cdot)$ (e.g., the logarithm) in order to smooth convergence issues when the objective is linear in the optimization variable (see [7]). Note that the introduction of such function does not modify the optimal value of the optimization variables (i.e., the solution is the same).

Notice also that the previous optimization problem is time-coupled and it requires the future channel realizations due to the expectation operator appearing in $C1$ and $C2$. In order to deal with such difficult problem involving expectations, we propose to use a stochastic approximation that has been proposed in the literature [7], [8]. In this approach, the constraints involving expectations are dualized, and their Lagrange multipliers are estimated stochastically at each period.

Let us start by dualizing constraint $C1$. Let $\boldsymbol{\lambda} \triangleq \{\lambda_k, \forall k \in \mathcal{K}_D\}$ be the vector of Lagrange multipliers associated to $C1$. The partial Lagrangian is given by $\mathcal{L}_{C1}(s, \boldsymbol{\lambda}) = -U(s) + \sum_{k \in \mathcal{K}_D} \lambda_k (s - \mathbb{E}_{\mathbf{h}}[r_k(\mathbf{h})])$. In order to find the optimum s we have to perform the following minimization:

$$\underset{0 \leq s \leq \frac{R_{BH} - \check{R}_{BH}(|\mathcal{K}_V|)}{\xi |\mathcal{K}_D|}}{\text{minimize}} \quad \mathcal{L}_{C1}(s, \boldsymbol{\lambda}). \tag{18}$$

Notice that we have introduced an additional constraint over s. As it is clear from the formulation, this constraint does not affect the optimum solution, but it will help in the numerical search of the optimum value of the new slack variable s. Given that, setting the gradient to zero, $\nabla_s \mathcal{L}_{C1}(s, \boldsymbol{\lambda}) = 0$ and solving yields:

$$s^\star(\boldsymbol{\lambda}) = \left((\dot{U})^{-1} \left(\sum_{k \in \mathcal{K}_D} \lambda_k \right) \right)_0^{\frac{R_{BH} - \check{R}_{BH}(|\mathcal{K}_V|)}{\xi |\mathcal{K}_D|}}, \tag{19}$$

where $\dot{U}(\cdot)$ is the derivative of $U(\cdot)$ and $(\dot{U})^{-1}(\cdot)$ is the inverse function of $\dot{U}(\cdot)$. Once we know the optimum $s^\star$, the problem (17) is updated as follows (where we have skipped in the objective function the term that does not depend on the optimization variables remaining in the optimization problem):

$$\begin{aligned} \underset{\mathbf{r}, \mathbf{p}, \mathbf{n}}{\text{maximize}} \quad & \sum_{k \in \mathcal{K}_D} \lambda_k \mathbb{E}_{\mathbf{h}}[r_k(\mathbf{h})] && (20) \\ \text{subject to} \quad & C2, \dots, C5 \ \text{of problem (17).} \end{aligned}$$

$$r_k(\mathbf{h}) \leq \frac{N_{\max}^{(k)} W}{M_D} \log_2 \left(1 + \frac{M_D p_k(\mathbf{h}) h_k}{N_{\max}^{(k)} \left(\sum_{\ell \in \mathcal{K}_D} p_\ell(\mathbf{h}) h_\ell - \frac{p_k(\mathbf{h}) h_k}{n_k(\mathbf{h})} + \frac{\sigma_{int}^2 \Gamma |\mathcal{K}_V|}{M_V - \Gamma(|\mathcal{K}_V|-1)} + \sigma^2 \right)} \right) \tag{14}$$

$$= \frac{N_{\max}^{(k)} W}{M_D} \log_2 \left(1 + \frac{M_D p_k(\mathbf{h}) h_k}{N_{\max}^{(k)} c_V \left(\sum_{\ell \in \mathcal{K}_D} p_\ell(\mathbf{h}) h_\ell + \sigma^2 \right)} \right) \tag{15}$$

Now, we proceed to dualize constraint $C2$. Let $\boldsymbol{\mu} \triangleq \{\mu_k, \forall k \in \mathcal{K}_D\}$ be the vector of Lagrange multipliers associated to $C2$. The partial Lagrangian is

$$\begin{aligned}
&\mathcal{L}_{C2}\left(r_k(\mathbf{h}); \boldsymbol{\lambda}, \boldsymbol{\mu}\right) = \\
&\quad - \sum_{k \in \mathcal{K}_D} \lambda_k \mathbb{E}_{\mathbf{h}}[r_k(\mathbf{h})] \\
&\quad + \sum_{k \in \mathcal{K}_D} \mu_k \left(\mathbb{E}_{\mathbf{h}}[r_k(\mathbf{h})] - \frac{R_{BH} - \check{R}_{BH}(|\mathcal{K}_V|)}{\xi |\mathcal{K}_D|} \right) \quad (21) \\
&= - \mathbb{E}_{\mathbf{h}} \left[\sum_{k \in \mathcal{K}_D} (\lambda_k - \mu_k) r_k(\mathbf{h}) \right] \\
&\quad - \sum_{k \in \mathcal{K}_D} \mu_k \left(\frac{R_{BH} - \check{R}_{BH}(|\mathcal{K}_V|)}{\xi |\mathcal{K}_D|} \right). \quad (22)
\end{aligned}$$

For given Lagrange multipliers $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, the optimization problem (20) is equivalently reformulated as (where we have skipped again in the objective function the term that does not depend on the optimization variables remaining in the optimization problem):

$$\begin{aligned}
\underset{\mathbf{r}, \mathbf{p}, \mathbf{n}}{\text{maximize}} \quad & \sum_{k \in \mathcal{K}_D} (\lambda_k - \mu_k) r_k(\mathbf{h}) \quad (23) \\
\text{subject to} \quad & C3, \ldots, C5 \text{ of problem (17).}
\end{aligned}$$

Notice that the expectations are no longer present in the formulation because the remaining constraints $C3 - C5$ are applied to instantaneous resource allocation variables (without expectations) and also because the maximization of the expected value of the objective function with respect to $\mathbf{r}$ and $\mathbf{p}$ in the current scheduling period, in this case, leads to the maximization of the term within the expectation. The problem now resides in the computation of the optimum Lagrange multipliers which requires knowing the statistics of $r_k(\mathbf{h})$. If we solve the dual problem of (23), i.e., $\sup_{\boldsymbol{\lambda} \succeq 0, \boldsymbol{\mu} \succeq 0} \inf \mathcal{L}(\mathbf{r}, \mathbf{p}; \boldsymbol{\lambda}, \boldsymbol{\mu})$, where $\succeq$ means element-wise inequality and $\mathcal{L}(\mathbf{r}, \mathbf{p}; \boldsymbol{\lambda}, \boldsymbol{\mu})$ is the Lagrangian, then the optimum multipliers could be found using a gradient approach as shown in (24) and (25). Note that, it is not possible to compute the value of the Lagrange multipliers in real time, and then solve (23), as they depend on the statistics of $r_k(\mathbf{h})$ that is a function not known a priori (it is the solution of the optimization problem itself). Thus, we propose to follow a stochastic approximation [8] - [7] and eliminate such uncertainty constraint by estimating the multipliers stochastically at each scheduling period (with a noisy instantaneous unbiased estimate of the gradient) as shown in (26) and (27) (note that this philosophy is similar to the instantaneous estimation of the gradient in the LMS algorithm [9]).

The advantages of the stochastic techniques are threefold: *i*) the computational complexity of the stochastic technique is significantly lower than that of their off-line counterparts; *ii*) stochastic approaches can deal with non-stationarity environments; *iii*) the distribution of the involved random variables $\mathbf{h}$ is not required.

It is importante to note that problem (23) is not convex due to the interference term in the rate expression. Given that, let us present the methodology employed to find the optimum power allocation. Firstly, the following constraint can be added without loss of optimality:

$$\sum_{\ell \in \mathcal{K}_D} p_\ell(\mathbf{h}) h_\ell \leq \sum_{\ell \in \mathcal{K}_D} P_T^{(\ell)} h_\ell. \tag{28}$$

With this, we guarantee that either constraint $C3$ of problem (23) or (28) will be active at the optimum. Secondly, we introduce the following slack variable q that is defined as

$$q = \sum_{\ell \in \mathcal{K}_D} p_\ell(\mathbf{h}) h_\ell. \tag{29}$$

Having introduced the previous constraint and variable, we may write problem (23) as follows:

$$\begin{aligned}
\underset{\mathbf{p}, q}{\text{maximize}} \quad & \sum_{k \in \mathcal{K}_D} (\lambda_k(t) - \mu_k(t)) r_k(\mathbf{h}) \quad (30) \\
\text{subject to} \quad & C3, \ldots, C5 \text{ of problem (17).} \\
& C6: q \leq \sum_{\ell \in \mathcal{K}_D} P_T^{(\ell)} h_\ell, \quad \forall k \in \mathcal{K}_D, \\
& C7: q = \sum_{\ell \in \mathcal{K}_D} p_\ell(\mathbf{h}) h_\ell.
\end{aligned}$$

The previous optimization problem (30) is concave with respect to the set of powers $\{p_k(\mathbf{h})\}$ for a fixed value of q. This means that we can always find efficiently the optimum value of $\{p_k(\mathbf{h})\}$ for a certain value of q [10]. Unfortunately, problem (30) is not jointly concave in $\{p_k(\mathbf{h})\}$ and q, so there is not an efficient method to obtain both $\{p_k^\star(\mathbf{h})\}$ and $q^\star$ simultaneously. For this reason, we propose a suboptimum approach to solve problem (30) where we perform the optimization in two stages: we first fix q and obtain $\{p_k^\star(\mathbf{h}; q)\}$, then we change the value of q within its range and solve the problem again. The final solution is the one that provides larger sum-rate for all selected values of q. The intuition behind the algorithm is simple. It is based on a exhaustive search approach for the value of q where the range of q has been quantized into small steps so as to provide an algorithm with finite iterations. So, if we want N iterations and the range of q is $q \in [0, Q_{\max}]$, then the step size is $\Delta = Q_{\max}/N$. The smaller the value of Δ, the better the precision. If fact, if $N \longrightarrow \infty \Longrightarrow \Delta \longrightarrow 0$, then the proposed algorithm provides the optimum solution of problem (30). Let

$$\lambda_k^{(q+1)} = \left(\lambda_k^{(q)} + \epsilon\left(s^\star(\boldsymbol{\lambda}^{(q)}) - \mathbb{E}_{\mathbf{h}}\left[r_k^\star\left(\mathbf{h};\boldsymbol{\lambda}^{(q)},\boldsymbol{\mu}^{(q)}\right)\right]\right)\right)_0^\infty, \quad \forall k \in \mathcal{K}_D, \tag{24}$$

$$\mu_k^{(q+1)} = \left(\mu_k^{(q)} + \epsilon\left(\mathbb{E}_{\mathbf{h}}\left[r_k^\star\left(\mathbf{h};\boldsymbol{\lambda}^{(q)},\boldsymbol{\mu}^{(q)}\right)\right] - \frac{R_{BH} - \check{R}_{BH}(|\mathcal{K}_V|)}{\xi|\mathcal{K}_D|}\right)\right)_0^\infty, \ \forall k \in \mathcal{K}_D, \tag{25}$$

$$\lambda_k(t+1) = \left(\lambda_k(t) + \epsilon\left(s^\star(\boldsymbol{\lambda}(t)) - r_k^\star(\mathbf{h};\boldsymbol{\lambda}(t),\boldsymbol{\mu}(t))\right)\right)_0^\infty, \quad \forall k \in \mathcal{K}_D, \tag{26}$$

$$\mu_k(t+1) = \left(\mu_k(t) + \epsilon\left(r_k^\star(\mathbf{h};\boldsymbol{\lambda}(t),\boldsymbol{\mu}(t)) - \frac{R_{BH} - \check{R}_{BH}(|\mathcal{K}_V|)}{\xi|\mathcal{K}_D|}\right)\right)_0^\infty, \quad \forall k \in \mathcal{K}_D. \tag{27}$$

us define the range of the variable q as

$$Q_{\max} = \min\left\{\left(\min_{k\in\mathcal{K}_V} P_T^{(k)} h_k\right)\frac{|\mathcal{K}_V|}{c_V - 1} - \sigma^2, \sum_{\ell\in\mathcal{K}_D} P_T^{(\ell)} h_\ell\right\}. \tag{31}$$

The algorithm to solve (30) is presented Algorithm 1.

Algorithm 1 Algorithm for Solving Resource Allocation Problem (30)

- set number of iterations N
- define vector $\check{\mathbf{q}} = [\check{q}_1, \ldots, \check{q}_N]$ with $0 \le \check{q}_i \le Q_{\max}$
- compute vector $\mathbf{u} = [u^\star(\check{q}_1), \ldots, u^\star(\check{q}_N)]$ where each component is the solution of the optimization problem (30)
- select $\check{q}^\star = \arg\max_{\check{q}_i} \ u^\star(\check{q}_i)$
- select the powers as the value of $\left\{0 \le p_k(\mathbf{h}) \le P_T^{(k)}\right\}$, $\forall k \in \mathcal{K}_D$ that achieve $u^\star(\check{q}^\star)$

C. Overall Resource Allocation Algorithm for the Uplink

In this subsection, we just present the overall algorithm to solve the resource allocation for the voice and data users based on the approaches presented in previous sections. The stochastic updates are also presented. The algorithm is summarized in Algorithm 2.

Algorithm 2 Algorithm for Solving the Resource Allocation Strategy for the UL Connections

1: define $\tilde{R}_{BH} \triangleq \frac{R_{BH} - \check{R}_{BH}(|\mathcal{K}_V|)}{\xi|\mathcal{K}_D|}$
2: initialize $\lambda_k(t) \ge 0$, $\mu_k(t) \ge 0$, $\forall k \in \mathcal{K}_D$
3: set $n_k^\star(\mathbf{h}) = N_{\max}^{(k)}, \quad \forall k \in \mathcal{K}_D$
4: **Data users**
5: compute $r_k^\star(\mathbf{h};\boldsymbol{\lambda}(t),\boldsymbol{\mu}(t))$ with $p_k^\star(\mathbf{h})$ and $n_k^\star(\mathbf{h})$ from (30) using Algorithm 1
6: update (dualized) primal variable:
7: $s^\star(\boldsymbol{\lambda}(t)) = \left((\dot{U})^{-1}\left(\sum_{k\in\mathcal{K}_D}\lambda_k(t)\right)\right)_0^{\tilde{R}_{BH}}$
8: update stochastic dual variables:
9: $\lambda_k(t+1) = \left(\lambda(t) + \epsilon\left(s^\star(\boldsymbol{\lambda}(t)) - r_k^\star(\mathbf{h};\boldsymbol{\lambda}(t),\boldsymbol{\mu}(t))\right)\right)_0^\infty$
10: $\mu_k(t+1) = \left(\mu(t) + \epsilon\left(r_k^\star(\mathbf{h};\boldsymbol{\lambda}(t),\boldsymbol{\mu}(t)) - \tilde{R}_{BH}\right)\right)_0^\infty$
11: **Voice users**
12: compute $\sigma_{int}^2 = \sum_{\ell\in\mathcal{K}_D} p_\ell^\star(\mathbf{h}) h_\ell + \sigma^2$
13: compute $\check{p}_k^\star(\mathbf{h}) = \frac{\sigma_{int}^2\Gamma}{h_k(M_V - \Gamma(|\mathcal{K}_V|-1))}, \quad \forall k \in \mathcal{K}_V$
14: $t \longleftarrow t+1$ and go to 4

IV. Numerical Evaluation

In this section we evaluate the performance of the proposed strategy. The scenario is composed of 1 BS, and 3 voice users and 6 data users. The number of available codes $N_{\max} = 13$. All the users are mobile with a speed of 3 m/s. The instantaneous channel gain, h_k, incorporates antenna gains, Rayleigh fading with unitary power, and a real path loss of a town in Perú known as San Juan (see details in [11]). The code gain of data codes $M_D = 16$ and the minimum SINR normalized with code gain for voice users is, $\frac{\Gamma}{M_V} = -13.7$ dB which corresponds to a rate of 12.2 Kbps. The noise power is $\sigma^2 = -102$ dBm. The scheduling period for the data users and voice users are 2 ms and 20 ms, respectively. The utility function is $U(\cdot) = \log(\cdot)$. The amount of backhaul capacity required by the 3 voice users is $\check{R}_{BH}(|\mathcal{K}_V|) = 90$ Kbps. The overhead for the data transmissions is $\xi = 1.2$. The step size for the update of the stochastic multipliers is $\epsilon = 10^{-3}$. For a more detailed description of the simulation parameters see [11].

For comparison purposes, we also show the resource allocation of the proportional fair (PF) strategy [8] with an instantaneous per-user backhaul constraint, $r_k(\mathbf{h}) \le \frac{R_{BH} - \check{R}_{BH}(|\mathcal{K}_V|)}{\xi|\mathcal{K}_D|}$, and an instantaneous sum constraint, $\sum_{k\in\mathcal{K}_D} r_k(\mathbf{h}) \le \frac{R_{BH} - \check{R}_{BH}(|\mathcal{K}_V|)}{\xi}$. The effective length of the exponential window in the PF has been set to $T_c = 500$ [8].

Fig. 1 and Fig. 2 show the time evolution of the estimation of the expected rates of the proposed stochastic scheduler and the PF scheduler (computed as $\hat{r}_k(t) = \frac{1}{t}\sum_{\tau=1}^{t} r_k(\tau)$). We also plot $s^\star(\boldsymbol{\lambda}(t))$ and the per-data user backhaul rate. The backhaul capacity considered in Fig. 1 is 6 Mbps and in Fig. 2 is 2 Mbps. Each line of the same color represents a different user. Initially, we assume that the queues at the output of the access network are sufficiently full so that all the bits demanded by the users are served. This makes the initial average rates violate the backhaul capacity constraint for a short period of time (see the initial transient in the figure). This is due to the stochastic approximation of the multipliers but when the average rates converge, they fulfill all the constraints of the original problem. As we can also see from Fig. 1, the limitation of the rates comes from the limited resources available at the access network (power and codes) as the backhaul capacity is not reached. It should be also emphasized that, the proposed stochastic approach provides a solution that introduces more fairness when compared to the PF approach as the average rates

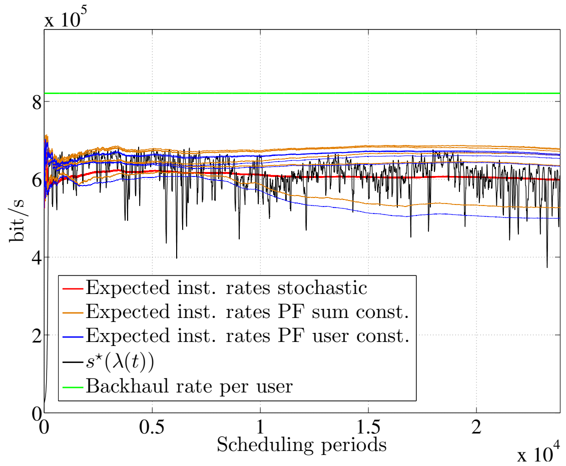

Fig. 1. Average bit rates per data user served in the air interface by different schedulers for a total backhaul capacity of 6 Mbps.

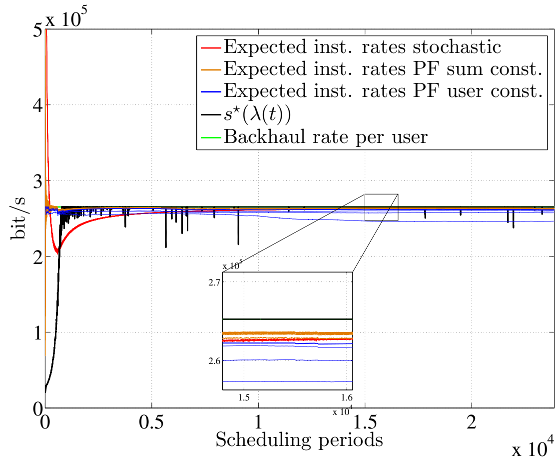

Fig. 2. Average bit rates per data user served in the air interface by different schedulers for a total backhaul capacity of 2 Mbps.

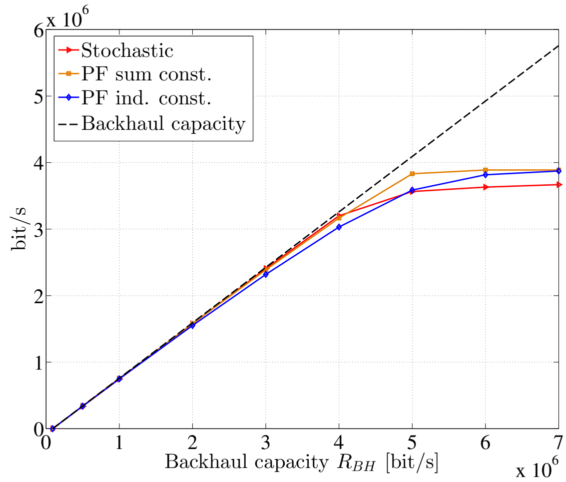

Fig. 3. Sum-rate served in the air interface for data users versus the total backhaul capacity.

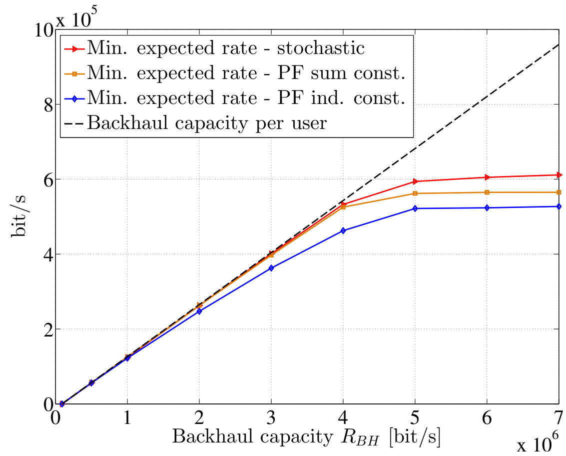

Fig. 4. Rate served in the air interface for the worst case data user versus the total backhaul capacity.

for the different users are quite similar. Considering now Fig. 2, the limitations comes from backhaul as the expected rates in the air interface converge to the maximum per-user backhaul capacity.

Fig. 3 shows the total rate demanded by data users as a function of the backhaul capacity. Note than, the stochastic approach performs slightly worse than the two PF schedulers when the system is limited by the access network and not by the backhaul network but works better than the PF with individual constraint for some backhaul capacities. However, the stochastic scheduler offers a greater fairness in terms of similar bitrates, so the rate for the worst case user is better for the stochastic scheduler than for the other approaches, as shown in Fig. 4.

V. Conclusions

In this paper, we have proposed a resource allocation strategy for the UL based on the maximization of the minimum average data rates. By the use of stochastic optimization tools, we are able to consider a backhaul capacity constraint in terms of the average data rate, and allow the access network to offer higher rates by taking advantage of good instantaneous wireless channel conditions. Simulations results showed that the proposed approach achieves more fairness in terms of similar rates among the users when compared to the traditional PF strategy and, for some backhaul capacities, the sum-rate is higher when compared to that obtained with the PF scheduler.

References

[1] A. Chowdhery, W. Yu, and J. Cioffi, "Cooperative wireless multicell OFDMA network with backhaul capacity constraints," in *IEEE Int'l. Conference on Comm. (ICC))*, Jun. 2011.

[2] L. Zhou and W. Yu, "Uplink multicell processing with limited backhaul via per-base-station successive interference cancellation," *IEEE Journal on Sel. Areas in Comm.*, vol. 31, pp. 1981–1993, Oct. 2013.

[3] Z. Cui and R. Adve, "Joint user association and resource allocation in small cell networks with backhaul constraints," in *IEEE Conf. on Information Sciences and Systems (CISS)*, Mar. 2014.

[4] W. Rhee and J. M. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," in *IEEE Vehicular Technology Conference*, pp. 1085–1089, May 2000.

[5] J. Rubio, O. Muñoz, and A. Pascual-Iserte, "A stochastic approach for resource allocation with backhaul and energy harvesting constraints," *IEEE Trans. on Vehicular Technology*, 2015. Accepted for publication.

[6] A. Goldsmith, *Wireless communications*. Cambridge U. P., 2005.

[7] A. Ribeiro, "Ergodic stochastic optimization algorithms for wireless communication and networking," *IEEE Trans. on Signal Processing*, vol. 58, pp. 6369–6386, Nov. 2010.

[8] X. Wang, G. B. Giannakis, and A. G. Marques, "A unified approach to QoS-guaranteed scheduling for channel-adaptive wireless networks," *Proceedings of IEEE*, vol. 95, pp. 2410–2431, Dec. 2007.

[9] S. Haykin, *Adaptive filter theory*. Englewood Cliffs, Prentice-Hall, 2002.

[10] S. Boyd and L. Vandenbergue, *Convex optimization*. Cambridge, 2004.

[11] "ICT-601102 STP TUCAN3G, UMTS/HSPA network dimensioning," in *Deliverable D41, available online: http://www.ict-tucan3g.eu*, Nov. 2013.