Adaptive Downlink OFDMA Resource Allocation

Ian C. Wong
Cellular Products Group R&D
Freescale Semiconductor, Inc., Austin, Texas USA
Email: ian.wong@ieee.org

Brian L. Evans
Wireless Networking and Communications Group
The University of Texas at Austin, Austin, Texas 78712
Email: bevans@ece.utexas.edu

[1] ***Abstract*—Optimizing OFDMA resource allocation with respect to communication performance requires solving a nonlinear mixed-integer programming problem. As a result, many researchers have fallen back on suboptimal heuristic algorithms. In a recent paper, we demonstrate that ergodic rate maximization is possible using a dual optimization framework that results in a practically optimal solution with complexity that is on the order of the number of subcarriers times the number of users. One of the primary disadvantages of considering ergodic rates is the assumption that the channel distribution information (CDI) is perfectly known at the transmitter. Therefore, this paper proposes an adaptive algorithm based on stochastic approximation methods that do not require knowledge of the CDI. This algorithm converges to the optimal solution with probability one, while for each OFDMA symbol, the complexity is on the order of the number of subcarriers times the number of users. There are no iterations in a given OFDMA symbol time; instead, the "iterations" are actually performed across time (symbols). Simulation results based roughly on a third-generation partnership project, long-term evolution (3GPP-LTE) OFDMA system corroborate our claims.**

I. INTRODUCTION

The instantaneous sum-rate maximization with proportional rate constraints have been studied previously in [1] [2]. The main emphasis of these papers, in terms of formulation, was on an instantaneous rate maximization with instantaneous proportional rate constraints. Furthermore, the solution methods proposed were suboptimal heuristics with complexity of $\mathcal{O}(MK\log(K))$ or higher.

In this paper, we use a dual optimization framework to solve the *ergodic sum-rate* maximization with *proportional ergodic rate* constraints. We show that the proportional rate constraints can actually be imposed by a weighted-sum rate dual, with the weights being the *dual optimal geometric multipliers* themselves that enforce the proportional rate constraints. We compared the performance of our algorithm with the previous algorithm that gives the best performance [2], and show that exploiting the temporal dimension using the ergodic formulation provides huge rate gains versus the previous state-of-the-art.

One main disadvantage of considering ergodic rates is the assumption that the *channel distribution information* (CDI) is perfectly known at the transmitter, and thus the expected values of the rates can be computed. Although methods to estimate the distribution function are available [3], they are typically more suitable for off-line processing rather than the online algorithms that are needed in practical wireless system implementations. Therefore, we also propose an adaptive algorithm based on *stochastic approximation* methods [4] [5] that do not require knowledge of the CDI, and is shown to converge to the optimal solution w.p.1, while requiring only $\mathcal{O}(MK)$ complexity per-symbol *without iterations*, since the iterations are actually done across time.

.

II. PROPORTIONAL RATE MAXIMIZATION

The ergodic rate maximization problem with proportional ergodic rate constraints can be formulated as

$$\begin{aligned}
&\max_{\boldsymbol{p}(\boldsymbol{\gamma})\in\boldsymbol{\mathcal{P}}} \mathbb{E}_{\boldsymbol{\gamma}}\left\{\sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}} R_{m,k}(p_{m,k}\gamma_{m,k})\right\}\\
&\quad\text{s.t. } \mathbb{E}_{\boldsymbol{\gamma}}\left\{\sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}} p_{m,k}\right\} \le \bar{P}\\
&\qquad \mathbb{E}_{\boldsymbol{\gamma}}\left\{\sum_{k\in\mathcal{K}} R_{m,k}(p_{m,k}\gamma_{m,k})\right\} \ge\\
&\qquad\quad \phi_m \mathbb{E}_{\boldsymbol{\gamma}}\left\{\sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}} R_{m,k}(p_{m,k}\gamma_{m,k})\right\}, \forall m\in\mathcal{M}
\end{aligned} \tag{1}$$

where the ϕ_m terms are the proportionality constants for each user m such that $\sum_m \phi_m = 1$. The constants ϕ_m

[1]This research was supported by an equipment grant from Intel.

can be interpreted as the portion of the total ergodic sum rate that should be allocated to each user m. We denote by $\boldsymbol{\phi} = [\phi_1, \ldots, \phi_M]^T$ the vector of proportionality constants. By introducing a dummy optimization variable $R \geq 0$ that represents the ergodic sum rate, we can rewrite the problem as

$$\begin{aligned} \max_{R\in\mathbb{R}^+, \boldsymbol{p}(\boldsymbol{\gamma})\in\boldsymbol{\mathcal{P}}} & R \\ \text{s.t. } & \mathbb{E}_{\boldsymbol{\gamma}}\left\{\sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}} p_{m,k}\right\} \leq \bar{P} \\ & \mathbb{E}_{\boldsymbol{\gamma}}\left\{\sum_{k\in\mathcal{K}} R_{m,k}(p_{m,k}\gamma_{m,k})\right\} \geq \phi_m R \end{aligned} \tag{2}$$

A similar reformulation as in (2) that uses a dummy variable is proposed in [6] to solve for the max-min rate, and in [7] for instantaneous proportional rates.

The Lagrangian of (2) is given by

$$\begin{aligned} L(R, \boldsymbol{p}(\boldsymbol{\gamma}), \lambda, \boldsymbol{\mu}) = R + \lambda\left(\bar{P} - \mathbb{E}_{\boldsymbol{\gamma}}\left\{\sum_{k\in\mathcal{K}}\sum_{m\in\mathcal{M}} p_{m,k}\right\}\right) \\ + \sum_{m\in\mathcal{M}} \mu_m\left(\mathbb{E}_{\boldsymbol{\gamma}}\left\{\sum_{k\in\mathcal{K}} R_{m,k}(p_{m,k}\gamma_{m,k})\right\} - \phi_m R\right) \end{aligned} \tag{3}$$

where $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_M]^T$ is the vector of geometric multipliers that are used to enforce the proportionality constraints. The dual objective can then be written as

$$\begin{aligned} \Theta(\lambda, \boldsymbol{\mu}) &= \max_{R\in\mathbb{R}^+, \boldsymbol{p}(\boldsymbol{\gamma})\in\boldsymbol{\mathcal{P}}} L(R, \boldsymbol{p}(\boldsymbol{\gamma}), \lambda, \boldsymbol{\mu}) \\ &= \lambda\bar{P} + \max_{R\in\mathbb{R}^+, \boldsymbol{p}(\boldsymbol{\gamma})\in\boldsymbol{\mathcal{P}}} \Big[R\left(1 - \boldsymbol{\mu}^T\boldsymbol{\phi}\right) \\ &\qquad - \lambda\mathbb{E}_{\boldsymbol{\gamma}}\left\{\sum_{k\in\mathcal{K}}\sum_{m\in\mathcal{M}} p_{m,k}\right\} \\ &\qquad + \sum_{m\in\mathcal{M}} \mu_m \mathbb{E}_{\boldsymbol{\gamma}}\left\{\sum_{k\in\mathcal{K}} R_{m,k}(p_{m,k}\gamma_{m,k})\right\}\Big] \end{aligned} \tag{4}$$

Focusing on the first term in the maximization $R\left(1 - \boldsymbol{\mu}^T\boldsymbol{\phi}\right)$, we observe that if $1 - \boldsymbol{\mu}^T\boldsymbol{\phi} > 0$, then the optimal solution would be $R^* = \infty$, since R is a free variable. This is clearly an infeasible solution for the ergodic sum rate. Furthermore, if $1 - \boldsymbol{\mu}^T\boldsymbol{\phi} < 0$, then the optimal solution would be $R^* = 0$. In this case, the ergodic sum rate is zero and is uninteresting from an optimization point of view. Thus, we would like to constrain the multipliers to satisfy $\boldsymbol{\mu}^T\boldsymbol{\phi} = 1$, which allows us to remove the dummy variable from consideration since $R\left(1 - \boldsymbol{\mu}^T\boldsymbol{\phi}\right) = 0$. Thus, (4) can be simplified to

$$\begin{aligned} \Theta(\lambda, \boldsymbol{\mu}) = \lambda\bar{P} + \\ K\mathbb{E}_{\boldsymbol{\gamma}_k}\left\{\max_{m\in\mathcal{M}}\left(\mu_m R_{m,k}(\tilde{p}_{m,k}\gamma_{m,k}) - \lambda\tilde{p}_{m,k}\right)\right\} \end{aligned} \tag{5}$$

where $\tilde{p}_{m,k} = [\mu_m/(\lambda\ln 2) - 1/\gamma_{m,k}]^+$. The dual problem can then be written as

$$g^* = \min_{\lambda\geq 0, \boldsymbol{\mu}\in\boldsymbol{\mathcal{U}}} \quad \Theta(\lambda, \boldsymbol{\mu}) \tag{6}$$

where $\boldsymbol{\mathcal{U}} = \left\{\boldsymbol{\mu} \geq \mathbf{0} \,\middle|\, \boldsymbol{\mu}^T\boldsymbol{\phi} = 1\right\}$. A possible method to find the solution to (6) is to use subgradient search [8, Ch. 6.3.1], which is a generalization of gradient-based search methods to possibly non-differentiable functions. From an initial guess λ^0, the subgradient method generates a sequence of dual feasible points according to the iteration

$$\lambda^{i+1} = \left[\lambda^i - s^i g_\lambda^i\right]^+ \tag{7}$$

where g_λ^i denotes the subgradient of $\Theta(\lambda^*(\boldsymbol{\mu}^i), \boldsymbol{\mu}^i)$ with respect to λ, and s^i is a positive scalar step-size. A similar subgradient method can be used to search for the optimal $\boldsymbol{\mu}^*$ that enforces the proportional rate constraints, given by the iterations

$$\boldsymbol{\mu}^{i+1} = \Pi_{\boldsymbol{\mathcal{U}}}\left[\boldsymbol{\mu}^i - s^i\boldsymbol{g}_{\boldsymbol{\mu}}^i\right] \tag{8}$$

where $\boldsymbol{g}_{\boldsymbol{\mu}}^i$ denotes the subgradient of $\Theta(\lambda^*(\boldsymbol{\mu}^i), \boldsymbol{\mu}^i)$ with respect to $\boldsymbol{\mu}$, and $\Pi_{\boldsymbol{\mathcal{U}}}[\cdot]$ denotes projection onto the set $\boldsymbol{\mathcal{U}}$.

The subgradient method is particularly attractive for solving the dual problem, since the inequality constraint evaluated at the optimal power vector for a given λ and $\boldsymbol{\mu}$ is itself the subgradient [8], i.e.

$$g_\lambda^i = \bar{P} - \hat{P}_{tot}^i \tag{9}$$

where

$$\hat{P}_{tot}^i = \mathbb{E}_{\boldsymbol{\gamma}}\left\{\sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}} p_{m,k}^*(\lambda^i, \mu_m^i)\right\} \tag{10}$$

is the average power given $\lambda^i, \boldsymbol{\mu}^i$; and

$$\boldsymbol{g}_{\boldsymbol{\mu}}^i = \bar{\boldsymbol{R}}^i - \boldsymbol{\phi}\bar{R}^i \tag{11}$$

where $\bar{\boldsymbol{R}}^i = \left[\bar{R}_1^i, \ldots, \bar{R}_M^i\right]$ with

$$\bar{R}_m^i = \mathbb{E}_{\boldsymbol{\gamma}}\left\{\sum_{k\in\mathcal{K}} R_{m,k}(p_{m,k}^*(\lambda^i, \mu_m^i)\gamma_{m,k})\right\} \tag{12}$$

as the vector of ergodic rates per user for a given $\boldsymbol{\mu}^i$, and

$$\bar{R}^i = \sum_{m\in\mathcal{M}} \bar{R}^i_m \tag{13}$$

is the ergodic sum rate. The optimal power given the current λ^i and $\boldsymbol{\mu}^i$, $p^*_{m,k}(\lambda^i, \mu^i_m)$, is similarly derived as in [9, Eq. 10] with w_m replaced by μ^i_m. The projection operation can be simply performed by clipping and rescaling the new iterate such that it is non-negative and that it satisfies $\boldsymbol{\mu}_i^T\boldsymbol{\phi} = 1$. Hence, we can write (8) as

$$\boldsymbol{\mu}^{i+1} = \frac{\left[\boldsymbol{\mu}^i - s^i\boldsymbol{g}^i\right]^+}{\boldsymbol{\phi}^T\left[\boldsymbol{\mu}^i - s^i\boldsymbol{g}^i\right]^+} \tag{14}$$

where $[\boldsymbol{x}]^+$ implements $\max(x_i, 0)$ for each element in the vector argument $\boldsymbol{x}$. The convergence properties of (14) for different step-size selection rules have already been studied previously (see e.g. [8, Ch. 6.3.1]). In our numerical experiments, we use the simple diminishing step-size rule

$$s^i = \frac{\beta}{i+\alpha} \tag{15}$$

where α and β are suitably chosen positive constants, which satisfies $s^i \to 0$ (for convergence) and $\sum_{i=0}^{\infty} s^i = \infty$ (for allowing us to go "anywhere").

We can interpret the multiplier vector $\boldsymbol{\mu}^i$ as a vector of priorities for the users, wherein we try to increase the priority of a user while it is still unable to get its allocated "portion of the pie" $\phi_m R^i$. Upon convergence, we arrive at the optimal $\boldsymbol{\mu}^*$ which is the vector of appropriate weights for each user such that the proportionality constraints are met, and its corresponding λ^* that enforces the average power constraint.

III. Adaptive Algorithms for OFDMA Rate Maximization

In the previous section, we assumed the availability of the channel distribution information (CDI) at the transmitter. Although there are methods that allow us to estimate this (e.g. goodness-of-fit tests followed by maximum likelihood parameter estimation [3]), they are typically quite computationally intensive, and are more suitable for offline processing. In our scenario, it is important to be able to perform the resource allocation in real-time, hence online adaptive algorithms are more desirable. In this section, we outline a framework based on *stochastic approximation* to perform adaptive OFDMA resource allocation that allows us to do without the CDI. Note that stochastic approximation methods have been studied in the context of wireless network scheduling for TDMA in [10], and for weighted-sum continuous rate maximization for a downlink OFDMA system [11].

A. Stochastic Approximation Solution to the Dual Problem

The fundamental stochastic approximation iteration we employ is based on the subgradient iterations given in (7) and (8), but performed across time, i.e.

$$\lambda[n+1] = \left[\lambda[n] - \beta_n g_\lambda[n]\right]^+_\epsilon \tag{16}$$

$$\boldsymbol{\mu}[n+1] = \Pi_{\boldsymbol{\mathcal{U}}}\left[\boldsymbol{\mu}[n] - \beta_n \boldsymbol{g}_{\boldsymbol{\mu}}[n]\right] \tag{17}$$

where $[x]^+_\epsilon = \max(x, \epsilon)$ for a small constant $0 < \epsilon \ll 1$ and is used in (16) as a modified projection operator to prevent λ from going to zero (which results in infinite power), and β_n is a real-valued step-size chosen to satisfy

$$\sum_{n=0}^{\infty}\beta_n = \infty, \beta_n \geq 0, \beta_n \to 0 \tag{18}$$

Furthermore, we employ an auxiliary filter to perform *subgradient averaging*

$$\begin{aligned} g_\lambda[n+1] &= (1-\alpha_n)g_\lambda[n] + \alpha_n\hat{g}_\lambda[n] \\ &= g_\lambda[n] + \alpha_n(\hat{g}_\lambda[n] - g_\lambda[n]) \end{aligned} \tag{19}$$

$$\begin{aligned} \boldsymbol{g}_{\boldsymbol{\mu}}[n+1] &= (1-\alpha_n)\boldsymbol{g}_{\boldsymbol{\mu}}[n] + \alpha_n\hat{\boldsymbol{g}}_{\boldsymbol{\mu}}[n] \\ &= \boldsymbol{g}_{\boldsymbol{\mu}}[n] + \alpha_n(\hat{\boldsymbol{g}}_{\boldsymbol{\mu}}[n] - \boldsymbol{g}_{\boldsymbol{\mu}}[n]) \end{aligned} \tag{20}$$

with α_n as a non-negative step-size chosen to satisfy

$$\alpha_n \geq 0, \frac{\beta_n}{\alpha_n} \to 0, \sum_{n=0}^{\infty}(\beta_n^2 + \alpha_n^2) < \infty \tag{21}$$

and where $\hat{g}_\lambda[n]$ and $\hat{\boldsymbol{g}}_{\boldsymbol{\mu}}[n]$ are approximations to the subgradient given the current CNR realization $\boldsymbol{\gamma}[n]$ and the current estimates for the multipliers $\lambda[n]$ and $\boldsymbol{\mu}[n]$. This method that employs averaging of the search directions are called averaged, aggregated, or mixed stochastic gradient or quasigradient methods [12]. Note that the conditions on step sizes α_n and β_n are to ensure w.p.1 convergence (see [13] for proof). A possible choice is given by

$$\beta_n = \frac{b_1}{b_2 + n} \tag{22}$$

$$\alpha_n = \frac{a_1}{a_2 + n^{0.4}} \tag{23}$$

with real constants $a_1 > 0$, $a_2 \geq 0$, $b_1 > 0$, and $b_2 \geq 0$. Although the diminishing step-size ratio requirement allows a simple convergence proof, it causes a degradation of the local rate of convergence. Fortunately, the use of small constant step-sizes to improve tracking capability was also recently shown to converge w.p.1 [4], and this is what we use for the simulations in section IV.

A suitable approximation to the subgradient would be to replace the expectations with the instantaneous (sample) subgradient, which can be computed via a single iteration of the "multi-level waterfilling" with "max-dual user selection" [9] procedure. We repeat this operation here for convenience:

$$\tilde{p}_{m,k}[n] = \left[\frac{\mu_m[n]}{\lambda[n]\ln 2} - \frac{1}{\gamma_{m,k}[n]}\right]^+ \quad (24)$$

$$\begin{aligned} m_k^*[n] = \arg\max_{m\in\mathcal{M}} \{&\mu_m[n]R_{m,k}\left(\tilde{p}_{m,k}[n]\gamma_{m,k}[n]\right) \\ &-\lambda[n]\tilde{p}_{m,k}[n]\} \end{aligned} \quad (25)$$

$$p_{m,k}^*[n] = \begin{cases} \tilde{p}_{m,k}[n], & m = m_k^*[n] \\ 0, & \text{otherwise} \end{cases} \quad (26)$$

where we use $\gamma_{m,k}[n]$ to denote the channel gain for user m and subcarrier k at time n. Observe that in the process of our stochastic subgradient iterations, we also generate the resource allocation procedure for time n given by (24)-(26).

The per-user instantaneous rate is then given as

$$R_m[n] = \sum_{k\in\mathcal{K}} R_{m,k}\left(p_{m,k}^*[n]\gamma_{m,k}[n]\right) \quad (27)$$

with instantaneous total power

$$P[n] = \sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}} p_{m,k}^*[n] \quad (28)$$

The subgradient approximations are then given as

$$\hat{g}_\lambda[n] = \bar{P} - P[n] \quad (29)$$

$$\hat{\boldsymbol{g}}_{\boldsymbol{\mu}}[n] = \boldsymbol{R}[n] - \boldsymbol{\phi}R[n] \quad (30)$$

where $\boldsymbol{R}[n] = [R_1[n], \ldots, R_M[n]]^T$ and $R[n] = \sum_{m\in\mathcal{M}} R_m[n]$. Using (29)-(30) in the subgradient averaging operations (19)-(20) completes our algorithm. Fig. 1 shows the block diagram for the proposed algorithm.

The complexity of this algorithm is significantly lower than our previously proposed algorithms assuming perfect CDI [9], since all that is needed is the multi-level waterfilling and max-dual user selection with $\mathcal{O}(MK)$, followed by $\mathcal{O}(M)$ updates for the rates, power, and multipliers. Hence, we do away completely with the initialization complexity, and have allowed our "iterations" to be performed over time and on the fly. .

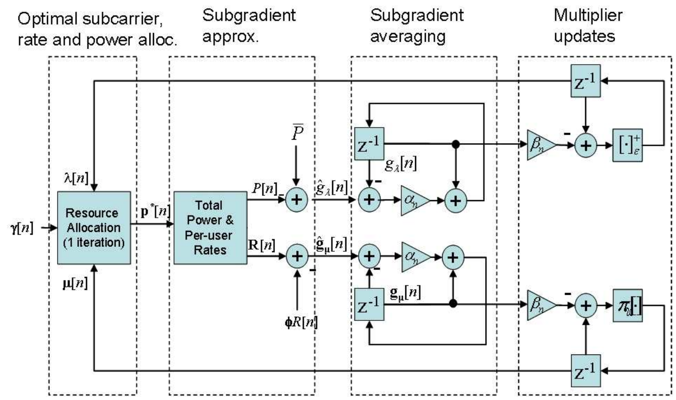

Fig. 1. Block diagram for adaptive OFDMA resource allocation for ergodic sum-rate maximization with ergodic proportional rate constraints.

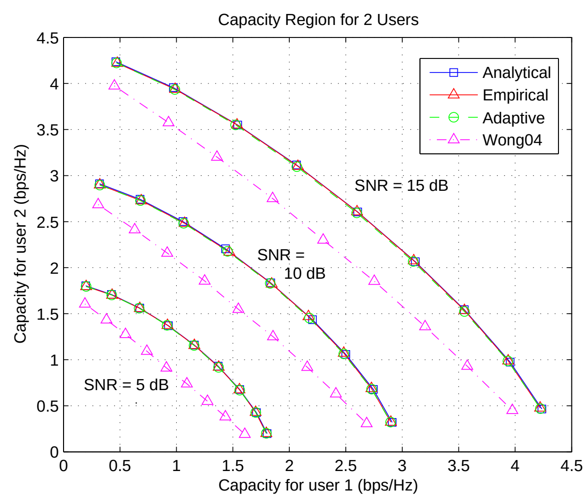

Fig. 2. Two-user capacity region for ergodic sum-rate maximization with proportional rate constraints.

IV. Results and Discussion

We consider an OFDMA system roughly based on a 3GPP-LTE downlink [14], with 128 subcarriers, 76 used subcarriers, 1.25 MHz bandwidth, 1.92 MHz sampling frequency, and a cyclic prefix length of 6 samples. We simulate the frequency-selective Rayleigh fading channel using the ITU-Vehicular A channel model. We generate 10000 IID channel realizations per data point, where for each user's channel realization $\boldsymbol{h}_m$, we generate a complex Gaussian random vector with N_t independent entries, each with variance corresponding to the power delay profile for the corresponding path. Fig. 2 shows the $M = 2$ user capacity region with $\phi_1 = 0.1$ to $\phi_1 = 0.9$ in 0.1 increments and $\phi_2 = 1 - \phi_1$ for the following:

1) *Analytical*: Numerical evaluation of the per-user ergodic rate integral (12)

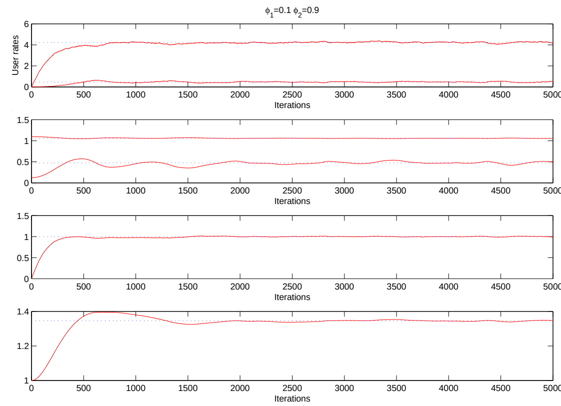

Fig. 3. Evolution across iterations of the exponentially averaged user rates and power, and their corresponding geometric multipliers. Theoretical values solved using a perfect CDI assumption is shown in dotted lines.

2) *Empirical*: Sample average of the per-user rates by using the pre-computed λ^* and $\boldsymbol{\mu}^*$
3) *Adaptive*: Sample average of the per-user rates of the algorithm in Sec. III-A with constant step-size $\alpha_n = \beta_n = 0.005$
4) *Wong04*: Sample average of the per-user rates using the current state-of-the-art algorithm for proportional rate OFDMA resource allocation [2]

Observe that in contrast to the weighted-sum rate capacity regions [9, Fig. 3], the rate points for all the methods are neatly spaced along the boundary of the rate region since we constrain $\bar{R}_1/\bar{R}_2 = \phi_1/\phi_2$, confirming that the algorithms indeed enforce the proportional rate constraints. We also observe that methods 1-3 give essentially identical results, confirming our analysis in the previous sections. On the other hand, using a per-symbol algorithm [2], which is more complex than our algorithms, has significantly poorer performance, because it is suboptimal to start with, and that it is unable to exploit the temporal dimension.

Fig. 3 shows the evolution of the exponentially averaged user rates $\bar{R}_m[n] = (1-\beta_n)\bar{R}_m[n-1] + \beta_n R_m[n]$ and average power $\bar{P}[n] = (1-\beta_n)\bar{P}[n-1] + \beta_n P[n]$, together with the multipliers $\lambda[n]$ and $\boldsymbol{\mu}[n]$ with initializations $\lambda[0] = \bar{P}$, $g_\lambda[0] = 0$, $\boldsymbol{\mu}[0] = \boldsymbol{\phi}/(\boldsymbol{\phi}^T\boldsymbol{\phi})$, and $\boldsymbol{g}_{\boldsymbol{\mu}}[0] = \mathbf{0}$ for an SNR of 15 dB and for proportionality constants $\boldsymbol{\phi} = [0.1, 0.9]^T$ (the results are similar for other $\boldsymbol{\phi}$ values). We can see that the iterates converge to their offline-equivalent optimal values, which are shown by the dotted lines.

V. Conclusion

In this paper, we derived the optimal algorithm for OFDMA resource allocation for ergodic sum-rate maximization subject to ergodic rate proportionality constraints. It is shown that the proportional rates can be enforced by a weighted-sum rate formulation using optimally chosen weights, which are themselves the dual-optimal geometric multipliers. We developed an adaptive algorithm that updates the geometric multipliers over time using a subgradient search and stochastic subgradient averaging. It is based on general *stochastic approximation* principles, which can be shown to converge to the optimal solution w.p.1 [13].

References

[1] Z. Shen, J. Andrews, and B. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2726–2737, Nov. 2005.

[2] I. C. Wong, Z. Shen, B. Evans, and J. Andrews, "A low complexity algorithm for proportional resource allocation in OFDMA systems," in *Proc. IEEE Workshop on Signal Processing Systems*, Oct. 2004, pp. 1–6.

[3] S. Ross, *Simulation*, 3rd ed. Academic Press, 2002.

[4] H. J. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, 2nd ed. Springer, 2003.

[5] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*. Springer, 1997.

[6] W. Rhee and J. M. Cioffi, "Increase in Capacity of Multiuser OFDM System Using Dynamic Subchannel Allocation," in *Proc. IEEE Vehicular Technology Conference*, Tokyo, Japan, May 2000, pp. 1085–1089.

[7] W. Yu, R. Lui, and R. Cendrillon, "Dual optimization methods for multiuser orthogonal frequency division multiplex systems," in *Proc. IEEE Global Telecommunications Conference*, vol. 1, Dec. 2004, pp. 225–229.

[8] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Athena Scientific, 1999.

[9] I. C. Wong and B. L. Evans, "Optimal OFDMA Resource Allocation with Linear Complexity to Maximize Ergodic Rates," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, Mar. 2008.

[10] X. Wang, G. Giannakis, and A. Marques, "A unified approach to QoS-Guaranteed Scheduling for Channel-Adaptive Wireless Networks," *Proc. IEEE*, submitted.

[11] Y. Yu, X. Wang, and G. B. Giannakis, "Channel-adaptive congestion control and OFDMA scheduling for hybrid wireline-wireless networks," *IEEE Trans. Wireless Commun.*, submitted for publication.

[12] B. T. Polyak and A. B. Juditsky, "Acceleration of stochastic approximation by averaging," *SIAM J. Control Optim.*, vol. 30, no. 4, pp. 838–855, 1992.

[13] I. C. Wong and B. L. Evans, *Resource Allocation in Multiuser Multicarrier Wireless Systems*. Springer, 2007.

[14] *3rd Generation Partnership Project, Technical Specification Group Radio Access Network; Physical layer aspects for evolved Universal Terrestrial Radio Access (UTRA)*, 3GPP Std. TR 25.814 v. 7.0.0, 2006.