A proximal minimization based distributed approach to power control in wireless networks: Performance and comparative analysis

Stefano Mutti, Alessandro Falsone, Kostas Margellos, Maria Prandini

Abstract— In this paper, we address power control in a wireless cellular network where multiple mobile users are served by a set of base stations, one per cell. Their reference base station has then to set their transmission power so as to maximize the level of service measured in terms of transmission throughput. Interference due to communication of devices on the same wireless channel is coupling the decisions of all base stations in the cellular network. We propose a distributed algorithm based on proximal minimization that makes the base stations reach consensus to a solution that guarantees an optimal throughput for all mobile devices, by appropriately setting the signal to interference plus noise ratio. The introduced algorithm is compared with a gradient-based distributed algorithm via an extensive simulation study, which reveals the advantage of proximal minimization in the case when the cost function is non-differentiable and the sub-gradient has to be computed.

I. INTRODUCTION

Power control in wireless cellular networks has been a topic of growing interest in the latest decades, possibly due to the pervasive use of wireless communications, [1], [2].

The large growth of the number of wireless devices has prompted the need for suitable resource management strategies so as to allow high quality of transmission while avoiding excessive power consumption, with the latter aspect being particularly critical for those devices using a battery with limited autonomy.

We consider a wireless cellular network where each cell is associated with a base station, and base stations exchange data via wired communication. Mobile users adopt Code Division Multiple Access (CDMA, [3]) to communicate on the same channel with the base station in their cell. If the same communication channel is used in all cells, then, interference occurs among all mobile users in the network, thus causing a decrease of the quality of transmission as measured by the Signal to Interference plus Noise Ratio (SINR). Consequently, the throughput of the device can deteriorate, since, as the SINR decreases, the number of retransmissions increases, with a reduction of the effective data transmission rate.

An adjustment of the transmission power is then needed to guarantee a certain transmission rate and throughput. Due

Stefano Mutti, Alessandro Falsone, and Maria Prandini are with Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy, e-mail: `stefano.mutti@mail.polimi.it`, {`alessandro.falsone, maria.prandini`}`@polimi.it`

Kostas Margellos is with the Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, United Kingdom, e-mail: `kostas.margellos@eng.ox.ac.uk`

to the wireless channel resource sharing, this leads to a network-wide constrained optimization problem where the transmission powers of all devices in the cellular network have to be jointly set so as to maximize the sum of their throughputs while not exceeding the power transmission capabilities of each one of them. The solution is far from being trivial like, e.g., "use the maximum transmission power per device" because this would have the positive effect of increasing the signal component but, at the same time, it would also increase the interference.

Solving the problem in a centralized way can be computationally challenging and also requires the transmission of much information among base stations, possibly overloading the wired communication network. Most of the approaches proposed in the literature are actually based on a distributed scheme and they can be classified into two main categories, i.e., autonomous and non-autonomous, the distinguishing feature being that, while in the former communications occur only between base station and mobile users, in the latter base stations collaborate and exchange information.

The approach described in [4] and further studied in [2], belongs to the first class and aims at minimizing the total transmission power subject to constraints on the SINR via a distributed iterative algorithm. At every iteration, each mobile user sets its transmission power at a certain level, communicates using that power level, and then refines it based on the information on the SINR provided by its reference base station. Autonomous distributed approaches to power control based on games involving non-cooperative users are proposed in [5] and [6]. A further autonomous approach in the literature consists of formulating power control as an open loop global optimization problem where the SINR is not a constraint but has to be optimized via a distributed scheme, [7], [8].

The solution of the power control problem in a distributed non-autonomous fashion involves communication among neighboring base stations and information sharing on their local solutions. The involved additional communication overhead is not an issue though, since, typically, base stations are connected via a wired backbone. In turn, the exchange of information between base stations leads to a faster optimization and to a more robust scheme.

In this work, we propose a novel distributed non-autonomous algorithm for power control in a single channel wireless cellular network. The algorithm is derived from the proximal minimization based approach to distributed optimization recently introduced in [9], [10]. In [11], the same problem is addressed via sub-gradient based distributed
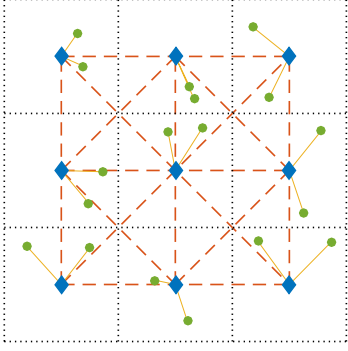
Fig. 1. Example of a cellular network with 9 wire-connected (red dashed lines) base stations (blue diamonds) and a number of mobile devices (green circles). Dotted squares denote coverage area of each base station.

optimization, which is adapted here to our modeling set-up where possibly multiple mobile users are linked to the same base station. A comparative analysis via an extensive simulation study is performed between the two methods.

The rest of the paper is organized as follows. In Section II we formulate the power control problem for a single channel wireless cellular network, extending the problem description in [11] to a set-up where multiple mobile users are possibly communicating with the same base station. In Section III, we propose a distributed non-autonomous algorithm for power control which is derived from the proximal minimization based approach to distributed optimization. A comparative analysis via extensive simulations with the alternative method in [11] derived from sub-gradient based distributed optimization is presented in Section IV. Finally, some concluding remarks are drawn in Section V.

II. POWER CONTROL PROBLEM FORMULATION

We consider a cellular network with m mobile users MU_j, $j = 1, \ldots, m$, served by n base stations BS_i, $i = 1, \ldots, n$, as in Figure 1. A common wireless channel is used for data exchange between mobile users and base stations, so that each mobile user introduces some interference in the data exchange of the others with their reference base station.

The channel is static, and the number of mobile users is assumed constant in the time scale of interest. We suppose that each base station, say BS_i, has an accurate estimate of the amplitude gains (called also channel coefficients) $h_{i,j}$ of the links from all mobile users MU_j, $j = 1, 2, \ldots m$. In particular, if MU_j is in a cell that is not an immediate neighbor of BS_i, then, it causes negligible interference and $h_{i,j}$ will be close to zero. If MU_j is either in cell i or in an immediate neighboring cell, BS_i can estimate the channel coefficient from pilot signals that are sent by the mobile user MU_j.

Each mobile user MU_j is served by a single base station and exchange data with it via wireless communication using a transmission power p_j.

Let us consider base station BS_i. Denote with $\mathcal{J}_i \subseteq \{1, 2, \ldots, m\}$ be the set of indices of the mobile users that communicate with BS_i. We can then compute the SINR of MU_j with $j \in \mathcal{J}_i$ at its reference base station BS_i as follows:

$$\varrho_j(p) = \frac{p_j h_{i,j}^2}{\sigma_i^2 + \sum_{s \neq j} p_s h_{i,s}^2}, \tag{1}$$

where $p = \begin{bmatrix} p_1 & p_2 & \ldots & p_m \end{bmatrix}^\top$ is the transmission power vector, $h_i = \begin{bmatrix} h_{i,1} & h_{i,2} & \ldots & h_{i,m} \end{bmatrix}^\top$ is the vector containing the channel coefficients of all communication uplinks from MU_j, $j = 1, \ldots, m$, to BS_i, and σ_i^2 is the receiver noise variance at the base station BS_i. Shadow fading can also be included by introducing a re-scaling lognormal distributed factor in the channel coefficient along each uplink, [12], [13].

The SINR ϱ_j in (1) is strongly affecting the quality of the transmission, and, in particular, the throughput achieved by the mobile user MU_j, which can be modeled (see e.g. [14]) as proportional to

$$\mathcal{U}_j(p) = \log(1 + \eta \varrho_j(p))$$

where η is a constant that depends on the modulation scheme. The expression above further simplifies to

$$\mathcal{U}_j(p) \approx \log(\eta \varrho_j(p)) \tag{2}$$

in high SINR regime. This implies that some minimum transmission power $p_{\min} > 0$ should be adopted on each link.

From the perspective of base station BS_i, a certain quality of transmission should be guaranteed to all mobile users that BS_i serves, while avoiding a too costly use of power. This can be achieved by choosing p so as to maximize the worst performance index

$$\min_{j \in \mathcal{J}_i} \left\{ \log(\varrho_j(p)) - \mathcal{V}_j(p) \right\}, \tag{3}$$

where $\mathcal{V}_j(\cdot)$ is a convex and differentiable function that represents the power cost for the mobile user MU_j and is then increasing as a function of p_j. The contribution of η to the cost can be neglected since it does not affect the solution.

Note that the above performance index depends on the transmission power of all mobile users in the network through $\varrho_j(p)$ (see (1)) and not only on those that are served by BS_i. This entails that all base stations need to cooperate to guarantee the best minimum quality of the transmission for all users by jointly solving the following constrained optimization problem

$$\max_p \sum_{i=1}^n \min_{j \in \mathcal{J}_i} \left\{ \log(\varrho_j(p)) - \mathcal{V}_j(p) \right\} \tag{4}$$

subject to: $p_{\min} \leq p_j \leq p_{\max}$, $j = 1, \ldots, m$,

where $p_{\max} > p_{\min}$ denotes a maximum transmission power and is assumed to be given.

III. Distributed non-autonomous algorithm based on proximal minimization

We now reformulate the problem so that it fits the framework of the distributed algorithm in [9], [10] for solving convex optimization problems with separable cost and local constraints on a global decision vector.

To this purpose, let us express vector p of the mobile users transmission powers as a function of vector $x = \begin{bmatrix} x_1 & \ldots & x_m \end{bmatrix}^\top$ through the change of variables $p = e^x$, where $e^x = \begin{bmatrix} e^{x_1} & e^{x_2} & \ldots & e^{x_m} \end{bmatrix}^\top$.

We then obtain the following reformulation of (4)

$$\min_{x \in X} \sum_{i=1}^{n} f_i(x) \tag{5}$$

where the cost function f_i is given by

$$f_i(x) = \max_{j \in \mathcal{J}_i} J_{ij}(x) \tag{6}$$

with $J_{ij}(x) = \log\left(\sigma_i^2 h_{i,j}^{-2} e^{-x_j} + \sum_{s \neq j} h_{i,s}^2 h_{i,j}^{-2} e^{x_s - x_j}\right) + \mathcal{V}_j(e^x)$ and the constraint set is

$$X = \{x : \log(p_{\min}) \leq x_j \leq \log(p_{\max}), j = 1, \ldots, m\}.$$

If we let $x^\star$ denote a solution to (5), then, the optimal transmission power vector $p^\star$ can be recovered as $p^\star = \log(x^\star)$ where $\log(\cdot)$ applied to a vector should be interpreted as the log function applied to each component of the vector, i.e., $p^\star = \begin{bmatrix} \log(x_1^\star) & \log(x_2^\star) & \ldots & \log(x_m^\star) \end{bmatrix}^\top$.

Remark 1: Note that function $J_{ij}(x)$ is convex and differentiable in x since it is the sum of two terms: the log of the sum of exponentials in x and the convex and differentiable function $\mathcal{V}_j(e^x)$. Then, $f_i(\cdot)$ is convex since it is the maximum of convex functions. In the case when there are multiple mobile users served by the base station i, $f_i(\cdot)$ is not guaranteed to be differentiable. Also, $f_i(\cdot)$ is known only to the base station BS_i because it depends on the channel coefficients h_i. Note finally that set X is convex, nonempty ($p_{\min} < p_{\max}$), and also compact since $p_{\min} > 0$, which guarantees that the set $X^\star$ of minimizers of (5) is non-empty. ∎

Solving the constrained optimization problem (5) in a centralized way can be difficult if the cellular network is large. It in fact involves introducing some central unit – possibly one of the base stations – and making all base stations send to the central unit all their (local) information on the estimates of the channel coefficients so as to allow for the computation of the optimal transmission power values of all mobile users by solving (5). These values are then communicated to the base stations, which in turn assign them to the mobile users that they are serving. This will possibly cause an overload of the wired communication network connecting the base stations.

We next propose a distributed approach to the solution of (5) that does not require to introduce any central unit, neither the base stations to share the channel coefficients of their mobile users, which are considered to be a sensitive information. The approach is described in Algorithm 1.

Algorithm 1 Distributed power control proximal algorithm

1: **Initialization**
2:    $k = 0$
3:    Consider $x^i(0) \in X$, for all $i = 1, \ldots, n$
4: **For** $i = 1, \ldots, n$ **repeat until convergence**
5:    $\hat{x}^i(k) = \sum_{j=1}^{n} a_j^i x^j(k)$
6:    $x^i(k+1) = \arg\min_{x^i \in X} \left\{ f_i(x^i) + \dfrac{1}{2c(k)} \|x^i - \hat{x}^i(k)\|^2 \right\}$
7:    $p^i(k) = \log(x^i(k))$
8:    $k \leftarrow k + 1$

The weights $a_j^i \geq 0$, $j = 1, \ldots, n$, at step 5 of Algorithm 1 encode the wired network (directed) communication graph, in that if $a_j^i = 0$, then, BS_i does not receive any information from BS_j. The set of BS_j such that $a_j^i > 0$ are called neighbors of BS_i.

Initially, each base station BS_i, $i = 1, \ldots, n$, makes its own guess on some tentative value $x^i(0)$ for the solution to (5) (step 3, Algorithm 1). One possible choice is $x^i(0) \in \arg\min_{x^i \in X} f_i(x^i)$. At iteration k, each BS_i constructs a weighted average $\hat{x}^i(k)$ of the tentative solutions $x^j(k)$, $j = 1, \ldots, n$, received by its neighbors and its local one (step 5, Algorithm 1). Step 6 of Algorithm 1 is a proximal minimization computation, where BS_i solves a local minimization problem, updating its tentative solution with a value within X that minimizes the sum of its local objective function $f_i(\cdot)$ and a weighted quadratic term, which accounts for the difference of its current tentative solution from the average $\hat{x}_i(k)$. The relative importance of the two terms is determined by $c(k) \in \mathbb{R}_+$. Note that since $f_i(\cdot)$ is convex (see Remark 1) and the quadratic penalty term is strictly convex, the resulting minimization problem admits a unique solution at each iteration k.

We next present some assumptions that are needed to prove converge of Algorithm 1 to an optimal solution $p^\star$ of the constrained optimization problem (4).

Assumption 1 (Connectivity and Communication): Let (V, E) be the directed graph with nodes $V = \{1, \ldots, n\}$ and edges $E = \{(j, i) : a_j^i > 0\}$. We assume that (V, E) is strongly connected, i.e., for any two nodes there exists a path of directed edges that connects them. ∎

Coefficients $\{c(k)\}_{k \geq 0}$ in step 6 of Algorithm 1 have to be chosen so as to satisfy the following assumption.

Assumption 2 (Coefficient $\{c(k)\}_{k \geq 0}$): The sequence $\{c(k)\}_{k \geq 0}$ with $c(k) \in \mathbb{R}_+$ is non-increasing and satisfies $\sum_{k=0}^{\infty} c(k) = \infty$, and $\sum_{k=0}^{\infty} c(k)^2 < \infty$. ∎

One possible choice for $\{c(k)\}_{k \geq 1}$ that satisfies 2 is $c(k) = \alpha/(k+1)^\beta$ for some $\alpha \in \mathbb{R}_+$ and $0.5 < \beta \leq 1$. Assumption 2 is identical to the conditions imposed in [15], [16] on the step-size of their subgradient algorithms.

As in other contributions on distributed optimization in the literature (see, e.g., [11], [17], [18], [19], [20], [15] to name a few), we impose the following assumption on the weights of the average tentative solution.

Assumption 3 (Weight coefficients): The coefficients a_j^i,

$i, j = 1, \ldots, n$, satisfy the following conditions:

1) $a_j^i \geq 0$ for all $i, j = 1, \ldots, n$,
2) $a_i^i > 0$ for all $i = 1, \ldots, n$,
3) $\sum_{j=1}^n a_j^i = 1$ for all $i = 1, \ldots, n$,
4) $\sum_{i=1}^h a_j^i = 1$ for all $j = 1, \ldots, n$,

where the last two conditions imply that the matrix with element (i, j) equal to a_j^i is doubly stochastic. ∎

Note that all base stations are required to agree on the same doubly stochastic matrices. This is a quite standard assumption in distributed optimization algorithms of this type (see also [15], [19], [16]). In the case when $a_j^i = a_i^j$, $i, j = 1, \ldots n$, i.e., when (V, E) is an undirected graph, the agreement can be achieved in one single iteration following the distributed approach in [20].

Theorem 1 (Optimality): Consider Assumptions 1-3 and Algorithm 1. We have that for some maximizer $p^\star$ of problem (4), all base stations reach consensus to $p^\star$, i.e.,

$$\lim_{k \to \infty} \|p^i(k) - p^\star\| = 0, \ i = 1, \ldots, n.$$

Proof: Given the convexity of $f_i(\cdot)$ and the convexity and compactness of X (see Remark 1), by applying the theorem on optimality of the distributed proximal algorithm in [9], [10], we get that for some minimizer $x^\star$ of problem (5) the following convergence property holds

$$\lim_{k \to \infty} \|x^i(k) - x^\star\| = 0, \ i = 1, \ldots, n.$$

Since the mapping $p^i(k) = \log(x^i(k))$ from x^i to p^i is a monotonically increasing bijective function and the cost that is minimized in (5) is obtained by multiplying by -1 the performance index that is maximized in (4), by defining $p^\star = \log(x^\star)$ the statement in the theorem follows immediately. ∎

Remark 2 (Resilience to failures): The optimality result in Theorem 1 is preserved when temporary failure of communication links occurs. This is because the asymptotic optimality result in [9], [10] holds with time-varying weight coefficients, under suitable long run connectivity conditions, and a time-varying variant of Assumption 3 (see Assumption 2 in [10]). Also, if one of the base stations definitely breaks down, its mobile users can be re-assigned to the closest base station, and the system will automatically adapt to the new configuration, if the distributed algorithm for transmission power control adjusts the weight coefficients so as to comply with Assumption 3. ∎

IV. SIMULATION-BASED PERFORMANCE ASSESSMENT AND COMPARATIVE ANALYSIS

In this section, we show the performance of the proposed distributed non-autonomous algorithm for power control in a single channel wireless cellular network, and compare it with the sub-gradient based alternative proposed in [11], which is here presented in Algorithm 2 to comply with our notation.

In Algorithm 2, $\Pi_X[\cdot]$ denotes the projection onto set X (which is straightforward since X is a box), and $\nabla f_i(\cdot)$ the sub-gradient of function $f_i(\cdot)$, which satisfies

$$\nabla f_i(x)^\top (y - x) \leq f_i(y) - f_i(x), \ x, y \in X.$$

Algorithm 2 Distributed power control sub-gradient algorithm

1: **Initialization**
2:   $k = 0$
3:   Consider $x^i(0) \in X$, for all $i = 1, \ldots, n$
4: For $i = 1, \ldots, n$ **repeat until convergence**
5:   $\hat{x}^i(k) = \sum_{j=1}^n a_j^i x^j(k)$
6:   $x^i(k+1) = \Pi_X[\hat{x}^i(k) - c(k)\nabla f_i(\hat{x}^i(k))]$
7:   $p^i(k) = \log(x^i(k))$
8:   $k \leftarrow k + 1$

Given that $f_i(\cdot)$ in (6) is convex (see Remark 1), the subgradient at $x \in X$ is well-defined. However, it is not differentiable in the case when the base station is serving multiple users. The sub-gradient computation involves the choice of the mobile user with the worst throughput at every iteration, given the current value x of the iterated weighted average $\hat{x}^i$. More specifically, we have that

$$\nabla f_i(x) = \nabla \bar{J}_i(x), \tag{7}$$

where $\nabla \bar{J}_i(\cdot)$ is the standard gradient of the function

$$\bar{J}_i(\cdot) = J_{ij_x}(\cdot) \text{ with } j_x = \arg\max_{j \in \mathcal{J}_i} J_{ij}(x), \tag{8}$$

which is convex and differentiable.

The same result in Theorem 1 holds for Algorithm 2 (see [11] for a proof). The two algorithms work in the same set-up and involve the same communication structure, both avoiding local information sharing. The main difference is how the tentative solution is updated at every iteration, and, in particular, the methods used by each base station to solve its local optimization problem.

A clear advantage of Algorithm 2 compared with Algorithm 1 proposed in this paper is that it is computationally less demanding since it does not involve any optimization over X. However, if we consider the multi-user case that is indeed the most common in practice, Algorithm 2 is typically affected by oscillations of the solution before convergence is reached. This is caused by the non-differentiability of $f_i(\cdot)$ and the fact that the sub-gradient calculation in (7) involves the identification of the mobile user with the worst throughput given the current average tentative solution (see (8)).

In the simulation results, we consider cellular networks with n base stations that are located on a regularly gridded square area with grid parameter equal to 5. Each base station is at the vertex of one or multiple squares and is wired connected to all the base stations that belong to its same squares, either on the same edge and or on the diagonal (see Figure 1). Each mobile user is served by the closest base station.

The channel coefficient $h_{i,j}$ of the communication link from the mobile user j to the base station i is set to zero if their distance is larger than or equal to 200, otherwise it decays to zero as the fourth power of the distance. A shadow fading factor modeled as lognormal with mean 1.05 and variance 0.1 is introduced along each uplink. The receiver

Algorithm 3 Sinkhorn-Knopp algorithm

1: $a_j^i \leftarrow 1$ if $(i,j) \in E$
2: $a_j^i \leftarrow 0$ if $(i,j) \notin E$
3: **repeat**
4:     $a_j^i \leftarrow a_j^i / \sum_{c=1}^{m} a_c^i, \ \forall i = 1, \ldots, m$
5:     $a_j^i \leftarrow a_j^i / \sum_{r=1}^{m} a_j^r, \ \forall j = 1, \ldots, m$
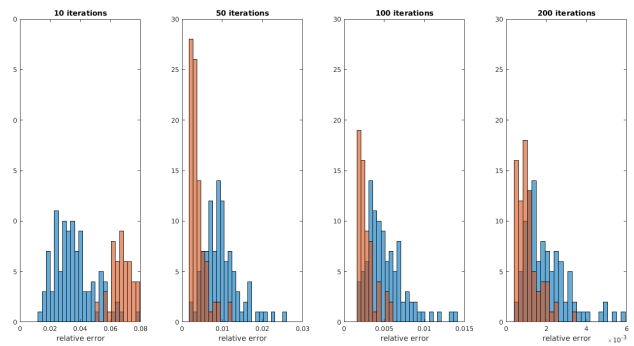6: **until** Assumption 3 is satisfied within a given tolerance

Fig. 2. Histograms of the normalized relative cost for the cellular network with 16 base stations and 16 mobile users at iterations $k = 10, 50, 100$, and 200, of Algorithms 1 (in blue) and 2 (in red).

Fig. 3. Cost function for a cellular network with 16 base stations and 16 mobile users for Algorithms 1 (in blue) and 2 (in red).

noise variance is set equal to 10^{-4} at each base station. We set $\log(p_{\min}) = 0$ and $\log(p_{\max}) = 7$ when defining the set X. The term penalizing the transmission power cost in (3) is given by $\mathcal{V}_j(p) = 10^{-3}p_j$.

Parameter $\{c(k)\}_{k \geq 0}$ appearing in both algorithms is set equal to $c(k) = \frac{\alpha}{(k+1)^\beta}$, where $\alpha = n$ and $\beta = 0.7$. The initial value for x^i, $i = 1, \ldots, n$, in both algorithms is set equal to the local optimal solution: $x^i(0) \in \arg\min_{x^i \in X} f_i(x^i)$.

Lastly, we set the weights coefficients so as to satisfy Assumption 3. To this end, we use the procedure described in [21] and formulate Algorithm 3, which is initialized with the adjacency matrix $\mathcal{A}$ of the graph (V, E) (i.e., the matrix which has 1 in position (i,j) if $(i,j) \in E$ and zero otherwise). Note that $(i,i) \in E$ since we require $a_i^i > 0$ in Assumption 3. Then, under Assumption 1, the adjacency matrix has full support (see [21] for a definition) and Algorithm 3 is guaranteed to converge to a doubly stochastic matrix with the same sparsity pattern of $\mathcal{A}$.

Given that each base station is in charge of setting the transmission power of the mobile users that it is serving, when plotting the performance at time k we shall refer to the transmission power vector $\hat{p}(k)$ that contains as elements the transmission powers computed at iteration k by the base stations for the mobile users that they are serving.

Performance evaluation and comparative analysis of the distributed Algorithms 1 and 2 are performed in terms of either the cost

$$\sum_{i=1}^{n} f_i(\log(\hat{p}(k))), \ k = 0, 1, \ldots,$$

or the normalized relative error cost

$$\frac{\sum_{i=1}^{n} f_i(\log(\hat{p}(k))) - \min_{x \in X} \sum_{i=1}^{n} f_i(x)}{\min_{x \in X} \sum_{i=1}^{n} f_i(x)}, \ k = 0, 1, \ldots.$$

All the numerical simulations were run using Matlab with CVX, [22], as optimization interface and Mosek$^{\text{TM}}$, as solver. In the first simulation study, we consider a cellular network with $n = 16$ base stations and $m = 16$ mobile users, one per base station.

We performed 100 runs of both Algorithms 1 and 2, where the position of each mobile user is extracted at random in a 5 by 5 square centered in a base station. The resulting histograms of the normalized relative costs at iterations $k = 10, 50, 100$, and 200, are plotted in Figure 2. The relative error decreases progressively for both, after an initial

transient phase Algorithm 2 performs slightly better than Algorithm 1. This is also apparent from Figure 3 where the plots of the cost obtained with the two algorithms are reported for one of the extracted configurations. Note that both plots are smooth curves. As for the overshoot in the cost for Algorithm 2, $c(k)$ is large initially and makes the algorithm take a large step in the wrong direction, thus increasing the cost.

If we now consider the multi-user set-up, with $n = 16$ base stations but $m = 20$ mobile users, and perform the same kind of experiments, we get the histograms of the normalized relative cost and the plots for the cost in Figures 4 and 5, respectively. Convergence is slower and oscillations deteriorate the performance of the solution obtained by Algorithm 2.

In Figure 6, we report the time needed to run 100 iterations of both algorithms for different instances of 5 different single-user cellular networks. As expected Algorithm 2 is less time consuming.

V. CONCLUDING REMARKS

In this paper we proposed a distributed transmission power control algorithm for a wireless cellular network. The algorithm is based on proximal minimization and represents an alternative to a gradient-based distributed algorithm that

Fig. 4. Histograms of the normalized relative cost for the cellular network with 16 base stations and 20 mobile users at iterations $k = 10, 50, 100$, and 200, of Algorithms 1 (in blue) and 2 (in red).
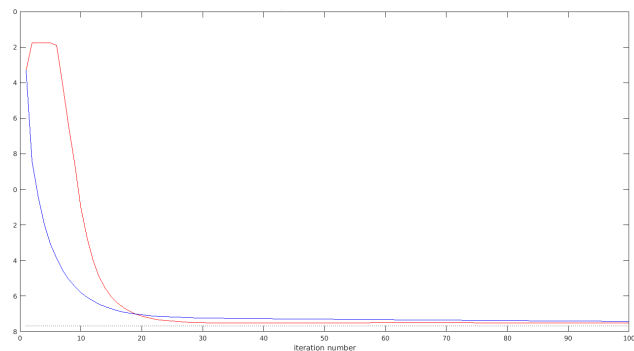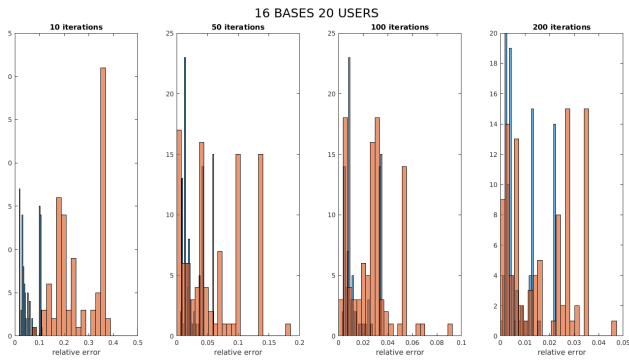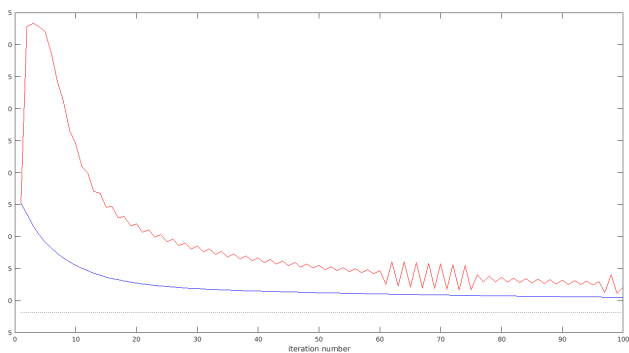

Fig. 5. Cost function for a cellular network with 16 base stations and 20 mobile users for Algorithms 1 (in blue) and 2 (in red). The dotted constant line is the optimal cost.

works in the same set-up, with the same guarantees on convergence and optimality. The two approaches show similar performance in the case when each base station serves a single mobile user in the cellular network. Admittedly, the proposed approach is computationally more demanding. However, in the relevant case of multiple mobile users per base station, it shows better performance than the gradient-based approach since, as an effect of the cost being non-differentiable, using a sub-gradient induces an oscillatory behavior.

REFERENCES

[1] M. Chiang, P. Hande, T. Lan, and C. W. Tan, "Power control in wireless cellular networks," *Foundations and Trends® in Networking*, vol. 2, no. 4, pp. 381–533, 2008.

[2] R. D. Yates, "A framework for uplink power control in cellular radio systems," *Selected Areas in Communications, IEEE Journal on*, vol. 13, no. 7, pp. 1341–1347, 1995.

[3] D. Zhao, "Cdma-based wireless cellular networks," in *Power Distribution and Performance Analysis for Wireless Communication Networks*. Springer, 2012, pp. 17–37.

[4] G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *Vehicular Technology, IEEE Transactions on*, vol. 42, no. 4, pp. 641–646, 1993.

[5] T. Alpcan, T. Başar, R. Srikant, and E. Altman, "Cdma uplink power control as a noncooperative game," *Wireless Networks*, vol. 8, no. 6, pp. 659–670, 2002.
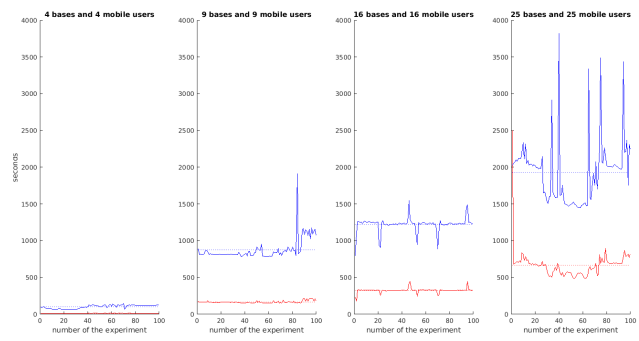
[6] D. Famolari, N. Mandayam, D. Goodman, and V. Shah, "A new framework for power control in wireless data networks: Games, utility, and pricing," in *Wireless Multimedia Network Technologies*. Springer, 2002, pp. 289–309.

[7] P. Hande, S. Rangan, M. Chiang, and X. Wu, "Distributed uplink power control for optimal sir assignment in cellular data networks," *Networking, IEEE/ACM Transactions on*, vol. 16, no. 6, pp. 1420–1433, 2008.

[8] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 5, pp. 1074–1084, 2006.

[9] K. Margellos, A. Falsone, S. Garatti, and M. Prandini, "Proximal minimization based distributed convex optimization," in *American Control Conference (ACC), 2016*. American Automatic Control Council (AACC), 2016, pp. 2466–2471.

[10] ——, "Distributed constrained optimization and consensus in uncertain networks via proximal minimization," *IEEE Transactions on Automatic Control*, pp. 1–16, 2017, to appear.

[11] S. S. Ram, V. V. Veeravalli, and A. Nedić, "Distributed non-autonomous power control through distributed convex optimization," in *INFOCOM 2009, IEEE*. IEEE, 2009, pp. 3001–3005.

[12] S. Grandhi and J. Zander, "Constrained power control in cellular radio systems," in *Proceedings of IEEE Vehicular Technology Conference (VTC)*, vol. 2, 1994, pp. 824–828.

[13] K.-L. Hsiung, S.-J. Kim, and S. Boyd, "Power control in lognormal fading wireless channels with uptime probability specifications via robust geometric programming," in *Proceedings of the 2005, American Control Conference, 2005.*, vol. 6, 2005, pp. 3955–3959.

[14] M. Chiang, "Geometric programming for communication systems," *Foundations and Trends in Communications and Information Theory*, vol. 2, no. 1‑2, pp. 1–154, 2005.

[15] A. Nedic, A. Ozdaglar, and P. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 922–938, 2010.

[16] M. Zhu and S. Martinez, "On distributed convex optimization under inequality and equality constraints," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 151–164, 2012.

[17] J. Tsitsiklis, *Problems in decentralized decision making and computation*. Ph.D. Dissertation, MIT, Cambridge, MA, 1984.

[18] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Transactions on Automatic Control*, vol. 31, no. 9, pp. 803–812, 1986.

[19] A. Olshevsky and J. Tsitsiklis, "Convergence speed in distributed convergence and averaging," *SIAM Review*, vol. 53, no. 4, pp. 747–772, 2011.

[20] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.

[21] R. Sinkhorn and P. Knopp, "Concerning nonnegative matrices and doubly stochastic matrices," *Pacific Journal of Mathematics*, vol. 21, no. 2, pp. 343–348, 1967.

[22] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

Fig. 6. Simulation time for running 100 iterations of Algorithms 1 (in blue) and 2 (in red) in 100 tests.