JOINT CELL SELECTION AND RADIO RESOURCE ALLOCATION IN MIMO SMALL CELL NETWORKS VIA SUCCESSIVE CONVEX APPROXIMATION

Stefania Sardellitti,[1] *Gesualdo Scutari,*[2] *and Sergio Barbarossa*[1]

[1]Dept. of Information Engineering, Electronics and Telecommunications
Sapienza University of Rome, Via Eudossiana 18, 00184, Rome, Italy
[2]Dept. of Electrical Engineering, State University of New York at Buffalo, Buffalo, USA
E-mail: stefania.sardellitti@uniroma1.it, gesualdo@buffalo.edu, sergio.barbarossa@uniroma1.it

ABSTRACT

It is widely recognized that one of the factors that are going to yield the most significant capacity increase in wireless networks is spatial reuse of radio resources through dense deployment of radio access points. This leads to the development of small cell networks where different size cells, e.g. macro cells, picocells, femtocells, relays, coexist under the same standard. Of course, dense deployment is able to unravel its potential benefits only provided that interference is properly managed. In this paper, we propose an algorithm able to perform cell association and radio resource allocation *jointly*, in order to maximize the sum rate in a MIMO (interference) network. Cell selection is inherently a combinatorial problem. To deal with the nonconvexity, we introduce a suitably chosen convex relaxation of the objective function and develop a fast algorithm converging to a locally optimal solution of the nonconvex problem.

***Index Terms*—** Resource allocation, Cooperation, MIMO systems, Interference channels, Successive convex approximation.

1. INTRODUCTION

The foreseen massive deployment of small cell base stations covering limited areas and coexisting with current macrocells and picocells is expected to produce a capacity boost, provided that appropriate interference management mechanisms are incorporated in the network [1]. On the other hand, high capacity wired links among small cell base stations create the opportunities for cooperative multipoint communications among base stations, which makes possible better interference management. In the scenario created by a massive deployment of small cell networks, one of the primary tasks is to devise a proper cell selection mechanism assigning mobile users to small cell base stations (BS), taking into account the coexistence of superimposed cells in the same area. The problem of jointly optimizing cell selection and radio resource allocation was considered in [2] for single antenna systems. Of course, the availability of multiple antennas at the BS facilitates dynamic assignment using receive beamforming. In [3] and [4], the authors studied the joint cell assignment/allocation problem in uplink and downlink systems, respectively; they proposed a joint optimization of the mobile transmit powers and receive beamforming in order to minimize the total transmit power, subject to Signal to Interference plus Noise Ratio (SINR) constraints. A joint network optimization (cell association) and beamforming for Coordinated Multipoint (CoMP) transmissions was recently considered in [5], [6]. In [5] the problem was formulated as a mixed integer second-order cone programming with the goal of minimizing the overall BS power consumption, while guaranteeing Quality of Service (QoS) to the mobile users. The assignment/allocation problem was then reconsidered in [7] in a more general MIMO set-up. In such a case, the goal is to find the optimal precoding matrices and BS assignment jointly. In the effort to devise distributed solutions, with limited exchange of information among the nodes, in [7], the problem is formulated as a game incorporating a pricing mechanism to "punish" mobile users for causing interference. A different perspective to the BS assignment problem was recently proposed in [8], where the goal is minimizing the number of active base stations still being able to guarantee a QoS to the mobile users. This is indeed a very promising approach in the current green paradigm, to save the overall energy consumption, as it allows to switch off the BS's dynamically, depending on the overall traffic load, without incurring in appreciable performance degradation.

In this paper, we propose a fairly general approach to the joint BS association and resource allocation problem for MIMO transceivers, with the goal of finding the precoding matrices and BS associations in order to maximize the sum rate. Each BS is enabled to decode a set of mobile users jointly, where the number of users assigned to each cell is to be determined based upon the interference conditions. We impose only an upper bound on the maximum number of decodable signals dictated by complexity constraints as well as by the need to guarantee the existence of a zero-forcing decoder. The latter constraint is dictated by the ratio between the number of receive and transmit antennas. The optimization variables are the precoding matrices and the binary values associating each mobile user to a base station. In general, the association problem is combinatorial. To deal with this inherent nonconvexity, we develop a provable convergent solution method hinging on the idea of Successive Convex Approximations (SCAs) [9]: The original nonconvex problem is replaced by a sequence of suitably chosen low complexity convex subproblems. The method is proved to converge to a local optimal solution of the nonconvex problem. Quite remarkably, the proposed approximation is such that every limit point generated by the algorithm is discrete in the BS association variables. Therefore, there is no need of any further, somehow arbitrary, discretization of the solution. Finally, numerical results show that the algorithm is remarkably fast and able to yield a significant performance improvement while enabling the receivers to decode a variable number of users, depending on the overall interference scenario.

2. SYSTEM MODEL AND PROBLEM FORMULATION

Let us consider a small cell network where I small cell base stations, or small cell enhanced Node B (SCeNBs) in LTE terminology,

The work of Barbarossa and Sardellitti was funded by the European Community 7th Framework Programme Project ICT-TROPIC, under grant nr. 318784. The work of Scutari was supported by the USA NSF under Grants CMS 1218717 and CAREER Award No. 1254739.

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

aim to serve J mobile users. One of the distinguishing features of small cell networks, with respect to macro-cellular networks, is that the number of active mobile users is of the same order of magnitude of the number of small cell access points. For notation simplicity and without loss of generality, in this work, we consider the case of $I = J$. We study the uplink transmissions, assuming the mobile users and the receiving BS's to be equipped with n_{T_i} and n_{R_i} antennas, respectively, with $i = 1, \ldots, I$. Our goal is to find out the set of mobile users to be served by each base station and the precoding matrix of each transmitter, in order to maximize the sum of rates in each cell. Just because the number of users served by each access point is a variable, some base stations can end up with no users to serve, whereas other stations may have multiple users to be served simultaneously. We impose an upper bound on this number. Under this scenario, the users assigned to the same cell do not interfere with each other, whereas of course there is intercell interference. Furthermore, in this work we assume that each mobile belongs to only one cell. Future developments may include the case where a user may be served by multiple cells. To identify which user is served by which base station, we introduce the so called cooperation matrix $\boldsymbol{A}$, whose coefficients $(a_{ij})_{i,j=1}^{I}$ are such that $a_{ij} = 1$ if mobile user j is associated to the i-th BS, and $a_{ij} = 0$ otherwise. We assume these binary variables to satisfy the following conditions: i) $a_{ij} = 0$ if $j \notin \mathcal{N}_i$ where $\mathcal{N}_i$ represents the *largest* set of users that BS i can serve, whose cardinality $|\mathcal{N}_i|$ is $D_i \triangleq |\mathcal{N}_i|$; ii) $\sum_j a_{ij} \leq D_i$, $\forall i$; and iii) $\sum_i a_{ij} = 1$, $\forall j$, since each mobile can belong only to one coalition. According to this notation, the nonzero entries of each vector $\mathbf{a}_i \triangleq (a_{ij})_{j=1}^{I}$ represent the set of users assigned to BS i. Additionally, the transmit covariance matrix $\boldsymbol{Q}_i \in \mathbb{C}^{n_{T_i} \times n_{T_i}}$, to be optimized by each mobile user, is subject to the following power constraints:

$$\mathcal{Q}_i \triangleq \left\{\mathbf{Q}_i \in \mathbb{C}^{n_{T_i} \times n_{T_i}} : \mathbf{Q}_i \succeq \mathbf{0},\ \text{tr}\left(\mathbf{Q}_i\right) \leq P_i,\ \text{and } \mathbf{Q}_i \in \mathcal{P}_i\right\},$$

where P_i is the average transmit power of user i, and $\mathcal{P}_i \subseteq \mathbb{C}^{n_{T_i} \times n_{T_i}}$ is an arbitrary convex and closed set suitable to accommodate additional local constraints, such as interference constraints, null constraints, per-antenna constraints, etc.

Throughout the paper, we will use the following notation: $\mathcal{Q} \triangleq \{\mathbf{Q} \triangleq (\mathbf{Q}_i)_{i=1}^{I} : \mathbf{Q}_i \in \mathcal{Q}_i, \forall i = 1, \ldots, I\}$ is the joint set of users' power constraints, $\mathbf{Q}_{-i} \triangleq (\mathbf{Q}_j)_{j \neq i}$ is the tuple of all users' covariance matrices except the one of user i, whose associated feasible set is $\mathcal{Q}_{-i} \triangleq \{\mathbf{Q}_{-i} : \mathbf{Q}_j \in \mathcal{Q}_j, \forall j \neq i\}$; similarly we define $\mathbf{a} \triangleq (\mathbf{a}_i)_{i=1}^{I}$. In the above setting, the sum-rate of each cell i is:

$$R_i(\mathbf{Q}, \mathbf{a}_i) = \log\det\left(\mathbf{I} + \left(\sum_{j=1}^{I} a_{ij}\mathbf{H}_{ij}\mathbf{Q}_j\mathbf{H}_{ij}^{H}\right)\widetilde{\mathbf{R}}_i\left(\mathbf{Q}, \mathbf{a}_i\right)^{-1}\right)$$

where $\mathbf{H}_{ij} \in \mathbb{C}^{n_{R_i} \times n_{T_j}}$ is the cross-channel matrix between the transmitter j and the receiver (BS) i, and

$$\widetilde{\mathbf{R}}_i\left(\mathbf{Q}, \mathbf{a}_i\right) = \mathbf{R}_{n_i} + \sum_{j=1}^{I}\left(1 - a_{ij}\right)\mathbf{H}_{ij}\mathbf{Q}_j\mathbf{H}_{ij}^{H} \tag{1}$$

is the covariance matrix of the multiuser interference plus noise, with $\mathbf{R}_{n_i} \succ \mathbf{0}$. The system design proposed in this paper consists in optimizing *jointly* the users' covariance matrices $\mathbf{Q} = (\mathbf{Q}_i)_{i=1}^{I}$ and the cooperation matrix $\boldsymbol{A}$ in order to maximize the sum-rate of all cells. More formally, we have the following

$$\begin{array}{ll} \max\limits_{\mathbf{Q},\mathbf{a}} & U(\mathbf{Q}, \mathbf{a}) \triangleq \sum\limits_{i=1}^{I} R_i(\mathbf{Q}, \mathbf{a}_i) \\ \text{s.t.} & \mathbf{Q} \in \mathcal{Q}, \\ & \mathbf{a}_i \in \{0,1\}^{I},\ \mathbf{1}_I^T\mathbf{a}_i \leq D_i,\ \sum\limits_{i=1}^{I} a_{ij} = 1,\ \forall i, j, \end{array} \tag{2}$$

where $\mathbf{1}_I$ denotes the column vector of all ones. The constraints on the entries a_{ij} are introduced to enforce a maximum number of users per cell D_i and the requirement for each mobile user to be served by a single BS. The value D_i has to be chosen in order to enable zero-forcing decoding which implies $D_i = n_{R_i}/n_T$ with $n_{R_i} \geq n_T$ and assuming all mobiles equipped with the same number of transmit antennas.

Unfortunately, problem (2) is NP-hard in general, even if one focuses only on the optimization of the covariance matrices $\mathbf{Q}$. Things are even more complicated because of the integer nature of the coefficients a_{ij}. It turns out that it is not possible to compute its global optimal solution in polynomial time. Motivated by the above NP-hardness, we propose computational affordable algorithms still able to obtain high-quality (albeit suboptimal in principle) solutions for problem (2). To this end, we start using the popular relaxation of the integer constraints, enabling the entries a_{ij} to belong to the following convex set

$$\mathcal{A} \triangleq \left\{(a_{ij})_{i,j=1}^{I} : a_{ij} \in [0,1], \sum_{j \in \mathcal{N}_i} a_{ij} \leq D_i, \sum_{i=1}^{I} a_{ij} = 1, \forall i, j\right\}. \tag{3}$$

The relaxed version of (2) can then be expressed as

$$\begin{array}{ll} \max\limits_{\mathbf{Q},\mathbf{a}} & U(\mathbf{Q}, \mathbf{a}) = \sum\limits_{i=1}^{I} R_i(\mathbf{Q}, \mathbf{a}_i) \\ \text{s.t.} & \mathbf{Q} \in \mathcal{Q}, \\ & \mathbf{a} \in \mathcal{A}. \end{array} \tag{4}$$

Note that (4) is still NP-hard [10]. In what follows, we will exploit the structure of (4) and, building on some recent techniques introduced in [9], we develop a fast distributed algorithm converging to a local optimal solution of (4). Quite interestingly, every stationary solution computed by the proposed algorithm is feasible for the original problem (2) and, in particular, the coefficients a_{ij} of any such solutions turn out to be integer.

3. ALGORITHMIC DESIGN

To solve the non-convex problem (4) efficiently, we develop an SCA-based method [9] where all the users solve a sequence of *decoupled strongly convex* optimization problems in the $\boldsymbol{Q}$-variables. At the basis of the proposed technique, there is a suitably chosen concave and separable approximation of the sum-rate function $U(\mathbf{Q}, \mathbf{a})$, which is preliminarily discussed next.

3.1. A (successive) convex approximation function for $U(\mathbf{Q}, \mathbf{a})$

We start observing that each rate $R_i(\mathbf{Q}, \mathbf{a}_i)$ can be written as

$$R_i(\mathbf{Q}, \mathbf{a}_i) \triangleq f_i^{(1)}(\mathbf{Q}) + f_i^{(2)}(\mathbf{Q}, \mathbf{a}_i) \tag{5}$$

where

$$f_i^{(1)}(\mathbf{Q}) \triangleq \log\det\left(\mathbf{R}_{n_i} + \sum_{j=1}^{I}\mathbf{H}_{ij}\mathbf{Q}_j\mathbf{H}_{ij}^{H}\right) \tag{6}$$

is a (twice-continuously $\mathbb{R}$-differentiable with Lipschitz continuous conjugate gradient [11] and) *concave* function on $\mathcal{Q}$, whereas

$$f_i^{(2)}(\mathbf{Q},\mathbf{a}_i) \triangleq -\log\det\left(\mathbf{R}_{n_i} + \sum_{j=1}^{I}(1-a_{ij})\,\mathbf{H}_{ij}\mathbf{Q}_j\mathbf{H}_{ij}^H\right)$$

is a twice-continuously ($\mathbb{R}$-)differentiable and *convex* function on $\mathcal{Q}\times\mathcal{A}_i$. For notational simplicity, let us define $\mathbf{Z}\triangleq(\mathbf{Q},\mathbf{a}) = (\mathbf{Z}_i)_{i=1}^I$, with $\mathbf{Z}_i \triangleq (\mathbf{Q}_i,\mathbf{a}_i)$ being the optimization variables of user i; associated with each $\mathbf{Z}$ let us also introduce the joint strategy set $\mathcal{Z}\triangleq\{\mathbf{Z}\triangleq(\mathbf{Q},\mathbf{a})\,:\,\mathbf{Q}\in\mathcal{Q}\text{ and }\mathbf{a}\in\mathcal{A}\}$. Finally, given the matrix tuples $\mathbf{A}\triangleq(\mathbf{A}_i)_{i=1}^I$ and $\mathbf{B}\triangleq(\mathbf{B}_i)_{i=1}^I$, with each $\mathbf{A}_i$ and $\mathbf{B}_i$ complex matrix whose product $\mathbf{A}_i^H\mathbf{B}_i$ is well-defined, let $\langle\mathbf{A}\,,\mathbf{B}\rangle\triangleq\sum_i\text{Re}\{\text{tr}\left(\mathbf{A}_i^H\mathbf{B}_i\right)\}$. Similarly, with a slight abuse of notation, we will use the same symbol also to denote the inner product between real vector tuples $\mathbf{a}\triangleq(\mathbf{a}_i)_{i=1}^I, \mathbf{b}\triangleq(\mathbf{b}_i)_{i=1}^I$, that is $\langle\mathbf{a}\,,\mathbf{b}\rangle\triangleq\sum_i\mathbf{a}_i^T\mathbf{b}_i$, with each $\mathbf{a}_i^T\mathbf{b}_i$ well defined.

Exploiting the concavity-convexity structure of each $R_i(\mathbf{Q},\mathbf{a}_i)$ as in (5), a concave approximation of the sum-rate function $U(\mathbf{Q},\mathbf{a})$ for each user i can be obtained by retaining the concave part of $U(\mathbf{Q},\mathbf{a})$ with respect to $(\mathbf{Q}_i,\mathbf{a}_i)$ while linearizing the rest. More formally, introducing the notation $\mathbf{Z}^\nu = (\mathbf{Q}_i^\nu,\mathbf{a}_i^\nu)_{i=1}^I\in\mathcal{Z}$ to denote the value assumed by $\boldsymbol{Z}$ at step ν, we define for each user i the "approximation" function

$$\widetilde{U}_i(\mathbf{Z}_i;\mathbf{Z}^\nu) \quad\triangleq\quad \widetilde{f}_i^{(1)}(\mathbf{Q}_i;\mathbf{Z}^\nu)+\widetilde{f}_i^{(2)}(\mathbf{a}_i;\mathbf{Z}^\nu) \tag{7}$$

with

$$\begin{aligned}\widetilde{f}_i^{(1)}(\mathbf{Q}_i;\mathbf{Z}^\nu) \triangleq\, & f_i^{(1)}(\mathbf{Q}_i,\mathbf{Q}_{-i}^\nu)+\langle\mathbf{\Pi}_i(\mathbf{Z}^\nu)\,,\mathbf{Q}_i-\mathbf{Q}_i^\nu\rangle\\ & -\tau_{Q_i}\,||\mathbf{Q}_i-\mathbf{Q}_i^\nu||^2\end{aligned} \tag{8}$$

and

$$\widetilde{f}_i^{(2)}(\mathbf{a}_i;\mathbf{Z}^\nu)\triangleq f_i^{(2)}(\mathbf{Q}^\nu,\mathbf{a}_i^\nu)+\langle\boldsymbol{\pi}_i(\mathbf{Z}^\nu)\,,\mathbf{a}_i-\mathbf{a}_i^\nu\rangle\,, \tag{9}$$

where: i) $\mathbf{\Pi}_i(\mathbf{Q}^\nu)$ in (8) is the linearization of the terms in $U(\mathbf{Q})$ that are not concave in $\mathbf{Q}_i$, that is

$$\mathbf{\Pi}_i(\mathbf{Z}^\nu)\triangleq\sum_{j\neq i}\nabla_{\mathbf{Q}_i^*}R_j(\mathbf{Z}^\nu)+\nabla_{\mathbf{Q}_i^*}f_i^{(2)}(\mathbf{Q}^\nu,\mathbf{a}_i^\nu), \tag{10}$$

where $\nabla_{\mathbf{Q}_i^*}R_j$ denotes the conjugate gradient of R_j [11], given by

$$\begin{aligned}\nabla_{\mathbf{Q}_i^*}R_j(\mathbf{Z})=&\Big(a_{ji}\mathbf{H}_{ji}^H-(1-a_{ji})\mathbf{H}_{ji}^H\,\widetilde{\mathbf{R}}_j(\mathbf{Q},\mathbf{a}_j)^{-1}\mathbf{F}_j(\mathbf{Q},\mathbf{a}_j)\Big)\\ &\cdot(\mathbf{I}+\widetilde{\mathbf{R}}_j(\mathbf{Q},\mathbf{a}_j)^{-1}\mathbf{F}_j(\mathbf{Q},\mathbf{a}_j))^{-1}\widetilde{\mathbf{R}}_j(\mathbf{Q},\mathbf{a}_j)^{-1}\mathbf{H}_{ji},\end{aligned}$$

and $\nabla_{\mathbf{Q}_i^*}f_i^{(2)}(\mathbf{Q},\mathbf{a}_i) = -(1-a_{ii})\mathbf{H}_{ii}^H\widetilde{\mathbf{R}}_i(\mathbf{Q},\mathbf{a}_i)^{-1}\mathbf{H}_{ii}$, with $\mathbf{F}_j(\mathbf{Q},\mathbf{a}_j)=\sum_{k=1}^I a_{jk}\mathbf{H}_{jk}\mathbf{Q}_k\mathbf{H}_{jk}^H$; ii) $\tau_{Q_i}>0$ in the proximal regularizations in (8) makes $\widetilde{f}_i^{(1)}(\mathbf{Q}_i;\mathbf{Z}^\nu)$ strongly concave; and iii) $\boldsymbol{\pi}_i(\mathbf{Z}^\nu)$ in (9) is the linearization of the convex part $f_i^{(2)}(\mathbf{Q},\mathbf{a}_i)$ w.r.t. $\mathbf{a}_i$, that is $\boldsymbol{\pi}_i(\mathbf{Z}^\nu)\triangleq\nabla_{\mathbf{a}_i}f_i^{(2)}(\mathbf{Z}^\nu) = (\pi_{ij}(\mathbf{Z}^\nu))_{j=1}^I$with $\pi_{ij}(\mathbf{Z}^\nu)\triangleq\text{tr}\left(\mathbf{H}_{ij}\mathbf{Q}_j^\nu\mathbf{H}_{ij}^H\widetilde{\mathbf{R}}_i\left(\mathbf{Q}^\nu,\mathbf{a}_i^\nu\right)^{-1}\right)$. Note that $\widetilde{U}_i(\mathbf{Z}_i;\mathbf{Z}^\nu)$ is a strongly concave function in $\mathbf{Q}_i$ and linear in $\mathbf{a}_i$, for any given $\mathbf{Z}^\nu$. Based on each $\widetilde{U}_i(\mathbf{Z};\mathbf{Z}^\nu)$, we can now define the candidate sum-rate "approximation": For any given $\mathbf{Z}^\nu\in\mathcal{Z}$, let

$$\widetilde{U}(\mathbf{Z};\mathbf{Z}^\nu)\triangleq\sum_{i=1}^I\widetilde{U}_i(\mathbf{Z}_i;\mathbf{Z}^\nu). \tag{11}$$

Note that $\widetilde{U}(\mathbf{Z};\mathbf{Z}^\nu)$ has many desirable properties, such as: i) it is separable in the users variables $\mathbf{Z}_i=(\mathbf{Q}_i,\mathbf{a}_i)$; ii) it is linear in $\mathbf{a}$; and iii) it is uniformly strong concave in $\mathbf{Q}\in\mathcal{Q}$ [12]; we will denote by $c_\tau>0$ the constant of strong concavity, that is c_τ is the largest positive scalar such that [11]

$$\begin{aligned}&\left\langle\mathbf{Q}^1-\mathbf{Q}^2,\nabla_{\mathbf{Q}^*}\widetilde{U}\left((\mathbf{Q}^1,\mathbf{a});\mathbf{Z}^\nu\right)-\nabla_{\mathbf{Q}^*}\widetilde{U}\left((\mathbf{Q}^2,\mathbf{a});\mathbf{Z}^\nu\right)\right\rangle\\ &\quad\leq\;-c_\tau\left\|\mathbf{Q}^1-\mathbf{Q}^2\right\|^2,\quad\forall\mathbf{Q}^1,\,\mathbf{Q}^2\in\mathcal{Q},\;\mathbf{Z}^\nu\in\mathcal{Z},\mathbf{a}\in\mathcal{A}.\end{aligned} \tag{12}$$

An explicit expression of c_τ is given in [12] and is omitted here because of space limitation; we only observe that $c_\tau\geq\min_i\{\tau_{Q_i}\}$, and the equality holds when all $\{\mathbf{H}_{ii}\}_{i=1}^I$ are column rank deficient. We are now ready to introduce the proposed convex approximation of the nonconvex problem (4), which just consists in replacing $U(\mathbf{Z})$ in (4) with the concave approximation $\widetilde{U}(\mathbf{Z};\mathbf{Z}^\nu)$, namely: Given $\mathbf{Z}^\nu\in\mathcal{Z}$,

$$\hat{\mathbf{Z}}(\mathbf{Z}^\nu)\triangleq\left(\hat{\mathbf{Q}}(\mathbf{Z}^\nu),\hat{\mathbf{a}}(\mathbf{Z}^\nu)\right)\in\underset{\mathbf{Z}\triangleq(\mathbf{Q},\mathbf{a})\in\mathcal{Z}}{\arg\max}\left\{\widetilde{U}(\mathbf{Z};\mathbf{Z}^\nu)\right\}, \tag{13}$$

where we denoted by $\hat{\mathbf{Z}}(\mathbf{Z}^\nu) = (\hat{\mathbf{Q}}(\mathbf{Z}^\nu),\hat{\mathbf{a}}(\mathbf{Z}^\nu))\triangleq(\hat{\mathbf{Q}}_i(\mathbf{Z}^\nu),\hat{\mathbf{a}}_i(\mathbf{Z}^\nu))_{i=1}^I$ the solution of (13). Note that the $\boldsymbol{Q}$-part of $\hat{\mathbf{Z}}(\mathbf{Z}^\nu)$ is unique for any given $\mathbf{Z}^\nu$, whereas the $\mathbf{a}$-part generally is not. Thanks to the separability of $\widetilde{U}(\mathbf{Z};\mathbf{Z}^\nu)$ as well as $\mathcal{Z}$ in the $\hat{\mathbf{Q}}_i(\mathbf{Z}^\nu)$'s and $\hat{\mathbf{a}}(\mathbf{Z}^\nu)$, each solution $\hat{\mathbf{Z}}(\mathbf{Z}^\nu)$ can be efficiently computed solving for each i *separately* the following convex optimization problems: Given $\mathbf{Z}^\nu$,

$$\hat{\mathbf{Q}}_i(\mathbf{Z}^\nu)=\underset{\mathbf{Q}_i\in\mathcal{Q}_i}{\arg\max}\left\{\widetilde{f}_i^{(1)}(\mathbf{Q}_i;\mathbf{Z}^\nu)\right\},\quad i=1,\ldots,I \tag{14}$$

and

$$\hat{\mathbf{a}}(\mathbf{Z}^\nu)\in\underset{\mathbf{a}\in\mathcal{A}}{\arg\max}\left\{\sum_{i=1}^I\widetilde{f}_i^{(2)}(\mathbf{a}_i;\mathbf{Z}^\nu)\right\}. \tag{15}$$

Note that (15) is an LP and it is not difficult to see that any solution must satisfy $\hat{\mathbf{a}}(\mathbf{Z}^\nu)\in\{0,1\}^{I^2}$, for any $\mathbf{Z}^\nu\in\mathcal{Z}$ (the feasible set of the problem is a integral polyhedron [13]).

3.2. Algorithmic framework

We are now ready to introduce the proposed distributed algorithm converging to a stationary solution of (4). It consists in solving successively the sequence of convex optimization problems in the form (14) and (15), starting from a feasible $\mathbf{Z}^0$. The formal description of the algorithm is given in Algorithm 1 below and its convergence properties are provided in Theorem 1 (the proof of the theorem is omitted because of space limitation; see [12]). Note that in Step 2 we allow a memory in the update of the iterate $\hat{\mathbf{Z}}(\mathbf{Z}^\nu)$ in the form of a convex combination via $\gamma^\nu\in(0,1]$.

Algorithm 1 Parallel Best-Response Algorithm for (4)

Data: $\mathbf{Z}^0=(\mathbf{Q}^0,\mathbf{a}^0)\in\mathcal{Z}$ for $i=1,\ldots,I$; $\{\gamma^\nu\}>0$, $(\tau_{Q_i})_{i=1}^I\geq\mathbf{0}$; set $\nu=0$;

(S.1): If $\mathbf{Z}^\nu$ satisfies a termination criterion: STOP.

(S.2): In parallel, for all $i=1,\ldots,I$, compute $\hat{\mathbf{Q}}_i(\mathbf{Z}^\nu)$ solving (14) and set $\mathbf{Q}_i^{\nu+1}\triangleq\mathbf{Q}_i^\nu+\gamma^\nu\left(\hat{\mathbf{Q}}_i(\mathbf{Z}^\nu)-\mathbf{Q}_i^\nu\right)$;

(S.3): Compute $\hat{\mathbf{a}}\left(\mathbf{Q}^{\nu+1},\mathbf{a}^\nu\right)$ solving (15) and set $\mathbf{a}^{\nu+1}\triangleq\hat{\mathbf{a}}\left(\mathbf{Q}^{\nu+1},\mathbf{a}^\nu\right)$;

(S.4): Set $\mathbf{Z}^{\nu+1}=\left(\mathbf{Q}^{\nu+1},\mathbf{a}^{\nu+1}\right)$,

(S.5): $\nu\leftarrow\nu+1$ and go back to (S.1).

Theorem 1 *Given the social problem (4), choose* $(\tau_{Q_i})_{i=1}^I\geq\mathbf{0}$, *and* $\{\gamma^\nu\}$ *so that one of the two following conditions is satisfied:*

$$a):\quad 0<\inf_\nu\gamma^\nu\leq\sup_\nu\gamma^\nu\leq\gamma^{\max}\leq 1\ \textit{and}\ 2\,c_{\boldsymbol{\tau}}\geq\gamma^{\max}L_{\nabla U},$$

with $c_{\boldsymbol{\tau}}$ *defined in (12); or*

$$b):\qquad c_{\boldsymbol{\tau}}>0,\;\gamma^\nu\in(0,1],\;\gamma^\nu\to 0,\;\textit{and}\quad\sum_\nu\gamma^\nu=+\infty.$$

Then, either Algorithm 1 converges in a finite number of iterations to a stationary point of (4), or every limit point $(\bar{\mathbf{Q}},\bar{\mathbf{a}})$ *of the sequence*

$\{\mathbf{Z}^\nu\}$ *(at least one such a point exists) is a stationary point of (4) such that* $\bar{a}_{ij} \in \{0, 1\}$*, for all* i, j*. Moreover, none of such points is a local minimum of (4).*

The algorithm implements a novel SCA decomposition: All the users solve *in parallel* a sequence of *decoupled* strongly convex optimization problems in the $\mathbf{Q}$ variables, followed by the update of the assignment vector $\mathbf{a}$. The algorithm is expected to perform better than classical gradient-based schemes (at least in terms of convergence speed) at the cost of no extra signaling, because the structure of the objective functions is better preserved. It is guaranteed to converge under very mild assumptions (the weakest available in the literature) while offering some flexibility in the choice of the free parameters; see [9] for some effective guidelines in the choice of these parameters.

4. NUMERICAL RESULTS

In this section we present some numerical results to assess the performance of the proposed optimization strategy. For all the examples, we consider a MIMO network with equal number $I = 6$ of mobile users and receiving base stations equipped, respectively, with $n_T = 2$ and $n_R = 8$ antennas. The pathloss exponent is assumed to be 2 and the thermal noise power is set to 1.e$-$3. We simulated Algorithm 1 using the diminishing step-size rule satisfying Theorem 1b) $\gamma^{\nu+1} = \gamma^\nu (1 - \alpha\gamma^\nu)$, with $\gamma^0 = 1$ and $\alpha = 1.$e-4. In Fig. 1 we plotted the optimal sum rate $U(\mathbf{Q}^\star, \mathbf{a}^\star)$ versus the iteration index of Algorithm 1, assuming that each base station can serve a maximum number of mobile users $D_i = D$, with $D \leq n_R/n_T$ and for all i. The different curves refer to different values of the upper bound D. The first striking feature observable from Fig. 1 is that the proposed algorithm is able to converge in a very few iterations. Furthermore, we may notice that, as D increases, the sum-rate increases as well and the rate gain with respect to the case $D = 1$ is remarkable. The reason is that, as the maximum number of mobile users served by each base station increases, there are more degrees of freedom to reduce the interference present in each cell. What happens is that the final BS assignment is the result of a trade-off between the need to limit interference, which would lead to associate all users to the same BS, and, at the same time, to have as many active links as possible, which would lead to associating one user per BS. Of course, the rate gain depends on the location of mobile users and access points as well as on the channel model (path loss exponent, etc.). In Fig. 2 we show an example of final assignment resulting as a solution of Algorithm 1, for $D = 3$ and $D = 4$. In this example, the optimal configurations resulting in these two cases coincide. This result is also confirmed by Figure 1. Note that the final base station selection tends to assign each user to its nearest base station, while grouping the potentially strongest interferers in the same cell. Finally, in Fig. 3 we plotted the optimal sum rate averaged over 100 channel realizations versus the network degree D for $I = 8$ (upper subplot) and $I = 6$ (bottom subplot). It can be observed that the average sum rate is an increasing function of both the maximum number of mobile users served by each base station and the number of active base stations.

5. CONCLUSIONS

In this paper we considered the joint BS assignment and precoding design problem for small cell MIMO networks. To deal with the inherently combinatorial nature of the assignment problem, we hinged on SCA techniques and proposed a low complexity fast algorithm with provable convergence to locally optimal solutions of the nonconvex problem. Quite interestingly, every limit point of the iterate was proved to be integer in the assignment variables.

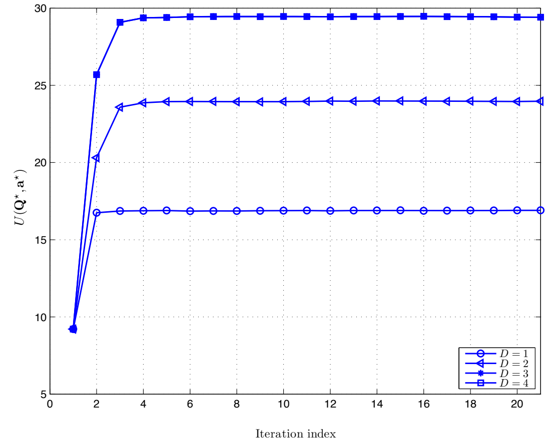

Fig. 1. Optimal sum rate vs. the iteration index, for different values of D.

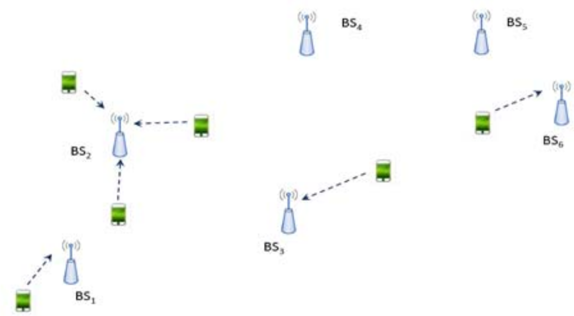

Fig. 2. Users-BS assignment, for $D = 3, 4$.

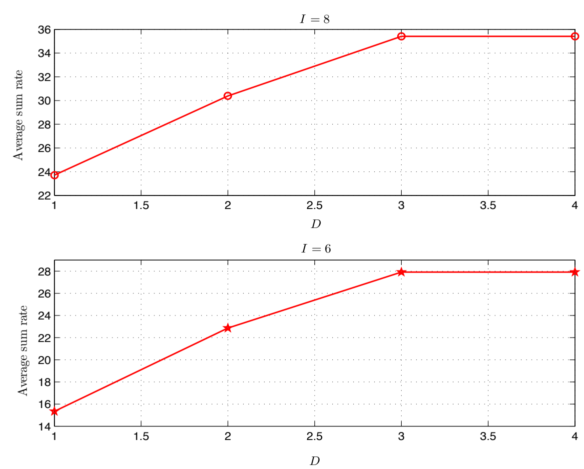

Fig. 3. Average sum rate vs. D.

6. REFERENCES

[1] J. Hoydis, M. Kobayashi, and M. Debbah, "Green small-cell networks," *IEEE Veh. Tech. Mag.*, vol. 6, no. 1, pp. 37–43, Mar. 2011.

[2] S.V. Hanly, "An algorithm for combined cell-site selection and

power control to maximize cellular spread spectrum capacity," *IEEE J. Select. Areas Commun.*, vol. 13, no. 7, pp. 1332–1340, Sep. 1995.

[3] F. Rashid-Farrokhi, L. Tassiulas, and K.J.R. Liu, "Joint optimal power control and beamforming in wireless networks using antenna arrays," *IEEE Trans. Commun.*, vol. 46, no. 10, pp. 1313-1324, Oct. 1998.

[4] M. Bengtsson, "Jointly optimal downlink beamforming and base station assignment," in *Proc. IEEE ICASSP*, pp. 2961–2964, May 2001.

[5] Y. Cheng, M. Pesavento, A. Phillip, "Joint network optimization and downlink beamforming for CoMP transmissions using mixed integer conic programming," *IEEE Trans. on Signal Process.*, pp. 3972–3987, Aug. 2013.

[6] D.H.N. Nguyen, T. Le-Ngoc, "Joint beamforming design and base-station assignment in a coordinated multicell system," *IET Commun.*, pp. 942–949, Oct. 2013.

[7] M. Hong and Z.-Q. Luo, "Joint linear precoder optimization and base station selection for an uplink MIMO network: A game theoretic approach," *Proc. of IEEE ICASSP 2012*, pp. 2941–2944.

[8] W.C. Liao, M. Hong, Z.Q. Luo, "Base station activation and linear transceiver design for utility maximization in heterogeneous networks," *Proc. of ICASSP 2013*, pp. 4419–4423; see also http://arxiv.org/abs/1309.4138.

[9] G. Scutari, F. Facchinei, P. Song, D.P. Palomar, and J.-S. Pang, "Decomposition by partial linearization: Parallel optimization of multi-agent systems," *IEEE Trans. on Signal Process.*, vol. 63, no. 3, pp. 641–656, Feb. 2014.

[10] M. Razaviyayn, M. Sanjabi, and Z.-Q. Luo, "Linear transceiver design for interference alignment: Complexity and computation," *IEEE Trans. on Inform. Theory*, vol. 58, no. 5, pp. 2896–2910, May 2012.

[11] G. Scutari, F. Facchinei, J.-S. Pang, and D.P. Palomar, "Real and Complex Monotone Communication Games," *IEEE Trans. on Inf. Theory*, to appear. [Online]. Available: `http://arxiv.org/abs/1212.6235`.

[12] S. Sardellitti, G. Scutari, S. Barbarossa, "Joint cell selection and radio resource allocation in MIMO small cell networks via successive convex approximation," to be submitted to *IEEE Trans. on Signal Process.*, Nov. 2013.

[13] A. Schrijver, *Theory of linear and integer programming.* John Wiley Sons: Wiley-Interscience, 1987.