

# Optimal User Association for Massive MIMO Empowered Ultra-Dense Wireless Networks

Antonis G. Gotsis, Stelios Stefanatos, and Angeliki Alexiou

Department of Digital Systems, University of Piraeus, Greece  
Email: {agotsis, sstefanatos, alexiou}@unipi.gr

**Abstract**—Ultra network densification and Massive MIMO are considered major 5G enablers since they promise huge capacity gains by exploiting proximity, spectral and spatial reuse benefits. Both approaches rely on increasing the number of access elements per user, either through deploying more access nodes over an area or increasing the number of antenna elements per access node. At the network-level, optimal user-association for a densely and randomly deployed network of Massive MIMO empowered access nodes must account for both channel and load conditions. In this paper we formulate this complex problem, report its computationally intractability and reformulate it to a plausible form, amenable to acquire a global optimal solution with reasonable complexity. We apply the proposed optimization model to typical ultra-dense outdoor small-cell setups and demonstrate: (i) the significant impact of optimal user-association to the achieved rate levels compared to a baseline strategy, and (ii) the optimality of alternative network access element deployment strategies.

**Index Terms**—ultra dense networks; massive MIMO; 5G; user association; optimization; integer linear programming

## I. INTRODUCTION

5G wireless networks target the provision of a sustainable solution to the radio capacity crunch predicted over the next decade [1]. Recent studies attempting to specify future network requirements state that 5G networks should support multiple orders of magnitude increased network and per user capacity levels, compared to today [2]. Although we are at the very beginning of defining a concrete 5G system concept, there are already strong indications for how the 5G enabling technologies portfolio will be formed [1], [3]–[6]. Following recent progress, we identify and focus on two key 5G evolution solutions: leftmargin=10pt

- Ultra Network Densification, namely multiple orders of magnitude higher infrastructure density levels compared to today’s macro-cell networks [7]. Ultra-Dense Networks (UDNs) will be characterized by comparable serving access nodes (small-cells) and served nodes (user devices) densities, up to the extreme point that each user is served by its own access node, exploiting the associated proximity and spectral reuse benefits.
- Massive MIMO, namely empowering today’s base stations with multiple orders of magnitude greater number (tens or even hundreds) of antenna elements compared to conventional MIMO sizes (4 or 8 antennas per base station which is the case for LTE/LTE-A), surpassing the number of active users [8]. Massive MIMO technology offers both spectral efficiency and resources reuse benefits

due to the exploitation of massively available spatial degrees of freedom.

Interestingly both technologies are built on the same design principle, that is the over-provisioning of infrastructure access elements per user. However, their focus is different, since UDNs tend to create highly localized access areas with improved channel conditions due to link distance reduction, whereas Massive MIMO leverages powerful spatial processing to create multiple interference-free user-access node links.

From a network-level perspective, the problem of selecting the best serving access node for each user, usually referred as user-association, is highly challenging in UDNs due to random topology deployment and significant load variations observed in different access nodes [9]. The problem becomes further exacerbated when Massive MIMO comes into play, since the presence of excess antenna elements per access node calls for an optimal exploitation of the available spatial degrees of freedom [10]. In this paper we formulate and solve, for the first time to the best of our knowledge, the optimal user association problem for Massive MIMO empowered homogeneous ultra-dense networks. We leverage the optimal solution to comment on the interplay of Massive MIMO and UDN technologies, and explore the performance trends as a function of access node and antenna element densities in future network setups.

### A. Related Work

The potential of infrastructure densification for achieving 5G capacity targets was recently demonstrated in [11]. Results based on simulation campaigns for indoor environments, revealed that UDN deployments corresponding to a single access node located at every room and serving at most three user devices, are able to support multiple orders of magnitude higher user rates, when combined with a moderate spectrum increase. The importance of optimizing user-association in UDNs has been stressed in the context of 3GPP small-cell enhancements study item [12], [13]. It has been demonstrated through system-level simulations that significant performance improvements could be achieved when considering both channel and load conditions during user-association decision. Nevertheless, only SISO network setups were considered and the applied user-association criteria were based on heuristic rules, rendering their optimality questionable. Interesting initial results for optimizing user-association in a limited 2x2 (2 access nodes and 2 users) SISO network setup have been recently

reported in [14]; however it is not clear how these findings could be generalized in scenarios involving more nodes and when the access nodes operate in the Massive MIMO regime.

A second line of works has recently emerged, striving to answer the “antenna element deployment dilemma”. The authors in [15], [16] explored which deployment strategy is optimal given an excess number of antenna elements over a serving area: to concentrate all elements on a single base station or to spread multiple single-antenna access nodes over the area. No single winner was identified, since conflicting performance trends were reported, depending on the considered environment type (indoor or outdoor). Differently from our work, these studies did not consider the potential of joint channel- and load-based user-association, since each user was assumed to connect to the closest access node. The most relevant work to ours is [10]. The authors first provide an approximate, yet fairly accurate expression, abstracting Massive MIMO impact on the achieved rate performance, in multi-cell setups. Then, based on this expression, they consider optimal load balancing strategies realized over a finite time horizon, for Massive MIMO HetNets.

## B. Contributions

In this paper, building upon the Massive MIMO abstraction model of [10] we:

- Formulate the problem of user-association for maximizing the guaranteed user rate across a Massive MIMO empowered ultra-dense network;
- Show that problem is computationally intractable for practical network cluster sizes and demonstrate that it admits an exact reformulation to a plausible form, which allow us to acquire the global optimal solution with reasonable complexity;
- Apply the transformed optimization model to acquire the optimal associations and achieved performance levels for a representative future network setup, considered by 3GPP;
- Leverage the acquired system-level results to quantify the impact of densification and Massive MIMO size enhancements to future networks, as well as explore the optimal strategy for deploying a given amount of antenna elements across the network, namely *concentrate* elements to fewer access nodes or *distribute* elements to a larger number of nodes;

The rest of this paper is organized as follows. Sec. II describes the system model, while Sec. III introduces the user-association problem, provides its mathematical formulation, and the steps for transforming it to a more plausible form. Sec. IV reports on numerical results and discuss key findings, whereas Sec. V concludes the work.

## II. SYSTEM MODEL

We consider a set of  $|\mathcal{M}| = M$  densely and randomly deployed multi-antenna access nodes (ANs) serving  $|\mathcal{K}| = K$  single-antenna users (UEs), similar to the overlay outdoor deployment scenario assumed in the context of 3GPP Release

12 small-cell densification theme [12]. We assume a single cluster of ANs for simplicity, and an over-provisioning of access elements per UE. In particular: i) we focus on the *ultra-dense network* (UDN) region, meaning that the number of ANs and UEs is comparable ( $K \approx M$ ) [17]; ii) consider operation in the *Massive MIMO* (MM) regime, meaning that each AN  $m$  is equipped with a very large number of antenna elements, let  $L$ , at least an order of magnitude greater than the number of UEs it accommodates ( $L \gg S_m$ , where  $S_m$  is the number of UEs served by AN  $m$ ), such that all UEs are served simultaneously over the same time-frequency resource unit [8]. It is also assumed that each UE is served by a single AN and no inter-AN coordination in the data plane is supported.

For a homogeneous network deployment scenario, it is reasonable to assume a common number of antenna elements per AN,  $L$ , and a common transmission power level per AN,  $p$ . We define the channel state information (CSI) factors for all possible UE-AN combinations, including the large-scale distance-based path-gain normalized to a given noise density level ( $g_{km}$ ) and small-scale rayleigh fading. Assuming user  $k^*$  served by AN  $m^*$ , then according to the recent work of [10], the achieved spectral efficiency (or simply rate) is given by  $r = \log_2(1 + \gamma_{k^*m^*})$ , where  $\gamma_{k^*m^*}$  is the *effective* SINR when the system is operating in the Massive MIMO region. The latter is well approximated by [10]:

$$\gamma_{k^*m^*} = \left( \frac{L - S_{m^*} + 1}{S_{m^*}} \right) \cdot \frac{p \cdot g_{k^*m^*}}{1 + \sum_{m \neq m^*} p \cdot g_{k^*m}}. \quad (1)$$

We report the major relevant observations of [10] for convenience: first the achieved SINR and equivalently the spectral efficiency or simply rate, does not depend on the small-scale channel fading (notice that in (1) only the large-scale factors are present), and second, each UE’s rate depends on the number of UEs accommodated by its serving AN. Different from [10], we assume a constant activity factor for each UE, namely all UEs inside each cluster are served simultaneously. Note that (1) is composed of:

- the factor  $p \cdot g_{k^*m^*} / \left( 1 + \sum_{m \neq m^*} p \cdot g_{k^*m} \right)$ , which is equivalent to the standard SINR formula for single-antenna networks, when accounting only for large-scale channel effects;
- the factor  $(L - S_{m^*} + 1)/S_{m^*}$ , which abstracts the Massive MIMO gain, compared to a reference conventional single-antenna setup.

In this paper, we search for identifying the optimal user-to-AN association decision which maximizes the worst rate level across all UEs, assuming that the network CSI information is perfectly known. The selected optimization criterion provides in essence the minimum guaranteed performance level, ensuring a fair handling of the various UEs, contrary to a sum-rate maximization criterion, which could resort to highly unbalanced achieved rates on a UE basis<sup>1</sup>.

<sup>1</sup>Note that one of the key requirements for 5G is the provision of uniform quality-of-service levels (see for example [2]).

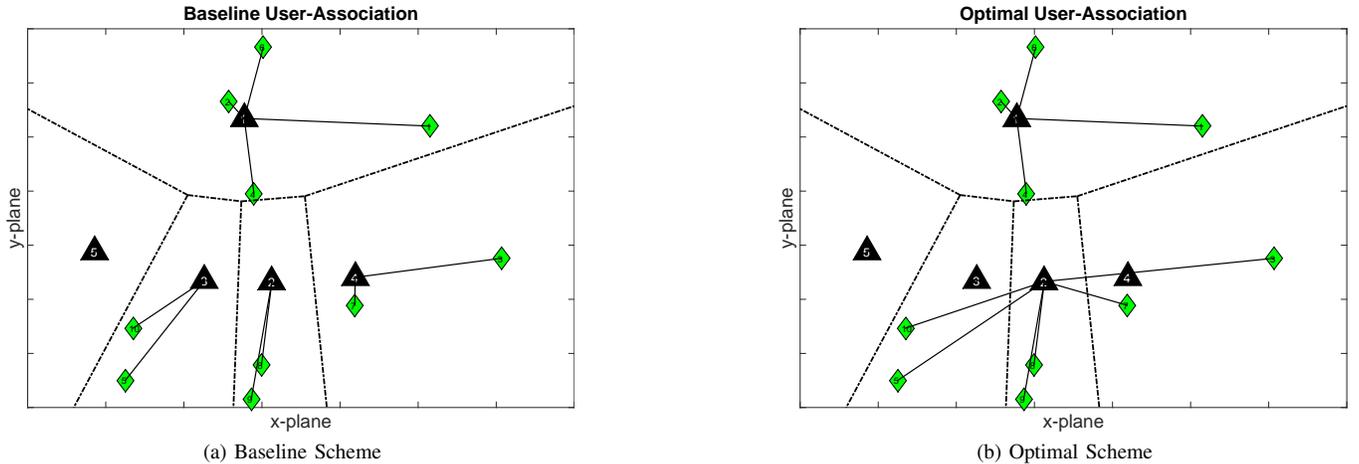


Fig. 1. A user association example for a random topology system snapshot involving a single cluster of 5 access nodes (each AN represented by a  $\triangle$ ) and 10 users (each UE represented by a  $\diamond$ ). Dashed lines depict the coverage limits for each AN considering only geographical information. Solid lines represent the user association decisions. Fig. 1a corresponds to the baseline user-association scheme, where each UE is served by its closest AN, providing the strongest received useful power. Fig. 1b corresponds to the optimal user-association decision. Notice that compared to the baseline scheme, the optimal scheme concentrates users served by ANs 2,3 and 4 to a single AN (2), since deactivating ANs 3 and 4 reduces network interference. For this example, optimal user-association enhances rate performance at a lower access node utilization level.

### III. USER ASSOCIATION IN MASSIVE MIMO UDNS: PROBLEM FORMULATION AND OPTIMAL SOLUTION

#### A. Motivation - Problem Statement

In traditional cellular networks the density of UEs is usually multiple orders of magnitude greater than that of ANs, and each UE is typically associated with the AN providing the best channel conditions (in terms of received useful power). This results in having non-idle ANs, balanced traffic loads offered per AN, and a uniform interference landscape. In Massive MIMO empowered UDNs (MM-UDNs) the following differentiating factors are noted: leftmargin=10pt

- The densities of ANs and UEs are comparable, hence multiple potentially serving ANs lie at the vicinity of each UE, with good channel conditions;
- The traffic load distribution among different ANs is highly unbalanced, opening up opportunities for dynamic load coordination among neighboring ANs;
- The interference landscape depends heavily on user association decisions; for example one could de-activate an AN serving a limited number of UEs (see for example in [13]), shift these UEs to a neighbor already active AN, thus reducing network interference.
- The excess number of available (spatial) degrees of freedom which could be utilized for enhancing useful received power and/or multiplexing more users on a per-AN basis, further perplex the user-association optimization decision making.

Based on the above observations we argue that in future networks, user association should be intelligently configured on a network basis, striking a balance between the following trade-offs: leftmargin=10pt

- *Maximizing useful received signal power* (by selecting the most proximal AN) vs *Maximizing received SINR*

taking into account the capability to also control network interference. An example is shown in Fig. 1, where UEs served by ANs 3 and 4 according to the baseline association scheme (Fig.1a), should be shifted to the farther AN-2 (allowing ANs 3 and 4 to be de-activated), according to the optimal association scheme (Fig.1b).

- *Load balancing* (by distributing the UE population to multiple ANs) vs *Load concentration* (by shifting UEs to a subset of (or even a single) Active AN), towards exploiting the spatial and spectral degrees of freedom, and in essence optimally tuning the Massive MIMO gain factor given in the effective SINR formula (1). For example in the configuration of Fig.1b it is not optimal to shift all UEs to a single-AN, since the Massive MIMO gain will become too limited.

#### B. Problem Formulation

We introduce: leftmargin=10pt

- a set of binary variables  $\{\alpha_{km}\}_{k \in \mathcal{K}, m \in \mathcal{M}} \in \{0, 1\}$ , reflecting the potential UE-to-AN associations, where  $\alpha_{km} = 1$  if UE  $k$  is associated with AN  $m$ ;
- a set of binary variables  $\{\rho_m\}_{m \in \mathcal{M}} \in \{0, 1\}$ , indicating the activity of each AN, where  $\rho_m = 1$  if AN  $m$  is active.

An AN is considered active if at least one UE is associated with it, else is turned off to avoid creating unnecessary interference. The following constraint expressions clearly hold:

$$\sum_m \alpha_{km} = 1, \forall k, \quad (2a) \quad \alpha_{km} \leq \rho_m, \forall k, \quad (2b)$$

where (2a) guarantees that each UE should be associated with a single AN, and (2b) that for an inactive AN, no UE could be associated with it. Utilizing the user-association variables, the number of UEs served by an AN  $m$  is simply given by

$S_m = \sum_{i \in \mathcal{K}} \alpha_{im}$ . Introducing the decision variables to (1), after a simple manipulation we may express the achieved effective SINR for an arbitrary UE-AN combination as:

$$\gamma_{km} = \frac{g_{km} \cdot (L+1) - g_{km} \cdot \sum_i \alpha_{im}}{(1/p) \cdot \sum_i \alpha_{im} + \sum_i \sum_{m' \neq m} g_{km'} \cdot \alpha_{im'} \cdot \rho_{m'}}. \quad (3)$$

The above fractional expression reveals the impact of different parameters on the achievable per-UE performance: leftmargin=10pt

- The numerator is optimized by assigning each UE to its most proximal AN, so as to maximize the experienced path-gain  $g_{km}$ , given that  $L \gg \sum_i \alpha_{im}$  holds in the Massive MIMO regime;
- The denominator is optimized by minimizing the number of co-served UEs (expressed through  $\sum_i \alpha_{im}$ ) and the number of interfering ANs (indicated by the  $\rho_{m'}$  variables that equal unity).

We aim at finding the user-association decision which corresponds to the maximum effective SINR guaranteed for all UEs (equivalently the max-min SINR). This corresponds to solving the following optimization problem:

$$\left\{ \begin{array}{l} \text{maximize } \theta \\ \theta, \{\alpha_{km}\}, \{\rho_m\} \\ \text{subject to } \gamma_{km} \geq \alpha_{km} \cdot \theta, \forall k, \forall m, (2a), (2b), (3) \end{array} \right\} \quad (4)$$

where  $\theta$  is the maximum guaranteed effective SINR. The constraint expression  $\gamma_{km} \geq \alpha_{km} \cdot \theta$  is interpreted as follows: if the UE  $k$  is served by the AN  $m$ , namely  $\alpha_{km} = 1$ , then the achieved SINR should be at least equal to the guaranteed common level  $\theta$ , namely  $\gamma_{km} \geq \theta$ , whereas if not, namely  $\alpha_{km} = 0$ , it is irrelevant, namely  $\gamma_{km} \geq 0$ .

With regard to the characterization of the optimization problem stated in (4) we observe that: leftmargin=10pt

- Both discrete (binary) and continuous variables are involved;
- The achieved SINR given in (3) corresponds to a (non-linear) fractional expression of binary (numerator) and products of binary (denominator) variables;
- The minimum SINR constraint expression involves the (non-linear) product of a binary and a continuous variable.

Therefore the problem in (4) is a mixed integer non-linear program (MINLP), which even for small dimensions, is extremely difficult to solve while providing global optimality guarantees [18].

### C. Reformulation towards acquiring the optimal solution

To cope with the intractability of formulation (4), we will demonstrate that the problem admits an exact linearization-reformulation, allowing us to transform it to a mixed integer linear program (MILP). The advantage of MILPs is that there exist powerful methods and tools for acquiring their global optimal solutions, at least for practical cluster sizes involving one to two tens of ANs and UEs.

Notice that the constraints in (2a) and (2b) are already linear with respect to  $\{\alpha_{km}\}$  and  $\{\rho_m\}$  variables. Next, by incorporating (3) into (4) we get:

$$\frac{g_{km} \cdot (L+1) - g_{km} \cdot \sum_i \alpha_{im}}{(1/p) \cdot \sum_i \alpha_{im} + \sum_i \sum_{m' \neq m} g_{km'} \cdot \alpha_{im'} \cdot \rho_{m'}} \geq \alpha_{km} \cdot \theta, \forall k, \forall m. \quad (5)$$

Although (5) involves multiple non-linear terms, we will show how this expression could be exactly linearized. Before introducing the exact reformulation procedure, we present two common reformulation techniques, answered in related mathematical programming literature, which deal with products involving two binary variables or a binary and a continuous variable respectively [19], [20]. The validity of these techniques is proven by simple inspection.

**Linearization Technique 1.** A product of two binary variables  $x$  and  $y$  can be replaced by a new auxiliary binary variable  $z = xy$ , along with a set of three linear constraint expressions:  $z \leq x$ ,  $z \leq y$ , and  $z \geq x + y - 1$ .

**Linearization Technique 2.** A product of a binary variable  $x$  and a continuous positive variable  $y$  can be replaced by a new continuous auxiliary variable  $z = xy$ , along with a set of four linear constraint expressions:  $y - z \leq K_y(1 - x)$ ,  $z \leq y$ ,  $z \leq K_y x$ , and  $z \geq 0$ , where  $K_y$  is large number guaranteed to be greater than the maximum value that  $y$  could take.

In what follows we will show how each product term appearing in (5) can be linearized, using the above techniques.

a) *Step 1:* With respect to the left-hand side of (5), the only non-linear term is the double-summation expression of the denominator, which involves products of two binary variables. Following Technique 1, we introduce the set of auxiliary binary variables  $\{z_{imj}\}_{i \in \mathcal{K}, m \in \mathcal{M}, j \in \mathcal{M} \setminus \{m\}}$ , through which we can replace the non-linear terms  $\sum_i \sum_{m' \neq m} g_{km'} \cdot \alpha_{im'} \cdot \rho_{m'}$  with the linear terms  $\sum_i \sum_{j \neq m} g_{kj} \cdot z_{imj}$  along with the following set of linear constraint expressions:

$$\begin{aligned} z_{imj} &\leq \alpha_{im}, z_{imj} \leq \rho_j, z_{imj} \geq \alpha_{im} + \rho_j - 1, \\ \forall i \in \mathcal{K}, m \in \mathcal{M}, j \in \mathcal{M} \setminus \{m\}, z_{imj} &\in \{0, 1\} \end{aligned} \quad (6)$$

Incorporating variables  $z_{imj}$  to (5), and applying a simple rearrangement of terms, we get:

$$\begin{aligned} g_{km} \cdot (L+1) - g_{km} \cdot \sum_i \alpha_{im} &\geq \left( (1/p) \cdot \sum_i \alpha_{im} \cdot \alpha_{km} \right) \cdot \theta \\ &+ \left( \sum_i \sum_{j \neq m} g_{kj} \cdot z_{imj} \cdot \alpha_{km} \right) \cdot \theta \end{aligned} \quad (7)$$

b) *Step 2:* With respect to the first term of the right-hand side of (7) we first linearize the user-association binary variable products, leveraging Technique 1 as above. We introduce the auxiliary binary variables  $\{v_{imk}\}_{i \in \mathcal{K}, m \in \mathcal{M}, k \in \mathcal{K}}$  and the set of linear constraint expressions in (8) to replace each

$\alpha_{im} \cdot \alpha_{km}$  product with a single  $v_{imk}$  term.

$$\begin{aligned} v_{imk} &\leq \alpha_{im}, v_{imk} \leq \alpha_{km}, v_{imk} \geq \alpha_{im} + \alpha_{km} - 1, \\ i &\in \mathcal{K}, m \in \mathcal{M}, k \in \mathcal{K}, v_{imk} \in \{0, 1\} \end{aligned} \quad (8)$$

c) *Step 3*: Likewise, the products  $z_{imj} \cdot \alpha_{km}$  appearing in the second term of the right-hand side of (7), are linearized with the introduction of binary auxiliary variables  $\{u_{imjk}\}_{i \in \mathcal{K}, m \in \mathcal{M}, j \in \mathcal{M} \setminus \{m\}, k \in \mathcal{K}}$  and the constraint expressions set (9):

$$\begin{aligned} u_{imjk} &\leq z_{imj}, u_{imjk} \leq \alpha_{km}, u_{imjk} \geq z_{imj} + \alpha_{km} - 1, \\ i &\in \mathcal{K}, m \in \mathcal{M}, j \in \mathcal{M} \setminus \{m\}, k \in \mathcal{K}, u_{imjk} \in \{0, 1\} \end{aligned} \quad (9)$$

Based on linearization steps 2 and 3, (7) can be (exactly) rewritten as:

$$\begin{aligned} g_{km} \cdot (L + 1) - g_{km} \cdot \sum_i \alpha_{im} &\geq \\ \left( (1/p) \cdot \sum_i v_{imk} \cdot \theta \right) &+ \left( \sum_i \sum_{j \neq m} u_{imjk} \cdot \theta \right) \end{aligned} \quad (10)$$

For a known target SINR level  $\theta$ , (10) is linear with respect to the initial and new auxiliary variables, hence problem (4) has been transformed to a feasibility MILP. When a target SINR is not specified (as in our study), it is required to optimize over the SINR as well, and locate the maximum guaranteed SINR level for all UEs. We observe that the non-linear terms in (10) involve products of a binary variable ( $v_{imk}$  or  $u_{imjk}$ ) and a bounded continuous variable (the SINR level  $\theta$ ), which can be linearized according to Technique 2. The procedure is described in detail below.

d) *Step 4*: Let  $\mathcal{Q}$  a large positive number which is at least equal to the maximum expected achieved SINR level. We introduce two sets of (continuous) auxiliary variables  $w_{imk}$  and  $n_{imjk}$  to replace the  $v_{imk} \cdot \theta$  and  $u_{imjk} \cdot \theta$  product terms respectively, along with the necessary linear constraint expression sets, (11) and (12):

$$\begin{aligned} w_{imk} &\leq \mathcal{Q}v_{imk}, w_{imk} \leq \theta, w_{imk} \leq \theta - (1 - v_{imk}) \mathcal{Q}, \\ w_{imk} &\geq 0, i \in \mathcal{K}, m \in \mathcal{M}, k \in \mathcal{K} \end{aligned} \quad (11)$$

$$\begin{aligned} n_{imjk} &\leq \mathcal{Q}u_{imjk}, n_{imk} \leq \theta, n_{imjk} \leq \theta - (1 - u_{imjk}) \mathcal{Q}, \\ n_{imjk} &\geq 0, i \in \mathcal{K}, m \in \mathcal{M}, j \in \mathcal{M} \setminus \{m\}, k \in \mathcal{K} \end{aligned} \quad (12)$$

Combining all the reformulation steps, the original MINLP formulation converts to the following mixed integer linear form:

$$\left\{ \begin{array}{l} \text{maximize } \theta \text{ s.t. (2a),(2b),(6),(8),(9),(11),(12),} \\ g_{km} \left( L + 1 - \sum_i \alpha_{im} \right) \geq \frac{1}{p} \sum_i w_{imk} + \sum_i \sum_{j \neq m} n_{imjk}, \\ \forall i \in \mathcal{K}, m \in \mathcal{M}, j \in \mathcal{M} \setminus \{m\}, k \in \mathcal{K} \end{array} \right\} \quad (13)$$

This problem can be solved leveraging powerful state-of-the-art ILP optimization tools, such as GUROBI, MOSEK or CPLEX, to acquire the optimal guaranteed SINR level and the user-association decision that realizes it. We remark that

every step applied throughout this section provides an exact transformation of the original problem variables, and no approximation/relaxation penalty is induced. The price we have to pay is the increase in the number of variables and constraint expressions, caused by the introduction of auxiliary variables ( $\{z, v, u, w, n\}$ ) and necessary expressions linking the original and the auxiliary variables. However the elimination of non-linearity is more critical, since today's optimization tools are mature enough to handle large-scale MILPs in typical desktop computing platforms<sup>2</sup>

## IV. RESULTS & DISCUSSION

### A. Simulation Setup

In this section we present the results of simulation experiments conducted with a two-fold perspective: `leftmargin=10pt`

- To highlight the benefits of joint and optimal load- and channel-based user-association in MM-UDNs, with respect to both achieved worse rate performance levels and the number of active ANs, for various infrastructure densification levels and Massive MIMO sizes;
- To explore the optimal antenna element deployment in MM-UDNs, considering strategies that either concentrate elements on less ANs or distribute them in higher number of ANs, given a fixed overall antenna element budget.

We follow the simulation guidelines used within the 3GPP small-cell study item and particularly we consider the outdoor small-cell cluster scenario 2a [12, A.1.2]. Figure 2 represents a snapshot generated based on settings reported in [12]. We extend the scenario by empowering small-cell ANs with massive number of antenna elements. We consider a fixed number of 10 UEs and vary the number of ANs from 2 to 10. Regarding power allocation, we assume a total cluster power budget which: i) is equally distributed to the deployed ANs (either active or inactive); ii) is independent from the number of ANs, hence with increased AN density the power level per AN ( $p$ ) is reduced accordingly; iii) its absolute level is configured such that a specific spatially averaged interference-free SNR level is targeted. Two user-association strategies are considered, as discussed in Sec.III-A and Fig.1: `leftmargin=10pt`

- The baseline channel-based approach, which associates each UE with the AN providing the maximum received signal power, and
- The optimal channel- and load-based approach proposed in this paper. For obtaining the optimal user-association decision for each system snapshot, we formulate the MILP problem in (13) using the CVX modeling framework [22], and use the state-of-the-art GUROBI package to solve it [21].

The results to be reported correspond to averaging over 200 independent system snapshots.

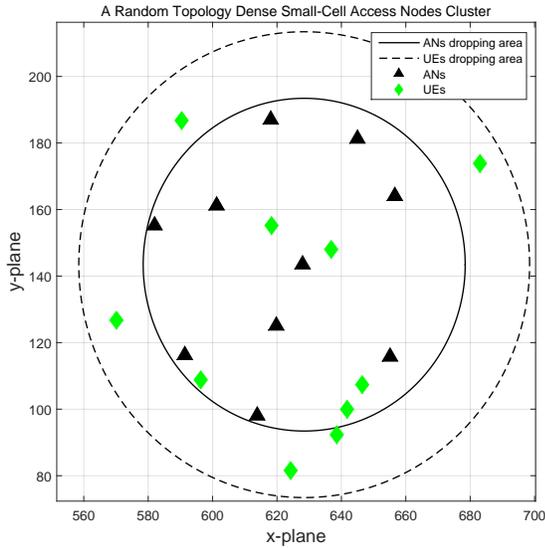


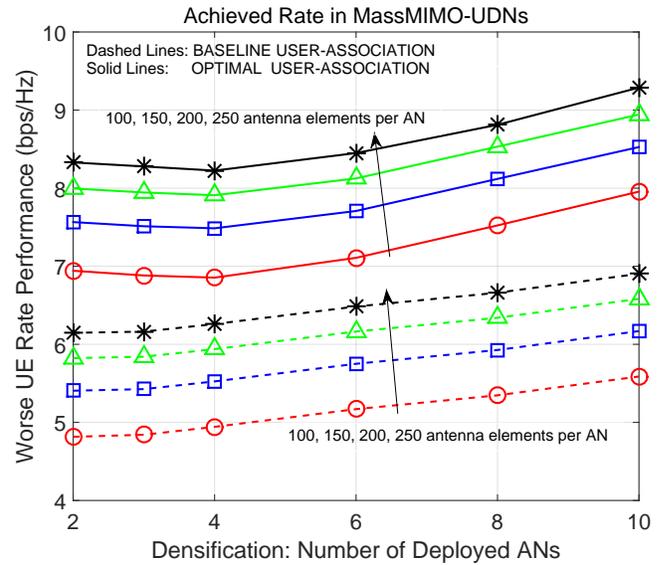
Fig. 2. A small-cell cluster snapshot formed by 10 ANs and 10 UEs. Each cluster is dropped in a random hot-zone of an overlay macro-cell. We ignore inter-cluster interference. Within a single cluster ANs are dropped uniformly over a circular area with radius 50m and UEs over a co-centered area with radius 70m. ANs-UEs large-scale path-gains follow the ITU-UMi model as in [12, A.1.2]. Carrier frequency is set to 3.5 GHz and noise level to -174 dBm/Hz.

## B. Results

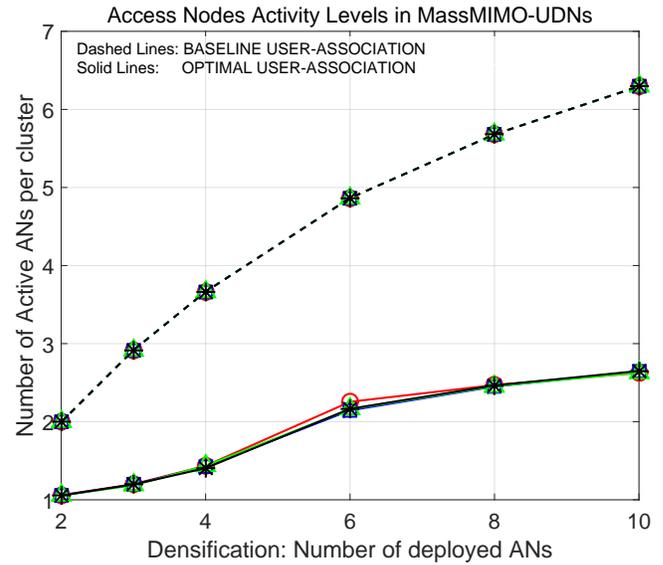
The first set of experiments demonstrates the importance of optimal user-association. An average target SNR of 30 dB is assumed. We compare the achieved worse UE rate performance for different infrastructure densification (number of deployed ANs within the considered small-cluster) and massive MIMO sizes (number of antenna elements per AN) for both user-association schemes. Figure 3a shows that on average, 33-40% rate performance gains are achieved by optimally selecting the serving AN per UE. The maximum gain is observed for the smallest MIMO size (100 elements). On top of the rate gains, Figure 3b shows that the improved performance levels are achieved with a limited number of active ANs, on average with 50% less active ANs than that of the baseline scheme. This means that load “concentration” is a preferable strategy for the particular MM-UDN setups. We also observe that the benefit from increasing MM size is more prominent than that of infrastructure densification. As an example, for an MM size of 200 elements per AN, we observe a 12.5% rate gain for a 2.5x infrastructure density increase (going from 4 to 10 ANs), whereas for the 6 ANs scenario, we observe a 20% rate gain for a 2.5x MM size increase (going from 100 to 250 elements per AN).

In the second set of experiments we consider a fixed amount of antenna elements (500) and study if it is preferable to concentrate these elements on sparse or dense AN deployments (resembling the macro vs small-cell network evolution

<sup>2</sup>As an example we report that a typical cluster instance involving 5 ANs and 10 UEs is solved in less than 5 secs using a laptop equipped with Intel Core i7-4710MQ CPU @ 2.50 GHz, using the GUROBI [21] solver called through CVX [22] in the MATLAB environment.



(a) Worse UE Rate



(b) Number of Active ANs

Fig. 3. The impact of network infrastructure densification (number of deployed ANs) and massive MIMO sizes (number of antenna elements per AN) for optimal and baseline user-association.

dilemma). We apply the optimal user-association scheme of (13) and examine the rate performance levels as well as the number of active ANs supporting these levels for different element deployment strategies, namely having 2, 4, 6 or 10 ANs per cluster and 250, 125, 83 or 50 elements per AN respectively. Various transmission power levels are assumed corresponding to target SNR of 30, 20, and 10 dB. Results are presented in Figure 4. It is deduced that the optimal strategy depends on the power budget. For high power availability (target SNR set to 30 dB) we observe a clear preference for concentrating antenna elements to a single or few ANs. This is justified by the fact that elements concentration allows to de-activate more ANs, reducing network interference levels,

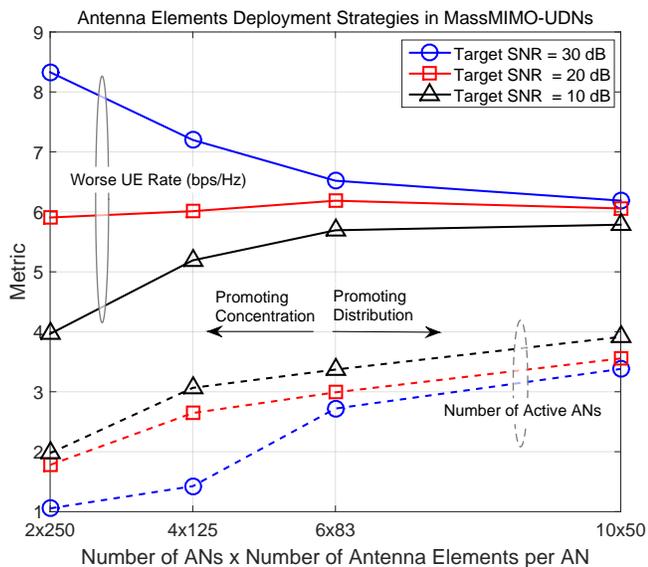


Fig. 4. Impact of antenna elements deployment strategies to optimal worse UE rate and number of active ANs.

while the ANs are allocated enough power to support non-proximal links. In other words, proximity benefits offered by activating more ANs are not enough to counterbalance the excess interference “leakage”. For limited power budgets (target SNR set to 10 dB) the distribution policy is optimal, since proximity benefits are critical for dealing with limited received signal levels. Regarding the number of active ANs supporting the optimal rates, we observe that it decreases as the target SNR is higher, since the optimal network operation point moves from the noise-limited to the interference-limited regime.

## V. CONCLUSION

Increasing the number of access elements per user is regarded as a major evolution path towards the next (5th) generation of wireless networks. This could be achieved by heavy densification of access nodes (Ultra-Dense Networks) and/or heavy increase in MIMO dimensions (Massive MIMO). Within this framework, we examined the problem of optimal channel- and load-based user-association. We reformulated the original computationally-intractable problem to a more plausible form, allowing us to acquire the optimal association decisions with reasonable complexity. We applied the optimization model to typical dense small-cell setups and demonstrated that optimal user-association significantly enhances rate performance levels and offers additional benefits by reducing the number of active access nodes. In addition we explored alternative strategies for optimally deploying the massively available number of access elements to ultra-dense networks.

## VI. ACKNOWLEDGEMENT

This work has been performed in the context of the ART-COMP PE7(396) “Advanced Radio Access Techniques for Next Generation Cellular Networks: The Multi-Site Coordination Paradigm” research project, implemented within the

framework of Operational Program Education and Lifelong learning, co-financed by the ESF and the Greek State.

## REFERENCES

- [1] NGMN Alliance, “5G White Paper (Executive Version),” <http://www.ngmn.org/>, Dec. 2014.
- [2] ICT-317669-METIS, “Project Deliverable, D1.1: Scenarios, requirements and KPIs for 5G mobile and wireless system,” <http://www.metis2020.com/>, Apr. 2013.
- [3] —, “Project Deliverable, D6.2: Initial report on horizontal topics, first results and 5G system concept,” Apr. 2014.
- [4] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, “What Will 5G Be?” *Selected Areas in Communications, IEEE Journal on*, vol. 32, no. 6, pp. 1065–1082, June 2014.
- [5] Q. Li, H. Niu, A. Papathanassiou, and G. Wu, “5G Network Capacity: Key Elements and Technologies,” *Vehicular Technology Magazine, IEEE*, vol. 9, no. 1, pp. 71–78, March 2014.
- [6] ICT-258512-EXALTED, “D2.4: The EXALTED system concept and its performance,” <http://www.ict-exalted.eu>, Feb. 2013.
- [7] I. Hwang, B. Song, and S. Soliman, “A Holistic View on Hyper-Dense Heterogeneous and Small Cell Networks,” *Communications Magazine, IEEE*, vol. 51, no. 6, 2013.
- [8] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, “Massive MIMO for Next Generation Wireless Systems,” *Communications Magazine, IEEE*, vol. 52, no. 2, pp. 186–195, February 2014.
- [9] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, “Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective,” *Wireless Communications, IEEE*, vol. 21, no. 3, pp. 118–127, June 2014.
- [10] D. Bethanabhotla, O. Bursalioglu, H. Papadopoulos, and G. Caire, “User Association and Load Balancing for Cellular Massive MIMO,” in *Information Theory and Applications Workshop (ITA), 2014*, Feb 2014.
- [11] ICT-317669-METIS, “Project Deliverable, D6.3: Intermediate System Evaluation Results,” <http://www.metis2020.com/>, Aug. 2014.
- [12] 3GPP TR 36.872 V12.1.0, “Small cell enhancements for E-UTRA and E-UTRAN – Physical layer aspects (Release 12),” Dec. 2013.
- [13] 3GPP, “R1-133246: Cell association and load balancing/shifting,” Barcelona, Spain, Aug 2013. [Online]. Available: <http://goo.gl/EKK6wd>
- [14] J. Wildman, Y. Osmanlioglu, S. Weber, and A. Shokoufandeh, “Delay Minimizing User Association in Cellular Networks via Hierarchically Well-Separated Trees,” in *IEEE International Conference on Communications (ICC’15)*, 2015, accepted for publication. [Online]. Available: <http://arxiv.org/abs/1501.02419>
- [15] S. Dierks, M. B. Amin, W. Zirwas, M. Haardt, and B. Panzer, “The Benefit of Cooperation in the Context of Massive MIMO,” in *OFDM 2014; 18th International OFDM Workshop 2014 (InOw’14); Proceedings of*, Aug 2014.
- [16] K. Hosseini, W. Yu, and R. Adve, “Large-Scale MIMO Versus Network MIMO for Multicell Interference Mitigation,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 8, no. 5, pp. 930–941, Oct 2014.
- [17] H. Tullberg, Z. Li, A. Hoglund, P. Fertl, D. Gozalvez-Serrano, K. Pawlak, P. Popovski, G. Mange, and O. Bulakci, “Towards the METIS 5G concept: First view on Horizontal Topics concepts,” in *Networks and Communications (EuCNC), 2014 European Conference on*, June 2014.
- [18] J. Lee and S. Leyffer, Eds., *Mixed Integer Nonlinear Programming*, ser. The IMA Volumes in Mathematics and its Applications, Vol. 154. Springer, 2012.
- [19] D.-S. Chen, R. Batson, and Y. Dang, *Applied Integer Programming - Modeling and Solution*. John Wiley & Sons, 2010.
- [20] T.-H. Wu, “A note on a global approach for general 0–1 fractional programming,” *European Journal of Operational Research*, vol. 101, no. 1, pp. 220–223, 1997.
- [21] GUROBI Optimization, “State-of-the-Art Mathematical Programming Solver, v5.6.2,” <http://www.gurobi.com/>, Oct. 2014.
- [22] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.1,” <http://cvxr.com/cvx>, Oct. 2014.