Improving Network Performance via Optimization-Based Centralized Coordination of LTE-A Cells

Giovanni Nardini[(1)], Giovanni Stea[(1)], Antonio Virdis[(1)], Marco Caretti[(2)], Dario Sabella[(2)]

(1) Dipartimento di Ingegneria dell'Informazione, University of Pisa, Italy
{g.nardini, g.stea, a.virdis}@iet.unipi.it

(2) Telecom Italia, Turin, Italy – {dario.sabella, marco.caretti}@telecomitalia.it

***Abstract*—This paper shows how to improve the overall network performance (cell throughput, fairness, and energy efficiency) via centralized coordination of LTE-A cells. We first present optimization models for small-scale coordination (i.e., three cells). Then, we show that extending the same solution to a higher number of cells is generally unfeasible, due to both an unfeasible amount of reporting on the UE side, and too high computational requirements. To overcome this limitation we then propose a layered solution which i) relies on small-scale coordination at the first level (e.g., three cells at the same site), and ii) coordinates groups of coordinated cells at a higher scale (i.e., tens of cells), using optimization models, reaping the benefits of a centralized architecture. We show through packet-level simulations that our scheme brings significant benefits, in terms of fairness, throughput, and energy efficiency.**

***Index Terms*—LTE-A, C-RAN, Cell Coordination**

I. INTRODUCTION

THE ever-increasing trend towards higher user bandwidth in LTE cellular networks [1] finds a natural opponent in inter-cell interference. Coordinating neighboring cells, so as to reduce the interference suffered by each UE, is also the key to achieving higher SINRs, hence higher throughput, better energy efficiency for the same throughput, and higher fairness for cell-edge UEs. Coordinated Scheduling (CS) is a CoMP (Coordinated Multi-Point Transmission and Reception) technique that allows several eNodeBs (eNBs) to coordinate service to a set of UEs: by deciding who addresses whom and using what Resource Blocks (RB), pairs of cell-UEs transmissions can be scheduled concurrently with a tolerable increase in interference. Moreover this kind of resource utilization could lead to an improvement of the energy efficiency of LTE networks.

Cells can indeed be coordinated using a distributed approach, whereby each cell contributes by running independent algorithms. This approach is hampered by the partial information on the state of the network possessed at each cell, and requires a considerable amount of computational power to be distributed at each site, with obvious reconfigurability and management problems. Recently, cloud-based architectures, such as C-RAN [2], have been proposed as a solution to obtain centralized control of a (possibly large) number of cells and can be seen as technology enabler of CoMP algorithms, especially at large scales. The idea is to use a central cloud (implemented by means of an IT platform) to run more computationally-intensive algorithms, relying on global information, thus achieving better performance than a distributed approach. Moreover, such an architecture reduces the amount of exchanged information, due to the hub-and-spoke topology, with respect to an eNB mesh.

In this paper we focus on centralized coordination based on cloud architectures, with the aim to analyze the feasibility of CS mechanisms in the central entity. For that purpose we first show that optimal solutions of CS problems *cannot* scale to a large number of cells, because the *amount of information* that the UEs are requested to report in order to achieve it gets unfeasible when the number of cells exceeds few units (e.g., three). We then present in this paper a solution for large-scale centralized coordination, whose starting point is small-scale coordination (e.g., a set of three neighboring cells). We thus show how to formulate and optimally solve this small-scale problem. Such optimal solution can act as a benchmark for any possible (small-scale) heuristic coordination algorithm. Building on the above, we show that larger-scale coordination (e.g., several tens of cells) can be achieved, with a computational power available in a cloud-based core, once small-scale coordination is in place, and discuss the overhead and performance.

As far as related work is concerned, [3] advocates coordinating the precoding matrices among neighboring eNBs, and shows that this increases the throughput of both the whole cell and cell-edge UEs. In [4], each cell indicates the highest interferers for its UEs, and a central entity makes a long-term resource sharing plan, so that mutually interfering cells never share resources. Such decision is only made on the number of UEs in each interference zone (the actual traffic is not considered), and is overly conservative (it is sufficient that only one UE of A perceives interference from B in order to constrain A and B to use mutually disjoint resources). A dynamic scheme is proposed in [5]: each cell allocates its frame to its UEs according to its own scheduler, assuming certain muting conditions. It then sends the frame and the associated per-RB muting requests to a central controller that arbitrates requests on a per-RB basis. This scheme requires more information to transit between cell and controller, and does not find globally optimal solutions.

In the rest of the paper, Section II describes the system model. Section III presents the coordination algorithms at both a small and a large scale. We evaluate the performance in Section IV, and report conclusions in Section V.

II. SYSTEM MODEL

We assume that the network is deployed on a 2D map as in Figure 1. Each hexagon has three antennas facing inwards (hence hosts three cells), which may be co-located with those of neighboring hexagons. A number of UEs is deployed in the hexagons. They are statically associated to one of the (inward-facing) cells, but they can measure the level

of interference perceived by each of the other two, so that they can report to the serving eNB the Channel Quality Indicator (CQI) when either or both the other two are muted. Each eNB is connected to a central cloud via high-bandwidth, low-latency links, which we assume to be ideal. We denote with A, B, C the three cells in a hexagon, and if the cell index is a value taken by variable x, then $x = A \Rightarrow (x+1) = B,\ (x-1) = C$. Let $N(x)$ be the number of UEs associated to cell x. UEs can be identified by couple x,j, where $1 \le j \le N(x)$. Second, given a cell x, we use two superscript symbols to denote the interference from the other two cells. The first symbol identifies cell x-1, whereas the second is for cell x+1. Symbol "+" means "active", and "-" is for "inactive". This way, $CQI^t_{x,j}$, where $t \in T = \{++, -+, +-, --\}$, denotes the four possible CQIs for a UE j associated to cell x: $CQI^{++}_{x,j}$ is the one reported when both x-1 and x+1 are active, etc. Set T represents the four *Interference Subbands (ISs)* for a UE. Our objective is to allocate resources effectively in the above settings.

III. Coordination

A. Small-scale coordination

Small-scale coordination is achieved by a small number of cells K (e.g., three). Let $s^t_{x,j}$ be the number of RBs given to UE x,j within IS t. Let $Q_{x,j}$ be that UE's buffer and let $r_{x,j}$ denote the (one and only) Transport Block Size (TBS) format that UE x,j will be scheduled with. We denote with $b^t_{x,j}$ a binary variable that is equal to 1 if UE x,j has a RB within IS t, and zero otherwise. Let M be the size of the frame (i.e, the number of RBs at each cell). Finally, let R be a large positive constant. A max-throughput problem can then be formulated as follows:

$$\max \sum_{x\in\{A,B,C\}} \sum_{j=1}^{N(x)} r_{x,j} \cdot \left(\sum_{t\in T} s^t_{x,j} - p_{x,j} \right)$$

$$s.t.$$

$$r_{x,j} \cdot \sum_{t\in T} s^t_{x,j} \le Q_{x,j} + p_{x,j} \qquad \forall x,j \quad (i)$$

$$r_{x,j} \le CQI^t_{x,j} + R \cdot \left(1 - b^t_{x,j}\right) \qquad \forall x,j,t \quad (ii)$$

$$b^t_{x,j} \le s^t_{x,j} \le M \cdot b^t_{x,j} \qquad \forall x,j,t \quad (iii)$$

$$p_{x,j} \le CQI^t_{x,j} + R \cdot \left(1 - b^t_{x,j}\right) - 1 \qquad \forall x,j,t \quad (iv)$$

$$\sum_{t\in T} \sum_{j=1}^{N(x)} s^t_{x,j} + \sum_{j=1}^{N(x-1)} s^{--}_{x-1,j} + \sum_{j=1}^{N(x+1)} s^{--}_{x+1,j} + \max\left\{ \sum_{j=1}^{N(x-1)} s^{-+}_{x-1,j}, \sum_{j=1}^{N(x+1)} s^{+-}_{x+1,j} \right\} \le M \qquad \forall x \quad (v)$$

$$\sum_x \left(\sum_{j=1}^{N(x)} s^t_{x,j} + \max\left\{ \sum_{j=1}^{N(x-1)} s^{-+}_{x-1,j}, \sum_{j=1}^{N(x+1)} s^{+-}_{x+1,j} \right\} \right) \le M \qquad (vi)$$

$$b^t_{x,j} \in \{0,1\}\ \forall x,j,t \qquad (vii)$$

$$r_{x,j}, p_{x,j} \in \mathbb{Z}^+\ \forall x,j, \quad s^t_{x,j} \in \mathbb{Z}^+\ \forall x,j,t \qquad (viii)$$

The objective function states that the total throughput should be maximized. Every UE x,j has a rate, and that rate is multiplied by all the RBs that are allocated to that UE, whatever the interference condition. $p_{x,j}$ denotes the padding, not to be counted as useful bits.

Constraint (*i*) states that each UE cannot transmit more than its queue's worth of bytes, including possible padding bits. Padding is necessary, otherwise queues may never be emptied. Constraint (*ii*) states that the rate cannot exceed the worst CQI among all the interference conditions. For instance, if a UE is allocated RBs with no interference ($b^{--}_{x,j} = 1$) and with interference from both cells ($b^{++}_{x,j} = 1$), it will use the smallest CQI, i.e. $r_{x,j} = CQI^{++}_{x,j}$. Note that, since R is a large constant, constraint (*ii*) is inactive if $b^t_{x,j} = 0$, hence those CQIs do not count as a limit. Constraint (*iii*) states that $s^t_{x,j} = 0$ if $b^t_{x,j} = 0$, and $s^t_{x,j} \ge 1$ if $b^t_{x,j} = 1$, thus ensuring consistency. Constraint (*iv*) states that a UE gets less than one RB's worth of padding. Constraint (*v*) states that the number of blocks M in a cell's frame accounts for all the RBs allocated to UEs of that cell x,j, whatever their ISs t, plus all the RBs allocated by other cells, which request cell x to be silent in those RBs. This last term can be further split into two: first, the RBs where the other cells require exclusive transmission (i.e., those two with a $--$ superscript). Second, RBs where other cells require only x to be muted (i.e., those in the *max* bracket). These last need not be disjoint. For instance, the RBs allocated by cell A should include all the RBs where:

- A's UEs transmit;
- B requests *both* A *and* C not to transmit;
- C requests *both* A *and* B not to transmit;
- B requests A not to transmit (whereas C may transmit);
- C requests A not to transmit (whereas B may transmit).

The last two terms can overlap, thus we take their maximum instead of their sum. Figure 2 shows an example of frame structure for three cells A, B, C. Constraint (*vi*) describes the fact that the groups of RBs where muting of one or two cells apply must occupy the same positions in the three frames.

The above one is a mixed integer-nonlinear problem (MINLP), with a size of $O\left(K \cdot N \cdot |T|\right) = O\left(K \cdot N \cdot 2^K\right)$ variables. Non-linearity comes from the product in both the objective function and constraint (*i*), whereas the *max* operator in constraints (*vi-vii*) can easily be linearized. The number of constraints is also $O\left(K \cdot N \cdot 2^K\right)$. MINLPs are NP-hard in general. The structure of this one is indeed similar to that of a multi-band-CQI scheduling (i.e. one where a MaxC/I allocation has to be made on per-subband CQIs), sub-bands being replaced by the ISs, with the added complexity that

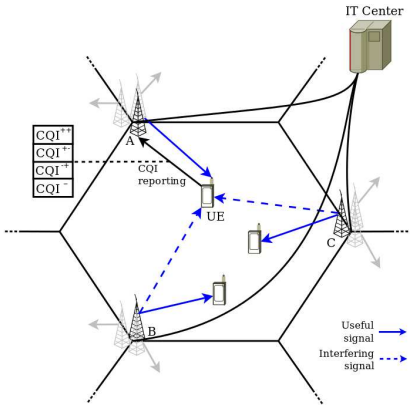

Figure 1 - Network architecture

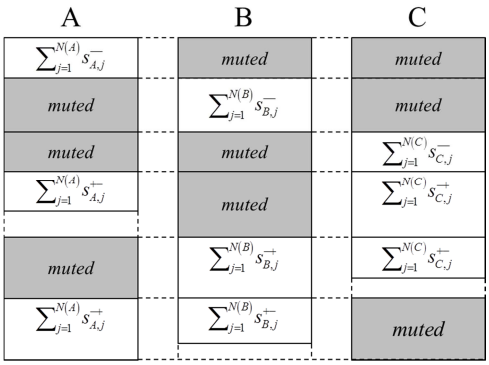

Figure 2 – Frame structure for three coordinated cells

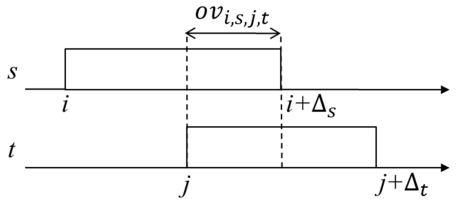

Figure 3 – Overlapping of two ISs *s* and *t*.

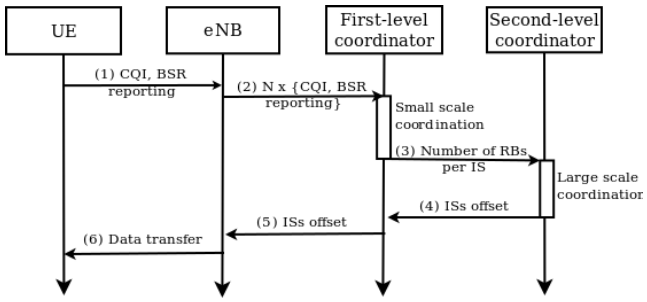

Figure 4 – Information exchange

resources are not partitioned statically (i.e., in advance), but dynamically (i.e., as a result), as per constraints (*vi*). Since the former problem has been proven to be NP-hard, this one is very likely to be the same as well. In any case, it is hardly feasible to attempt to solve it in a TTI's time, even for a small number of UEs (i.e., 10-20). The fact that the reporting information required increases exponentially with the number of coordinated cells clearly indicates that this model cannot scale to larger dimensions.

This problem *can* however be reformulated as a mixed-integer-*linear* problem (MIP), through a careful reformulation (omitted due to lack of space) at the price of increasing the number of variables to $O\left(2^{2^k}\right)$ ($4 \cdot 10^9$ variables for 6 cells). Such reformulation allows one to take advantage of considerably faster MIP solvers (e.g., CPLEX, [6]).

The solution to the small-scale coordination problem yields an allocation vector for the three cells (i.e., the set of $s^t_{x,j}$ values). From the latter, the dimension of each IS *d* of a cell, call it Δ_d, can be easily obtained. Note that ISs can be arranged in several ways in the cells frames (provided that mutual exclusion constraints are met) without affecting optimality. We exploit this degree of freedom to achieve larger-scale coordination. Before moving on to that, we remark that the above model accommodates different objectives. For instance, a Coordinated Proportional Fair (CPF) could be achieved by simply substituting the objective with:

$$\max \sum_{x \in \{A,B,C\}} \sum_{j=1}^{N(x)} r_{x,j} \cdot \left(\sum_{t \in T} s^t_{x,j} - p_{x,j} \right) \Big/ R_{x,j} \,, \qquad (1)$$

where $R_{x,j}$ is the long-term PF rate achieved by UE (*x,j*).

B. Large-scale coordination

As we can only afford to coordinate (very) few cells using the above modeling, we exploit *layering* to achieve larger-scale coordination: the first layer coordinates triples of cells as explained above, targeting UEs in the most appropriate way given their (electromagnetic) position; the second layer coordinates neighboring *triples*, up to a reasonable number (e.g. ten). Both layers are implemented in the cloud.

For each couple of cells *x,y* belonging to triples *i* and *j* respectively, we define an Interference Coefficient (IC) $\alpha_{x,y}$, which measures the average interference that *x*'s UEs will suffer from cell *y*. An Interference Graph (IG) can thus be constructed, whose nodes are cells and whose edges are the couples *x,y* such that $\alpha_{x,y} > 0$. Note that this also allows one to set $\alpha_{x,y} = 0$ when the IC is below a threshold, to simplify the problem (which depends on how connected the IG is). Moreover, ICs need not be symmetric, i.e., with anisotropic cells.

For each couple of ISs *s,t*, the overall interference is:

$$\beta_{s,t} = \sum_{\substack{x \text{ active in } s, \\ y \text{ active in } t}} \left(\alpha_{x,y} \right)$$

Call $ov_{i,s,j,t}$ the number of overlapping RBs between IS*s s, t*, if allocated starting from RB *i* and *j,* as shown in Figure 3. We pre-compute values $\gamma_{i,s,j,t} = ov_{i,s,j,t} \cdot \beta_{s,t}$. Furthermore, call T_n the set of ISs belonging to triple *n* and define variables $x_{i,s} \in \{0,1\}$, where $x_{i,s} = 1$ means that IS *s* begins at the *i*-th RB. The problem formulation is thus the following:

$$\min \sum_{i,j \in [1,\ldots,M]} \sum_{s,t \in S} \left(\gamma_{i,s,j,t} \cdot x_{i,s} \cdot x_{i,t} \right)$$

s.t.

$$\sum_{i=1}^{M} x_{i,s} = 1 \qquad \forall s \qquad (i)$$

$$\sum_{i=M-\Delta_s}^{M} x_{i,s} = 0 \qquad \forall s \qquad (ii)$$

$$\left(\sum_{j=\max\{1,i-\Delta_t\}}^{\min\{M,i+\Delta_s\}} \sum_{t \in T_n} x_{j,t} \right) \le M \cdot \left(1 - x_{i,s}\right) \qquad \forall i, \forall n, \forall s, t \in T_n \qquad (iii)$$

$$x_{i,s} \in \{0,1\} \qquad \forall i, \forall s \qquad (iv)$$

Constraints (*i-ii*) state that IS *s* begins once, early enough in the frame for it to fit in. Constraint (*iii*) enforces mutual exclusion among ISs of the same triple. In fact, when $x_{i,s} = 1$, all $x_{j,t}$, for each *j* between $i - \Delta_t$ and $i + \Delta_s$, must be zero.

The above problem is a QSAP (Quadratic Semi-Assignment Problem). This kind of problems is NP-hard and, in general, difficult to solve optimally even at small scales. However, CPLEX can solve these problems, also capitalizing on the fact that many $\gamma_{i,s,j,t}$ are null.

IV. Performance Evaluation

We now show some performance metrics related to our proposal. We first show the benefit of our approach with respect to the amount of information exchanged and the associated latency, and then profile the large-scale optimization models. Finally, we show that our centralized coordination improves throughput, energy efficiency and fairness using detailed packet-level simulation.

A. Information exchange

In Figure 4, the information exchange required in a traditional distributed architecture is reported. Messages (3-4) have to be transmitted through the X2 interface. The (one-way) latency introduced by the X2 depends on the backhaul technology and can vary from few milliseconds, when fiber optic is used, to 60 ms, when DSL access is used. This delay can be removed if the algorithm is executed in a C-RAN architecture, i.e. when first- and second-level coordination are co-located into a single IT center. Furthermore, in a cloud environment, we can take advantage from the fact that both first- and second-level coordination can exploit hardware resources from all the virtual eNodeBs. Thus, the processing time would also be reduced. The overall gain in terms of latency results in improved coordination, as the algorithms can exploit fresher reporting (e.g., CQI).

B. *Profiling of the large-scale optimization model*

The solution time of the large-scale model is affected by the number of variables and constraints. In Figure 5 and 6, the number of binary variables and memory requirements are reported as a function of the number of RBs (25, 50 and 100) and triples (ranging from 7 to 19), respectively. Memory usage has been estimated following guidelines in [7], which yield a lower bound on the actual required memory. The number of constraints is of the same order as variables, thus it is not reported. The figures show that the size of the problem does not grow too fast with the number of triples, which is promising. Figure 7-8 report the evolution of the objective function, when seven triples are coordinated. CPLEX is run on a machine with 8 Intel Core I7 CPUs at 2.80 GHz, 8 GB of memory and Ubuntu 12.10 OS. The figures show that the number of UEs hardly affects the convergence time, whereas the number of RBs does: in fact, in Figure 8, the downward step with 50 RBs takes place around t=180s. i.e., much later than with 25 RBs.

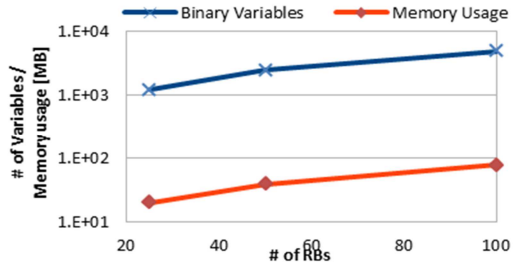

Figure 5 – Memory usage and binary vars against the # of RBs, 7 triples

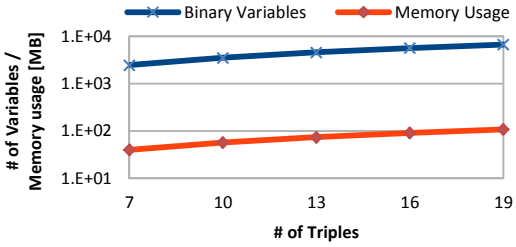

Figure 6 – Memory usage and binary vars against the # of triples, 50 RBs

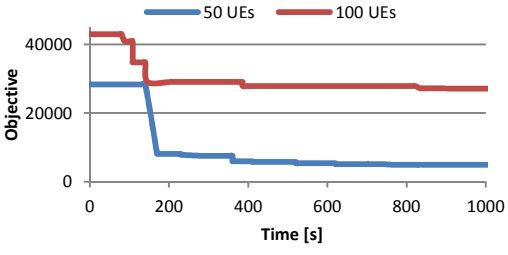

Figure 7 – Evolution of the objective, 7 triples and 50 RBs per cell

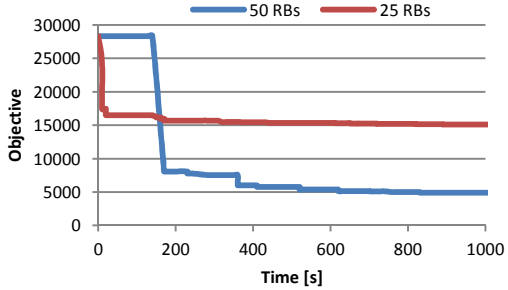

Figure 8 – Evolution of the objective, 7 triples and 50 UEs per cell

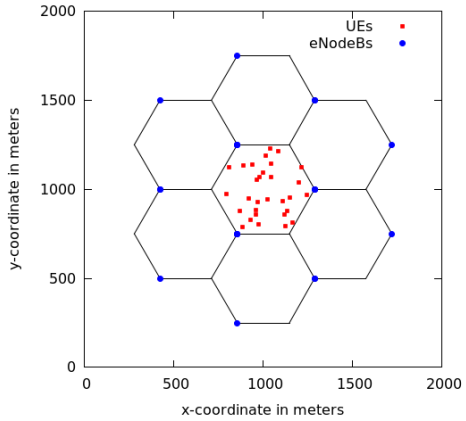

Figure 9 – Simulation scenario

Table 1 – Parameters used in the packet-level simulation

Parameter	**Value**
Number of RBs	50
Pathloss model	Urban Macro
eNB Tx power	46 dBm
eNB P_{base}	260 W
eNB P_{off}	150 W
eNB ρ	3.76 W
Inter-site distance	500 m
Traffic model	Periodic, 20ms inter-arrival time
Packet size	40 Bytes
Scheduling	Proportional Fair
Number of UEs	[75, 150, 225]

C. *Packet-level simulation*

We now report packet-level simulation results. The evaluation has been carried out using SimuLTE [8], a packet-level simulator based on OMNeT++ [9]. Figure 9 illustrates the simulation scenario, which consists of 7 hexagonal areas, each of them having three cells located on three vertices and radiating inward (each blue dot thus represents three anisotropic radiating antennas). UEs are randomly deployed only in the central hexagon, whereas surrounding triples only produce interference on UEs served by the three cells of the central hexagon. Interference is produced assuming that the neighboring hexagons have a similar load as the central one. Only downlink traffic is simulated. The power model, taken from [10], is an affine function of the number of RBs, i.e., $P = P_{base} + \rho \cdot n$, where P_{base} is the *baseline* power, and $n \leq M$ is the number of allocated RBs. Idle eNBs consume a power $P_{off} < P_{base}$. Table 1 lists all the scenario parameters. First, we evaluate a single-hexagon case, where small-scale coordination is executed in the central hexagon, while surrounding triples produce constant interference. Figure 10 reports the results of the above scenario with 75 UEs and shows that coordination between three cells in the same hexagon yields considerable benefits in terms of both cell throughput and depleted power.

We then compare the large-scale coordination algorithm proposed in the previous section against two baseline scenarios: one, called B1, where no coordination takes place (each cell allocates resources independently). Another, called B2, where small-scale coordination is independently executed within the three cells of each hexagon. In the latter,

no large-scale coordination occurs, hence ISs of different hexagons are unaligned.

Figure 11-13 show the cumulative distribution function of user throughput. A significant gain at the 5-th percentile is clearly noticeable when large-scale coordination is employed, resulting in better service for cell-edge UEs. Comparing our solution to B2, there is little difference for cell-center UEs. This can be explained with the fact that their major interferers are eNBs radiating in the same hexagon, thus their performance is weakly affected by inter-hexagon coordination. When the load approaches saturation (225 UEs) coordination mechanisms are ineffectual, since there is no way to avoid overlap of ISs. The improvement over B1 is still relevant, however.

In Figure 14-15, the overall throughput and power consumption of the central hexagon are reported. The results show that centralized coordination is beneficial from the overall system point of view as well.

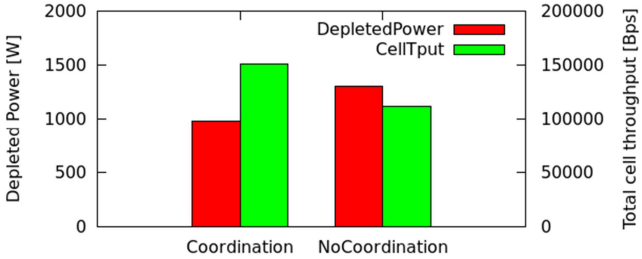

Figure 10 - Single-triple scenario

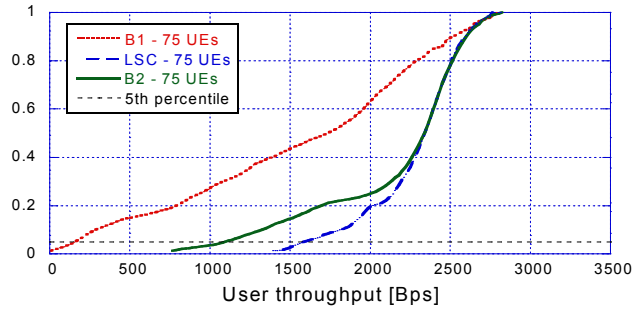

Figure 11 - CDF of UE throughput, 75 UEs

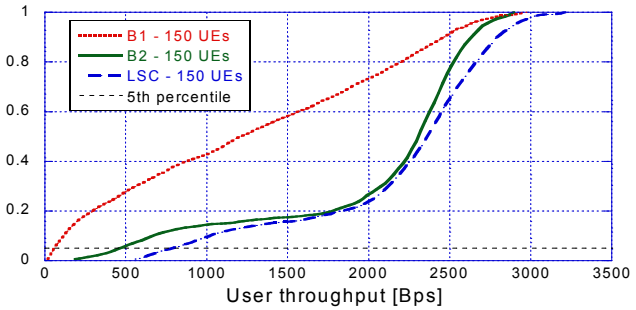

Figure 12 - CDF of UE throughput, 150 UEs

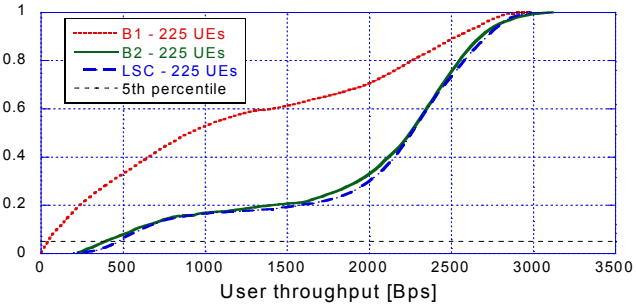

Figure 13 - CDF of UE throughput, 225 UEs

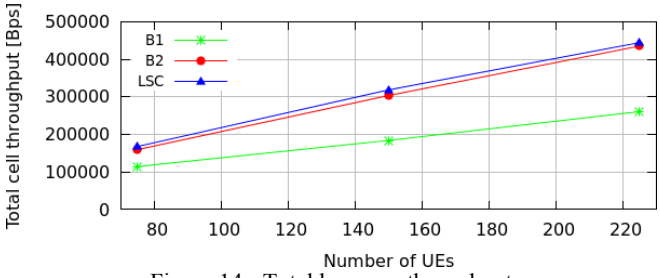

Figure 14 - Total hexagon throughput

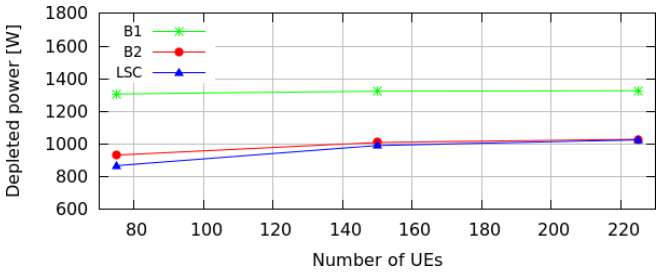

Figure 15 - Power consumption

V. Conclusions and Future Work

This paper presented models for centralized cell coordination at both small and large scales. We showed that a layered solution, whereby small-scale coordination is a first building block of a large-scale scheme, achieves significant benefits as far as fairness is concerned, with additional benefits in terms of cell throughput and energy efficiency. Future work will include designing fast heuristics for both small- and large-scale coordination.

VI. Acknowledgements

The subject matter of this paper includes description of results of a joint research project carried out by Telecom Italia and the University of Pisa. Telecom Italia reserves all proprietary rights in any process, procedure, algorithm, article of manufacture, or other result of said project herein described. Authors would like to thank prof. Antonio Frangioni of the University of Pisa for his useful suggestions.

References

[1] 3GPP - TS 36.300 Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2.

[2] C-RAN – The Road Towards Green RAN, v. 2.5 white paper, China Mobile Research Institute, Oct. 2011

[3] J. Zhang, J.-C. Yu, and J. L. L. Liu, "Intercell Interference Coordination through Limited Feedback," International Journal of Digital Multimedia Broadcasting, 2010

[4] M. Y. Arslan, J. Yoon, K. Sundaresan, S. V. Krishnamurthy, and S. Banerjee, "A Resource Management System for Interference Mitigation in Enterprise OFDMA Femtocells,"IEEE/ACM Transactions on Networking, Nov. 2012

[5] M. Rahman and H. Yanikomeroglu, "Enhancing cell-edge performance: a downlink dynamic interference avoidance scheme with inter-cell coordination," IEEE Transactions on Wireless Communications, vol. 9, pp. 1414-1425, April 2010

[6] ILOG CPLEX Software, http://www.ilog.com

[7] Guidelines for estimating CPLEX memory requirements based on problem size, CPLEX website: http://tinyurl.com/o66dfoc

[8] SimuLTE webpage. http://www.simulte.com.

[9] OMNeT++, http://www.omnetpp.org

[10] EARTH EU project website, https://www.ict-earth.eu/