Coordinated Downlink Multi-Point Communications in Heterogeneous Cellular Networks

A. Barbieri, P. Gaal, S. Geirhofer, T. Ji, D. Malladi, Y. Wei, and F. Xue
Corporate R&D, Qualcomm Inc., San Diego, CA.
Emails: {alanb, stefang, pgaal, tji, dmalladi, ywei, fxue}@qualcomm.com

***Abstract*—In this paper we assess how coordination among base stations can be exploited to improve downlink capacity in fourth generation (4G) cellular networks. We focus on heterogeneous networks where low-power pico cells are deployed within the coverage area of an existing macro network with the aim of offloading traffic from the (potentially congested) macro cells to low-power cells. Firstly, we describe an enhanced inter-cell interference coordination scheme which is shown to achieve a significant capacity gain in such deployments by leveraging a loose coordination among neighbor base stations. Secondly, we explore how a tighter coordination among base stations can be exploited to further improve the network capacity. Even though the schemes described in this paper apply to long term evolution (LTE) wireless networks, we point out that most of the findings and conclusions we draw apply to any cellular network.**

I. Introduction

Mobile data traffic demand has been increasing unrelently in the last few years because of the widespread adoption of smartphones and the increasing usage of data-intense mobile applications. For example, AT&T reported a 50x increase in the US mobile data traffic between 2006 and 2009 [1]. Such striking increase does not seem to slow down soon, in fact according to recent forecasts data traffic demand is projected to double every 12 months for the next few years [2], leading to a stunning 1000x increase in capacity demand in 10 years. To cope with such capacity crunch issue, which is a big concern for mobile operators around the world, four possible approaches have been identified:

- **Increase the spectrum allocated to cellular networks**. Although, from a technical perspective, this is a relatively simple way to increase capacity, spectrum is a scarce resource and thus licenses have become increasingly expensive. Furthermore, spectrum fragmentation around the globe and the cost of multi-band radio frequency (RF) transceivers significantly limit the total spectrum that could be utilized in a cellular network.
- **Advanced physical layer TX and RX techniques**. Increasingly complex communication techniques have been studied and in some cases adopted for future broadband communication systems, so as to squeeze more bits per second for a given bandwidth. Examples include techniques to exploit multiple antennas according to the multiple-input multiple-output (MIMO) paradigm, increased constellation densities, etc. Even though such techniques have been shown to effectively improve spectral efficiency, the relative gains are typically minor and the chances to extract large capacity gains from pure PHY layer techniques in future generation broadband wireless networks are considered small [3], [4][1].
- **More offloading to other radio technologies**. The coexistence of multiple radio networks covering the same locations, on either licensed or unlicensed spectrum, allows in principle to offload traffic from congested networks to lightly loaded networks. The widespread adoption of WiFi, and specifically the deployment of access points in high-density areas (hotspots), allowed to dramatically improve the experience of users allowed to access the WiFi network in those areas. Nevertheless, field observations have shown that cellular data demand *increased* after WiFi offloading due to the improved user experience.
- **Increasing cell density**. Area spectral efficiency of a cellular network can be increased by increasing the cell density and shrinking the cells' footprint [5]. Though, in the dense deployments typically found in highly populated urban areas today, cell splitting gains achievable by adding further macro cells can be substantially limited by the inter-cell interference [6]. Furthermore, deploying macro base stations in a dense urban environment may be prohibitively expensive. Hence, embedding low-power nodes into existing macro-cells-based networks so as to obtain a so called heterogeneous network (HetNet), has emerged as a viable and cost-effective way to increase network capacity [6].

In this paper we focus on the latter approach, with particular emphasis on networks composed of a mix of macro and pico base stations (BS)[2]. It has been recently shown that introducing pico nodes within an existing macro cellular network provides both coverage and capacity improvements by offloading users from the macro network to a pico cell, whenever possible [6]. Though, because of the reduced footprint of the pico base cells, the amount of users which can be offloaded is limited in most scenarios of practical interest [7]. In [8] it has been shown

[1] An interesting exception is the case where the number of TX and RX antennas can be significantly increased. For mobile devices with size constraints, this is only suitable at higher frequency band, which is only usable for small cells due to physical propagation limitation

[2] The two main differences between a macro BS and a pico BS are the transmission power, which is typically 10-20dB lower in pico BS [6], and the antenna height and gain. Furthermore, pico BSs may have reduced equipment size, reduced power consumption (leading to a reduced OPEX), and sometimes a reduced set of features (e.g., fewer number of supported concurrent users).

that a significant network capacity boost can be achieved by increasing the cell coverage of pico cells, an approach known as cell range expansion (CRE), as long as the resulting inter-cell interference problem is dealt with suitably. In particular, [6] and [9] proposed an enhanced inter-cell interference coordination (eICIC) method which, by means of a loose coordination among neighbor macro and pico base stations, and through suitable improvements at the user equipment (UE) side, allows to achieve significant capacity gains in practical network deployments.

In case of eICIC only loose coordination among macros and picos is needed, which is advantageous from a deployment perspective. At the same time, it poses the question whether tighter coordination among cells may lead to further performance improvement. The focus of this paper is to provide some insights into this question by considering different forms of coordinated multipoint transmission (CoMP), with special emphasis on the potential benefits of such schemes on the downlink capacity of co-channel heterogeneous networks.

CoMP has been an active area of research, both in academia as well as in industry. For example, an ongoing work item in 3GPP targets support of CoMP in future releases of LTE. At the same time, various forms of CoMP may be differentiated. One important example is the case of coordinated scheduling/beamforming (CS/CB), in which the coordination among cells strives to align scheduling and beam decisions such as to minimize interference to UEs scheduled on the same time/frequency resources. On a high-level such schemes can achieve two different kind of gains. First, coordinating scheduling decisions, i.e., which specific UEs are selected for transmission, can help alleviate strong interference conditions. In addition, a judicious selection of beamforming weights may further contribute to minimizing interference. Ideally, scheduling and beam selection should be carried out jointly to optimize performance gains.

While CS/CB strives to align scheduling decisions, it assumes that a UE's serving cell remains fixed. Some forms of CoMP have relaxed this assumption by allowing a dynamic switching of the serving cell (dynamic cell selection or DCS) or by allowing multiple cells to serve the same UE simultaneously (joint transmission or JT). While both DCS and JT may have potential for larger gain, they also present increased implementation and standardization complexity. For example, from a backhaul perspective both DCS and JT require that the data intended for a specific UE is available at all cells that may potentially be engaged in data transmission to the UE. This represents an important consideration as it increases backhaul traffic proportionally with the number of participating cells. Furthermore, in order to select which cells should participate in the transmission to a specific UE, increased channel state feedback is required. This impacts system overhead and UE implementation complexity.

In this paper, we present a detailed system design for a CS/CB-based CoMP scheme. The choice of CS/CB is rooted in the observation that from a practical perspective, CS/CB strikes an attractive tradeoff between performance and complexity [10].

Outline of the paper

In this paper, we provide an overview of the state-of-the-art inter-cell coordination techniques that achieves significant capacity gains in heterogeneous network. Furthermore, we propose a detailed downlink CoMP system design for an LTE-based heterogeneous network and we show its performance, obtained through computer simulations, under realistic assumptions.

This paper is organized as follows: Section II describes the system and network models while Section III gives an overview of a state-of-the-art inter-cell coordination scheme for heterogeneous LTE networks. Section IV explores state-of-the-art advanced coordination techniques (CoMP) and in Section V the detailed design of a novel CoMP scheme is described, with performance results shown in Section VI. Finally, Section VII concludes the paper.

II. System model

We consider a heterogeneous LTE network deployed according to the models described in [11]. 19 macro eNBs are dropped following a hexagonal layout with an inter-site distance (ISD) of 500m, which corresponds to the D1 scenario of the 3GPP evaluation methodology. Each macro node includes three sectors with antenna boresights pointing in the three horizontal directions separated by 120 degrees. For each sector, N pico cells and M UEs are randomly dropped within the sector area. In [11, Table A.2.1.1.2-4] different criteria for the random placement of low-power nodes and UEs are specified. In this paper we focus on the so-called configuration 1 and configuration 4b, which are respectively defined as follows:

- **Config1**. Locations of low-power BSs and UEs are independent and uniformly distributed within the sectors area. 25 UEs are dropped per macro sector.
- **Config4b**. 30 UEs are dropped in clusters (namely, non-uniform distribution) within the macro sector area and each low-power BS is placed in the vicinity of a cluster of UEs (hotspot scenario).

Path loss and shadowing values are computed according to the rules described in [11].

In Section I it was mentioned that, in order to achieve the large capacity gains promised by eICIC, pico cells' footprint must be enlarged so as to increase the traffic offload from the macro network. Since transmission power cannot be increased, such cell coverage increase can be achieved by changing the handover threshold between macro and pico cells. In particular, in LTE the network can instruct a UE to trigger a reporting event when the serving cell becomes weaker than a neighbor cell plus or minus a bias, determined by the network [12]. By setting such bias to a sufficiently negative value and avoid handing a pico UE over to a macro unless the event above has been triggered, the handover from the pico to the macro is effectively delayed and the pico cell coverage increases accordingly. Fig. 1 shows the association statistics,

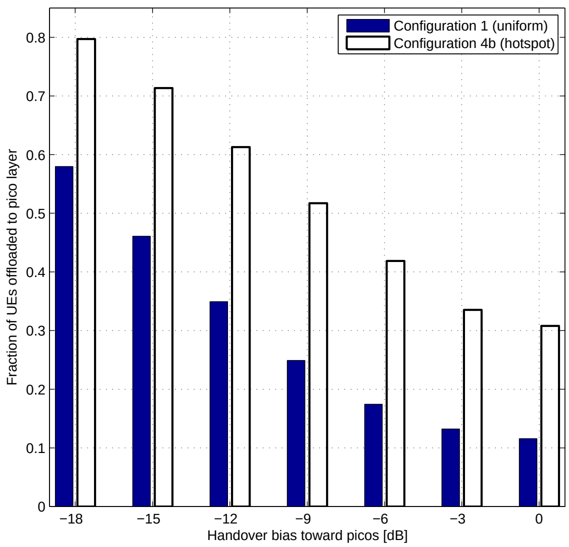

Fig. 1. Macro to pico offloading statistics as a function of the the pico handover bias, assuming 4 pico BSs per macro sectors.

as a function of the pico handover bias, for both config1 and config4b, assuming 4 picos per macro sector. These results have been obtained by assuming that a UE is associated to the strongest cell according to the reference signal received power (RSRP), of course taking into account the pico bias as well[3]. From Fig. 1 it is clear that, without any bias toward the picos, the offloading from macros to picos is not as significant, but it increases remarkably when a large enough bias is employed. Section III describes the issues stemming from a large handover bias toward picos and shows an effective way to cope with them.

We also emphasize that, without a suitable range expansion of the pico cells, the actual network performance improvements may not justify the cost of deploying and operating pico BSs. In fact, since backhaul is usually the major cost associated to the deployment of a pico BS, it has to be fully utilized to make such investment economically sound. On the other hand, for small cells without range expansion the amount of users actually offloaded may be very low, thus leading to an under-utilization of the pico BSs backhaul.

Fig. 2 shows an examplary heterogeneous network where pico BSs are deployed in the coverage area of macro BSs. This figure highlights the presence of an X2 connection between neighbor BSs [13], whose role is described in Section III, and also shows using different colors the possible UE associations, namely macro UEs, pico CRE UEs (UEs associated to a pico BS but the UEs being in the CRE area), and pico center UEs (UEs associated to a pico, where the power received from the pico by the UEs is also the strongest among all BSs).

[3]Fig. 1 assumes that all UEs are instructed to use the same handover bias toward the picos. Though, Release-8 UEs may not support bias values as large as newer UEs (see Section III), so in a real network where a mix of legacy and newer UEs is present, the actual offloading may be reduced.

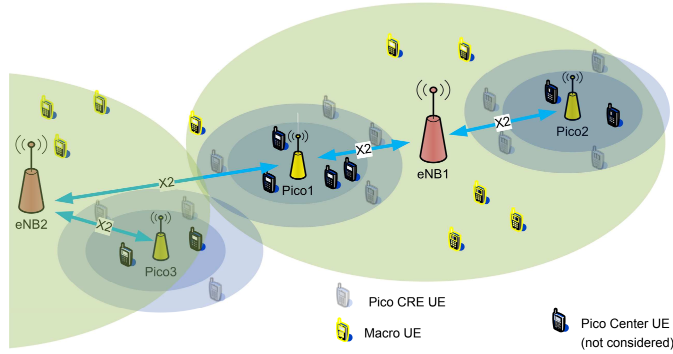

Fig. 2. Examplary heterogeneous network.

Another important consideration concerns backhaul technology. Fundamentally, two cases may be differentiated. First, for backhauls with large capacity and small delay, CoMP may target a tight coordination in which scheduling decisions, beam weights and potentially even the selection of serving cells (for DCS or JT) can be jointly determined within a CoMP cluster. From a practical perspective, this may be the case when a fiber-based backhaul is available and processing is concentrated at a single entity possibly coinciding with the macro or residing at some other place in the network. In this case, the setup effectively becomes a distributed antenna array. Second, when the backhaul has a delay that exceeds several milliseconds or when the backhaul throughput is limited, it may no longer be possible to support such a centralized architecture. In this case, scheduling and beam selection needs to be carried out at both the macro and picos separately but subject to some coordination that takes place over the backhaul. It should be appreciated that both deployments are important from a practical perspective as some operators may have fiber-based backhauls available while such technology may not be available elsewhere. We will refer to the idealized backhaul case as "RRH-based CoMP" and to the limited backhaul setup as "distributed" CoMP.

Backhaul considerations also impact standardization. For example, RRH-based CoMP in which scheduling is concentrated at a single entity does not rely on a standardized form of exchanging coordination information over the backhaul. In contrast, in a distributed CoMP system, coordination information needs to be exchanged as scheduling is performed in a distributed fashion. This may or may not entail standardizing a format for coordination messages depending on whether macros and picos are supplied by different vendors. If macros and picos come from the same vendor, a proprietary format for the message exchanges may be used.

III. An Enhanced Inter-Cell Interference Coordination Scheme

When a low-power cell range is expanded by changing the handover bias, some UEs may experience an unusually low signal-to-interference ratio (SIR). In particular, for those UEs which are close to the handover boundary (*i.e.*, in the so-called CRE area) the signal coming from at least one neighbor macro

BS may be significantly stronger than the signal from the serving pico BS. If the handover bias magnitude is larger than a few dB, which is a necessary condition to achieve significant offloading according to Fig. 1, user experience in the CRE area may be negatively impacted for the following two reasons:

i Low SIR UEs may not be able to perform basic acquisition tasks, e.g., detecting and decoding LTE primary and secondary synchronization sequences (PSS/SSS) and/or the physical broadcast channel (PBCH) [14]. If a UE could not detect a weak cell, then the strong cell does not have sufficient knowledge to offload the UE to the weak but unloaded cell.

ii Even if we assume that the UE was able to connect to the weak pico cell, its low operating SIR may prevent connected-mode data communication between the BS and the UE to happen, or such communication may be possible but with a very low throughput. When this happens, offloading from the macro layer happens at the cost of a significant user experience degradation for those UE which happen to be in a CRE area.

A set of techniques shown to effectively cope with the two issues described above have been described in [6] and [9]. It is shown in [6] that the adoption of such techniques in a heterogeneous macro-pico LTE network with cell range expansion brings significant capacity gains. For instance, by adding two picos per macro BS, a capacity improvement of about 260% can be achieved over a macro only deployment, assuming an average network load of 75% (see [6, Figure 9]). We hereby summarize the main findings and design aspects:

A. Network-side enhancements

i **Time-domain resource partitioning**. We mentioned that the SIR of CRE UEs on the data channel may be very low because of the interference from neighbor macro BSs. By preventing the macro BSs to transmit any data on a (periodically repeating) pattern of subframes it is possible to significantly improve the SIR experienced by the CRE UEs on those special subframes. Hence, Rel-10 specifications of the LTE standard introduced the concept of almost blank subframe (ABS) which is a subframe where no transmission on the physical downlink shared channel (PDSCH) is allowed from, e.g., macro BSs [14] (see Fig. 3 for an example).

ii **Inter-cell coordination**. In order to efficiently exploit the time-domain resource partitioning described above, the victim BSs (which are the picos in the considered scenario) must have perfect knowledge of the ABS patterns employed by the strongest neighbor aggressor BSs (macros). Furthermore, when deciding the amount of resources to blank, a macro BS should take into account both the benefit of the victim pico BS(s) as well as its own throughput reduction because of the dimensionality loss. In particular, average load of the involved cells may be taken into account when determining the network-wide optimal proportion of blanked resources. All the operations described above require a suitable backhaul information exchange between macro BSs and pico BSs, which in LTE Release 10 is realized through the X2 Application Protocol (X2AP) [13].

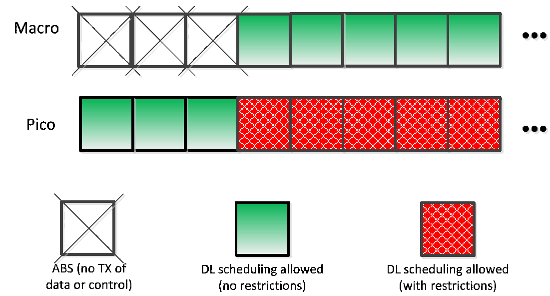

Fig. 3. Example of a time-domain resource partitioning pattern.

B. UE-side enhancements

i **Interference cancellation (IC)**. In order for a UE to connect, it needs to first acquire the cell and achieve synchronization. In LTE this requires reliable detection of the acquisition signals and the decoding of the physical broadcast channel, which carries the master information block (MIB). For this purpose all active eNBs periodically broadcast suitable control signals, including the already mentioned PSS/SSS, PBCH, and the cell reference signals (CRS). UEs in the CRE area may not be able to directly acquire the low-power BS because of large interference, thus an interference cancellation is mandatory for the acquisition of weak cells [6]. Advanced UEs deploying such advanced IC algorithm can aquire cells which are several dB weaker than the other cells, thus allowing UEs in CRE areas to acquire and maintain synchronization with the low power node.

ii **Double CQI**. Because of the adoption of a time-domain resource partitioning scheme among the BSs, channel quality may change abruptly between subframes. In particular, a pico UE affected by strong interference coming from a neighbor macro BS may experience a much better channel quality on those subframe where the macro BS refrain from PDSCH transmission (*i.e.*, ABSs). In order to improve the scheduler rate prediction accuracy, which in turn affects the achievable throughput, Rel-10 LTE specifications introduced the concept of dual channel quality indicator (CQI) [15], namely Rel-10 UEs can be instructed by the network to concurrently measure two independent CQI values, using different subframes for interference estimation.

IV. Tightening the coodination among nodes: CoMP

A. Coordination among nodes described earlier could be tightened in many different ways

CoMP aims to achieve additional gains on top of eICIC by tightening the coordination among cells. As mentioned earlier, different forms of CoMP can be differentiated:

- **CS/CB.** Coordinated scheduling/beamforming aligns scheduling decisions and/or beamforming weights across multiple cells with the objective of minimizing interference to co-scheduled UEs in the same CoMP cluster. A UE's serving cell is not changed as part of the CoMP operation.
- **DCS.** Dynamic cell selection considers changing a UE's serving cell on a per-subframe basis. The change of the serving cell may be transparent to the UE, meaning that the UE is unaware of this change. In the context of LTE this transparency is supported through UE specific reference signals.
- **JT.** Joint transmission represents the case where multiple cells are simultaneously transmitting to a single UE. Similar to DPS this operation may be fully transparent to the UE. It is important to differentiate coherent vs. non-coherent JT depending on whether a coherent combining of the signals form multiple cells is targeted at the UE. The former case requires additional phase feedback between the cells whereas the latter targets opportunistic combining at the UE.

Disregarding practical constraints, coherent JT may offer the biggest potential for performance gains. In fact, this observation has been made in the past, both in an academic context and in a more practical framework [16]. As the latter reference illustrates in detail, coherent JT conceptually enables cells to completely null interference to other co-scheduled UEs, which has the potential for large performance gains. However, such transmitter-side interference nulling is extremely sensitive to CSI imperfections and requires a large number of cells to cooperate. As [16] shows, when practical constraints are taken into account, gains deteriorate swiftly.

Non-coherent JT lacks phase information between cooperating cells and does not target interference nulling but rather a form of opportunistic combining. Performance gains due to non-coherent JT are not obvious though; the joint transmission may boost SINR conditions at the UE but at the same time leads to a dimension loss as cells could instead have scheduled separate UEs. For DCS, such dimension loss does not occur and it may be possible to exploit channel variations opportunistically; however in practice only few UEs, namely those located at the edge of two cells, may benefit thereby reducing system-level gain. Some gains due to improved load balancing may be achieved at low loads.

Motivated by the above, this paper focuses on CS/CB based CoMP schemes in which scheduling and beam selection gains can be achieved. While Section V provides a complete system design, we provide a semi-analytically study first to provide a high-level overview of the performance trends and achievable gains.

B. Discussion on expected gains of the beam-selection algorithms

In the heterogeneous networks set-up, depending on who makes decision first, CoMP schemes can be categorized into two sets. One is macro-first, in which macro cells make scheduling and MCS selection prior to pico cells, allowing the latter to optimize scheduling and beam selection accordingly. The other is pico-first, the opposite to macro-first.

In both algorithms, three types of gains could be exploited. The first is certainly the beam selection gain, due to the fact that the *later* scheduler can adjust its beams for enhancing signal strength towards its intended receiver and weakening interference towards the un-intended receiver. The second is the multi-user gain – also called the UE selection gain – coming from the fact that there are multiple UEs with each one a different channel. The third is the link adaptation gain, allowed when the channel is slowly changing. With more accurate channel feedback, the scheduler can predict the channel quality and base its scheduling decision towards improving the system throughput.

In a practical cellular environment, the significance of the three gains manifests differently due to several factors. In the macro-first case, the targeted beneficiary of coordination are the pico UEs. Consider the beam-selection gain first. Because the typical number of transmit antennas is 2 or 4, it is easy for the pico transmitters to null the macro sector's beam at a destination UE by precoding, provided that the channel feedback representing both the pico-to-UE and macro-to-UE channels is accurate. Yet, typically that is not easy to achieve. Thus receiver side processing is more significant in dealing with macro interference. If the macro sector's signal is rank one, then it is easy for the pico UE to cancel that interference using MMSE. In this sense, the beam selection gain is not pronounced much. If the macro's signal is rank two, then a UE with two receive antennas cannot attain an interference-free direction even in the high SNR case. One thus would focus on the other two types of gains in coordination. For UE selection, as pointed out in theory, a gain of $\log\log n$ scale is expected, where n is the number of UEs served by one sector. For link adaptation, in slowly changing environment, selecting the best UE to serve at each time is "water-filling" over time and among UEs. This gain hinges on how the magnitude and eigen-vectors of the channel change over time, and is more significant than the beam-selection gain. A factor that might limit the gain is the fairness issue, e.g. the scheduling must maintain certain fairness among strong and weak UEs. One example is the proportional fair scheduler, which maximizes the sum of logarithms of the achieved throughputs instead of the total throughput.

In the pico first case, the beam-selection gain is even more suppressed. This is because in this case, the picos first choose their beamforming directions independently, then the macro optimizes its precoding matrix accordingly. It would be easy for the macro to find a better beam to avoid interference to all picos if there are only one or two. However, as the number of picos grows, it becomes much harder to choose a beam good for every pico, if possible at all. The situations for the UE selection gain and link adaptation gain are similar to the macro-first scheduling.

A complication in the cellular environment is the fact that a receiver is exposed to multiple interferers from nearby cells which are not in the coordinating cluster. Improved gains would be expected if more coordination (thus more complexity) is introduced. We address this below.

C. Coordination across CoMP clusters

The analysis presented so far focused on coordination within a CoMP cluster. As described, this coordination may be assumed as perfect in the case of RRH-based CoMP (facilitated through a centralized scheduler) or as subject to constraints in case of the distributed architecture.

Another aspect that is important to address, though, is the coordination across CoMP clusters. Clearly, in the case of RRH-based CoMP, the centralized processor may only manage a relatively small number of cells due to processing constraints. Boundaries between adjacent CoMP clusters are therefore inevitable and are important to take into account. The distributed CoMP architecture, on the other hand, is less impacted as processing anyway occurs in a distributed fashion at each of the cells similar to macro/pico HetNet without CoMP. Boundary issues may therefore, at least in principle be avoided.

Boundary issues between CoMP clusters are important to address as part of CoMP algorithms as UEs in boundary areas may not be able to participate in CoMP. In HetNet setups, this may introduce significant issues if resource partitioning patterns are not aligned across CoMP clusters or if CRS interference cancellation of adjacent cells' reference signals is not supported. While a detailed discussion of all pertinent aspect goes beyond the scope of this paper, it has been demonstrated that CRS interference cancellation and resource partitioning remain crucial in HetNets, even if CoMP is supported [17].

V. A DETAILED CoMP DESIGN FOR RRH

We now describe an advanced yet practical CoMP scheme for macro-RRH heterogeneous networks. The backhaul links between the central unit and the remote radio-heads is assumed to be ideal, namely characterized by zero latency and infinite capacity. The proposed CoMP scheme boils down to the following two enhancements on top of the hetnet eICIC features described in Section III.

- **UE-side**. A novel CSI reporting scheme that allows the network to have an estimate of the channel quality experienced by a UE as a function of the transmission beam employed by one, suitably selected, strong interfering cell.
- **Network-side**. By leveraging the ideal backhaul among the nodes forming a cluster, namely one macro BS along with its attached RRHs, one single centralized scheduling algorithm can joinly take optimal scheduling decisions for all such nodes, hence achieving optimal intra-cluster coordinated scheduling. Such centralized scheduling algorithm leverages the advanced CSI reporting scheme mentioned above.

The two techniques above are described in detail in the following.

A. CSI reporting enhancements

In heterogeneous networks, UEs associated with picos may experience strong interference from the macro cell, especially when they are located in the range expansion region. The concept of resource partitioning and almost blank subframes is an effective way of mitigating this interference as discussed in Section III. On shared subframes, however, the question arises whether CoMP may offer performance gains by exploiting the centralized system architecture for improved interference coordination.

To enable the tight interference coordination offered by a centralized scheduler, changes to the UEs' feedback reporting are required. In particular, the UE feedback needs to be augmented such as to provide channel quality indication (CQI) not only under one serving assumption but also conditioned on various interference hypotheses. This concept is illustrated in Fig. 4 where UE1 is assumed to be associated with RRH1. In non-ABS subframes, where the macro eNB1 is transmitting, UE1 may generally see strong interference from the eNB1; however, the impact of this interference depends on the specific precoder and transmission rank chosen at the macro. For example, if the macro transmission uses rank-1, then UE1 may be able to partially suppress the interference when it is equipped with a suitable MMSE receiver.

The feedback supporting the proposed CoMP scheme is coined *multi-hypothesis* feedback in line with the above description. In this framework, pico UEs such as UE1 provide CQI feedback for each admissible precoder/rank with which the macro may choose to transmit (LTE uses codebook-based precoding so there is a finite number of precoding choices). In this way, a centralized scheduler can accurately evaluate the impact that a certain precoder/rank decision at the macro will have on specific pico UEs. This leads to improved scheduling and rate prediction as will be explained in detail as part of the centralized scheduling procedures. Besides providing feedback under different precoder assumptions, muting of the macro eNB may also be considered as a separate hypothesis. It is important to note that multiple transmission hypotheses are only assumed for the dominant macro interferer; residual interference from other sources is common across the different feedback hypotheses, as also illustrated in Fig. 4.

It may seem as though multi-hypothesis feedback would increase feedback payload proportionally with the total number of considered precoder/rank hypotheses. This is not the case, however, thanks to the structure of the feedback reporting in LTE. In fact, multi-hypothesis feedback is only needed for the CQI, a metric that essentially provides a quantized version of the received SINR conditions. Intuitively, by providing such CQI under different hypotheses, the impact of additional interference becomes apparent. On the other hand, other feedback information such as precoding or rank feedback from the UE (to its serving pico cell) is likely influenced to a lesser degree by the macro interference. It therefore only needs to be conveyed once as opposed to for each of the multiple hypotheses.

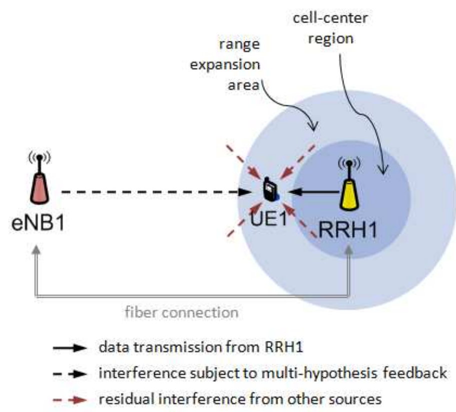

Fig. 4. Illustration of multi-hypothesis feedback.

B. Centralized scheduling algorithm

We consider a macro-RRH deployment where each macro *sector* is assumed to be fiber-connected with N RRHs, the backhaul being assumed ideal. Note that, even though in a sectorized deployment like the one considered in this paper a single macro eNB controls multiple sectors, we don't assume any scheduling coordination among sectors belonging to the same macro eNB – and, of course, no scheduling coordination among different eNBs, except the loose coordination discussed in Section III. A potential improvement to the proposed scheme leveraging inter-sector scheduling coordination, is discussed at the end of this Section.

Since each cluster (a cluster being composed of one macro sector and N RRHs) takes scheduling decisions independently of neighbors, we focus on a single cluster for the rest of this Section. Furthermore, for the sake of clarity, we assume that all UEs in the system are advanced UEs featuring the multi-hypothesis feedback discussed in Section V-A, although we point out that supporting legacy UEs is a straightforward extension of the proposed algorithm.

Let's now assume that all UEs associated to a given RRH are instructed by the network to periodically feedback a multi-hypothesis CQI (MH-CQI) where the dominant interferer assumed by the UE for the sake of MH-CQI is the macro cell the considered RRH is fiber-connected to. Note that, in practice, this may be challenging for those RRH UEs for which the macro signal is very weak (e.g., much weaker than the signal from the serving RRH). Such UEs may not be able to acquire synchronization to the macro cell and, depending on the practical implementation of the MH-CQI estimation algorithm, may therefore not be able to compute a MH-CQI at all. On the other hand, UEs in such conditions are by definition not significantly affected by interference coming from the considered macro cell, thus the need for a MH-CQI feedback for those UEs vanishes.

For each scheduling resource,[4] the proposed cluster-wide joint scheduling algorithm can be formalized as the following optimization problem:

$$\max_{A} \sum_{c \in Cluster} \frac{\eta\left(A(c); A(m)\right)}{\overline{T}\left(A(c)\right)} \tag{1}$$

where $A(c) \in \{U_c, \emptyset\}$ is the assignment function which, for each cell c, determines the UE to be served on the considered scheduling resource, being U_c the set of active users served by cell c. Note that $\emptyset$ indicates the muting hypothesis. Furthermore, in eq. (1), m is the macro sector index, $\eta\left(A(c); A(m)\right)$ is the spectral efficiency achievable when cell c schedules user $A(c)$ assuming that the macro scheduling decision is $A(m)$. Note that such dependence on the macro's decision stems from the use of MH-CQI. Finally, $\overline{T}(u)$ indicates a filtered measure of the throughput achieved by user u.

Note that eq. (1) is an extension of the proportional fairness (PF) criterion to the considered joint scheduling problem, where the cluster-wide utility becomes the *sum* of the PF metrics for each cell in the cluster. We also point out that the optimization criterion described above applies to subframes where the macro is allowed to transmit PDSCH, namely non-ABS subframes. On ABS subframes a baseline scheduler, which takes independent scheduling decisions for each cell in the cluster, can be adopted.

For each scheduling resources, the optimization problem in eq. (1) can be efficiently solved through the following procedure:

1) All the possible transmission hypotheses at the macro side are exhausted, including muting (*i.e.*, no macro transmission on the considered scheduling resource) and all transmission beams belonging to the codebook.
2) For each transmission hypothesis, an optimal scheduling decision is taken for the macro cell, conditioned on the hypothesis. For instance, if the hypothesis consists of a transmission with a specific number of layers and beam(s), only the UEs which fed back a precoding matrix indicator (PMI) and rank indicator (RI) compatible with that transmission are accounted for. Among such "compatible" UEs, selection is made according to the PF criterion.
3) For each RRH fiber-connected to the considered macro, an optimal scheduling decision is taken assuming the hypothesized macro transmission. In particular, among the multiple CQI fed back by each RRH UE, one is selected according to the macro transmission hypothesis, and such CQI is used for the sake of rate prediction. Once the MH-CQI are pruned to single CQI for each candidate RRH UE based upon the macro hypothesis, a single-cell scheduling algorithm based on the PF criterion is independently run for each RRH[5].
4) Given the scheduling decisions taken by the macro cell and all the RRHs for a given macro hypothesis, a cluster-

[4]We define as "scheduling resource" a set of consecutive physical resource blocks (PRBs) of a single subframe which determines the scheduling.

[5]Note that there is no scheduling coordination among the RRHs. This mainly stems from the selected MH-CQI structure, where the dominant interference is always assumed to be the macro cell and no additional information about neighbor RRHs' transmission hypotheses is considered by the UE when an MH-CQI is computed.

wide utility metric is computed as the sum of the single-cell PF utility metrics of each of the involved nodes. Note that the macro PF utility metric associated to the macro muting hypothesis is zero.

5) The hypothesis corresponding to the maximum cluster-wide utility metric is selected and the scheduling decisions (of all cells in the cluster) associated to such hypothesis are finalized. Rate prediction for the sake of modulation and coding scheme (MCS) selection for the scheduled RRH UEs shall be based on the CQI associated to the selected macro hypothesis.

Note that, besides the potential beamforming gain described in Section IV, the proposed scheme could also entail a rate prediction gain for the RRH UEs. In fact, thanks to the MH-CQI, the scheduler gets to know the specific channel quality experienced by an RRH UE when the selected macro transmission beam is employed, thus allowing a rate prediction matched to the actual macro transmission strategy.

It is worth discussing the algorithm's behavior when re-transmissions are pending. We assume that re-transmissions are prioritized and thus always preempt new transmissions so as to minimize the average packet delay[6]. Hence, if an RRH has a pending re-transmission, namely a packet transmitted 8 (in FDD) subframes before didn't get decoded correctly, such re-transmission will always be picked as a final scheduling decision, regardless of the macro hypothesis and the other UEs pending new transmission in the considered RRH. Additional UEs will be considered for the resources not already taken by the scheduled re-tranmission(s). If the macro cell has a pending re-transmission on a given resource, the only tranmission hypothesis compatible with the considered retransmission is selected, and there's no optimization over the hypotheses. Scheduling at the RRHs proceeds as before.

C. Extension to inter-sector coordination

As we pointed out earlier, since in sectorized deployments a single eNB may control multiple sectors, a natural extension of the RRH-CoMP scheme we proposed is to increase the cluster size by defining as a "cluster" the set of *all* macro sectors belonging to the same macro eNB plus all the RRHs fiber-connected to such eNB. A single centralized scheduling algorithm can be defined for the whole cluster, thus implicitly introducing inter-sector coordination, which could lead to further performance enhancements (especially for those UEs suffering from significant inter-sector interference, e.g., RRH UEs whose interference spilling from a macro sector different from the one they have been associated to is significant).

Besides the scheduling algorithm, the proposed MH-CQI scheme must be enhanced, too, in order to effectively see any inter-sector coordination. In particular, taking into account the practical uplink overhead limitations, the following proposal is advocated (because of the lack of space, the detailed analysis of such proposal is postponed to a future publication):

- Each RRH UE determines (or, alternatively, is instructed by the network) one dominant macro interferer and a second dominant macro interferer, e.g., based on received signals strengths. Both selected macro cells must belong to the same eNB to which the RRH is fiber-connected to.
- Exhaust all possible tranmission hypotheses of the first dominant interferer, assuming transmission from the second interferer, and compute the CQI values accordingly (that would be the same as the baseline MH-CQI described in Section V-A).
- Add an additional CQI value assuming muting hypotheses for both the first and second interferers. This, in general, will be the highest reported CQI, since it assumes no interference from *two* potentially strong interferers.

The scheduler shall exhaust combinations of hypotheses for all the sectors belonging to the considered eNB. In particular, for all hypotheses where two sectors are muted, the additional CQI mentioned above shall be used. For the optimal hypothesis selection, handling of re-transmission, etc., the same strategies discussed in Section V-B can be employed.

VI. Performance results

An extensive simulation analysis has been carried out with the aim of evaluating the throughput in the considered macro-RRH heterogeneous LTE network. Three schemes are compared:

i Co-channel deployment, no CRE, no eICIC (*i.e.*, Rel-8)
ii 18dB CRE bias toward the RRHs, eICIC, transmission mode 4 (*i.e.*, Rel-10)
iii 18dB CRE bias toward the RRHs, eICIC, RRH CoMP with centralized scheduling and multi-hypothesis CQI as described in Section V, transmission mode 9 (*i.e.*, Rel-11 or beyond)

Note that transmission mode 9 (TM9) has a higher overhead than TM4, and such overhead is accounted for in the throughput values shown below.

The network layout described in Section II (with 4 RRHs deployed per macro sector) has been used in the computer simulations, and both configuration 1 (uniform) and configuration 4b (clustered) have been considered. It is assumed that each sector of a macro eNB is assigned a different CRS offset while RRHs CRS offsets are chosen randomly. Further simulation pararameters are described in Table I.

We also point out that, for schemes ii and iii, a static time-domain resource partitioning is applied, namely all macro eNBs blank 3 subframe out of 8 according to a periodically repeating pattern which is common among all macro eNBs.Such 37.5% macro blanking has been shown to achieve the optimal eICIC edge user throughput performance and has therefore been assumed for all the throughput simulations in this Section. In particular, Table II shows the edge[7], median, and mean user throughput as a function of the percentage

[6]Note that the DL HARQ in LTE is asynchronous, namely the delay between a retransmission and the original transmission doesn't have to be fixed (the minimum roundtrip value is provided, though).

[7]"Edge" throughput is defined as the throughput of the worst 5%-ile UEs in the system

TABLE II
eICIC PERFORMANCE AS A FUNCTION OF THE PERCENTAGE OF RESOURCES BLANKED BY THE MACRO BSS, FOR CONFIGURATION 1 (VALUES IN [MBPS]).

	Co-channel	**eICIC**				
	0% ABS	12.5% ABS	25% ABS	37.5% ABS	50% ABS	62.5% ABS
Edge	0.38	0.27 (-31%)	0.50 (+29%)	0.59 (+55%)	0.51 (+33%)	0.41 (+7%)
Median	0.83	1.22 (+47%)	1.44 (+73%)	1.54 (+86%)	1.67 (+102%)	1.67 (+102%)
Mean	2.34	2.61 (+12%)	2.69 (+15%)	2.74 (+17%)	2.91 (+24%)	3.06 (+31%)

TABLE I
SIMULATION PARAMETERS

Parameter	**Value**
Macro eNBs	19 sites, 3 sectors/site
Inter-site distance	500 m
RRHs	4 per macro sector
eICIC CRE bias	18dB
Antennas	2 TX, 2 RX
Antenna patterns	Sectorized (macros), omni (RRHs and UEs)
Macro vertical antenna tilt	10 degrees
Line-of-sight modeling	No
TX powers	46 dBm (macros), 30 dBm (RRHs)
Antenna gains	14 dB (macros), 5 dB (RRHs), 0 dB (UEs)
Bandwidth	10 MHz (50 RBs)
Fading model	Uncorrelated antennas, TU-3
Traffic	Full buffer
Control region	3 OFDM symbols
Feedback	5 ms periodicity for both RI and CQI/PMI
CQI backoff loop	Enabled. Dual (clean/unclean) when ICIC is used
Target block error rate	10%
Receiver type	Linear MMSE with perfect interference estimation

TABLE III
THROUGHPUT COMPARISON (VALUES ARE IN [MBPS]).

Configuration 1 (uniform)			
	Co-channel	**eICIC**	**RRH-CoMP**
Edge	0.38	0.59	0.57 (-3%)
Median	0.83	1.54	1.54 (0%)
Mean	2.34	2.74	2.48 (-9%)
Configuration 4b (clustered)			
	Co-channel	**eICIC**	**RRH-CoMP**
Edge	0.40	0.75	0.74 (-1%)
Median	1.03	2.32	2.13 (-8%)
Mean	3.10	3.51	2.98 (-15%)

of blanked resources, for configuration 1. eICIC performance relative to baseline i (co-channel) is also shown in the table.

Table III shows the throughput comparison between the three considered schemes, for both configuration 1 and configuration 4b. As mentioned above, for both eICIC and RRH-CoMP 37.5% of resources are blanked by the macro eNBs. In fact, we remark again that the proposed CoMP scheme is built on top of eICIC such that all the eICIC features are still used when CoMP is employed. Results in Table III shows that the throughput *degrades* when CoMP is enabled on top of eICIC. We remark that such degradation is due to the larger overhead stemming from the use of different transmission modes. In fact, if we ignored the additional overhead due to TM9, CoMP would exhibit a throughput gain over eICIC, although small[8].

In order to better understand the behavior of the proposed scheme we assessed to which extent the selected macro transmission strategies are impacted when the centralized scheduling is employed. We emphasize that if we assume a baseline CQI feedback, rather than the proposed multi-hypothesis feedback scheme, it would still be possible to employ the centralized scheduling which in such case would boil down to a baseline distributed scheduling, since the lack of MH-CQI feedback doesn't allow to perform any scheduling coordination among the cells. Table IV shows the macro transmission strategy statistics, for both the Rel-10 eICIC baseline and the proposed RRH-CoMP scheme, as a function of the percentage of assigned blanked macro resources. These results give a good indication of how the macro scheduling decisions may change when the scheduler is meant to optimize a *cluster-wide* utility function, rather than a *local* metric. On the other hand, since different transmission beams have been collapsed in a single number (only the number of transmission layers are differentiated in the table), these results don't give a full insight because different transmission beams would end up in the same number eventually. Nevertheless, we think that these results are anyway very useful because they show how frequently the cluster-wide scheduling would end up changing the number of transmission layer, or even decide to mute the macro, with respect to the baseline distributed scheduler.

VII. CONCLUSIONS

The deployment of heterogeneous networks composed by a mix of cells with significantly different characteristics (including transmission power and deployment cost) is an economically viable way to overcome the capacity crunch expected in the next few years. Such heterogeneous networks pose interesting technical challenges, including a potentially significant inter-cell interference, which can be efficiently mitigated through the enhanced inter-cell interference coordination scheme which we reviewed in Section III.

In this paper we asked ourselves the following question: how can the coordination among cells be further leveraged so as to enhance the network performance with respect to the eICIC scheme mentioned above? For this purpose, we reviewed various schemes belonging to the so-called coordinated multi-point framework, where tight coordination among nodes is exploited in various ways to improve inter-cell interference

[8]We point out that the additional overhead introduced by TM9 is mainly due to the UE-specific reference signal (UE-RS). The dimensionality loss introduced by UE-RS depends on the number of transmit layers and is approximately 10% assuming up to two layers [14, Section 6.10.3].

TABLE IV
MACRO TRANSMISSION STATISTICS FOR R10 EICIC AND THE PROPOSED RRH-COMP SCHEME, FOR CONFIGURATION 4B.

Macro TX strategy →	**Muting**		**Single layer**		**Dual layer**	
ABS ↓	Rel-10	CoMP	Rel-10	CoMP	Rel-10	CoMP
12.5%	12.5%	22.7%	62.4%	58.9%	25.1%	18.4%
25%	25.0%	28.8%	52.7%	53.1%	22.3%	18.1%
37.5%	37.5%	39.0%	44.2%	45.2%	18.3%	15.8%
50%	50.0%	50.6%	35.2%	36.2%	14.8%	13.2%
62.5%	62.5%	62.7%	26.1%	27.0%	11.4%	10.3%
75%	75.0%	75.0%	17.6%	17.8%	7.4 %	7.2 %

management and eventually enhance the user experience. We proposed a novel CoMP scheme, applicable when the backhaul among the cells belonging to a cluster of cells is fast enough (e.g., fiber-based backhaul), that mainly consists of an improved CQI feedback scheme and an advanced, centralized, cluster-wise optimal, downlink scheduling algorithm.

Even though the performance gains promised by CoMP are significant, the practical limitations of real cellular deployments pose serious challenges and are shown to reduce the effective gains remarkably. Among such practical limitations it is worth mentioning: (a) the CQI feedback limitations, e.g., in terms of uplink data rate consumed by the feedback, (b) practical backhaul constraints, e.g., latency and throughput[9], (c) excessive downlink overhead because of the increasingly large amount of reference signals required in most CoMP schemes, and (d) control channel limitations, which may be a significantly relevant issue in network characterized by bursty traffic and a large number of low-data-rate concurrent traffic flows.

Because of some of the practical limitations mentioned above, the scheme we proposed in this paper failed to exhibit any significant downlink capacity gain on top of the state-of-the-art interference management scheme that we reviewed as part of this contribution. Nevertheless, it is worth noting that CoMP schemes such as the one proposed in this paper could provide benefits other than capacity improvements, e.g., in terms of mobility enhancements. Furthermore, there are many areas of improvements for the current state-of-the-art CoMP schemes, which may eventually lead to a significant capacity boost. In particular, it is worth mentioning potential feedback enhancements, which are of course constrained by the maximum uplink data rate devoted to feedback. Time division duplex (TDD) networks may be suitable candidates for such future studies, since channel reciprocity could be leveraged in this case for the sake of channel estimation *at the transmitter side*, as long as technical challenges such as TX-RX imbalances can be overcome.

REFERENCES

[1] "The mobile internet report," Morgan Stanley, Tech. Rep. [Online]. Available: http://www.morganstanley.com/institutional/techresearch/pdfs/Theme_6_Da%ta_Growth.pdf

[2] "Cisco visual networking index: global mobile data traffic forecast update, 2010–2015," Cisco, Tech. Rep. [Online]. Available: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705%/ns827/white_paper_c11-520862.pdf

[3] M. Dohler, R. Heath, A. Lozano, C. Papadias, and R. Valenzuela, "Is the phy layer dead?" *Communications Magazine, IEEE*, vol. 49, no. 4, pp. 159 –165, april 2011.

[4] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell networks: a survey," *Communications Magazine, IEEE*, vol. 46, no. 9, pp. 59 –67, september 2008.

[5] M.-S. Alouini and A. Goldsmith, "Area spectral efficiency of cellular mobile radio systems," *Vehicular Technology, IEEE Transactions on*, vol. 48, no. 4, pp. 1047 –1066, jul 1999.

[6] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3gpp heterogeneous networks," *Wireless Communications, IEEE*, vol. 18, no. 3, pp. 10 –21, june 2011.

[7] A. Khandekar, N. Bhushan, J. Tingfang, and V. Vanghi, "Lte-advanced: Heterogeneous networks," in *Wireless Conference (EW), 2010 European*, april 2010, pp. 978 –982.

[8] R. Madan, J. Borran, A. Sampath, N. Bhushan, A. Khandekar, and T. Ji, "Cell association and interference coordination in heterogeneous lte-a cellular networks," *Selected Areas in Communications, IEEE Journal on*, vol. 28, no. 9, pp. 1479 –1489, december 2010.

[9] A. Barbieri, A. Damnjanovic, T. Ji, J. Montojo, Y. Wei, D. Malladi, O. Song, and G. Horn, "LTE femtocells: system design and performance analysis," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 3, April 2012.

[10] RAN1#66bis, "High-level proposals on DL-CoMP," Qualcomm Incorporated, Tech. Rep. R1-113385, Oct. 2011.

[11] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects," 3rd Generation Partnership Project (3GPP), TR 36.814, Mar. 2010. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36814.htm

[12] ——, "Evolved Universal Terrestrial Radio Access (E-UTRA); Radio resource control (RRC); Protocol specification," 3rd Generation Partnership Project (3GPP), TS 36.331, Dec. 2011. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36331.htm

[13] ——, "Evolved Universal Terrestrial Radio Access (E-UTRA); X2 Application protocol (X2AP)," 3rd Generation Partnership Project (3GPP), TS 36.423, Dec. 2011. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36423.htm

[14] ——, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation," 3rd Generation Partnership Project (3GPP), TS 36.211, Dec. 2011. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36211.htm

[15] ——, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures," 3rd Generation Partnership Project (3GPP), TS 36.213, Dec. 2011. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36213.htm

[16] S. Annapureddy, A. Barbieri, S. Geirhofer, S. Mallik, and A. Gorokhov, "Coordinated joint transmission in WWAN," in *Communication Theory Workshop, 2010. IEEE*, May 2010.

[17] RAN1#66bis, "Phase 2 evaluation results for HetNet CoMP scenarios," Qualcomm Incorporated, Tech. Rep. R1-112544, Oct. 2011.

[18] M. Costa, "Writing on dirty paper (corresp.)," *Information Theory, IEEE Transactions on*, vol. 29, no. 3, pp. 439 – 441, may 1983.

[19] V. Cadambe and S. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *Information Theory, IEEE Transactions on*, vol. 54, no. 8, pp. 3425 –3441, aug. 2008.

[9] Note that, even if we assume fiber-based backhaul, this is usually limited to a few nodes (the so called "cluster"), while any coordination among nodes belonging to different clusters, if allowed, is usually through a much slower backhaul (see Section IV-C).