Deliverable D4.4

Energy saving: innovative concepts and performance evaluation

Date of delivery	03/03/2015
Contractual date of delivery	28/02/2015
Project number	C2012/1-8
Editor(s)	Yves Louet (SUP)
Author(s)	Yolanda Fernández (TTI), Adrián Sánchez (TTI), Kimmo Hiltunen (ERICSSON), Yves Louet (SUP), Mourad Khanfouci (MERCE), Gregory Gougeon (SIR), Laurent Maviel (SIR), Mathieu Brau (SIR), Yves Lostanlen (SIR)
Dissemination level	PU/RE/CO
Workpackage	4
Version	
Total number of pages	86

Abstract:

This deliverable provides performance evaluation of innovative concepts for energy savings in heterogeneous network deployments. We focus on Radio Resource Management (upper layer) reconfiguration and on analog/digital radio front end (lower layer) flexibility. For RRM reconfiguration we consider eNodeB sleep mode and cell reactivation mechanisms based on uplink interference measurements, compensation-based ON/OFF energy savings with graph representation and learning algorithms for small cell load balancing. For front end flexibility we consider joint power amplifier linearization/crest factor reduction approach and reconfigurable power amplifiers with adaptive operation points. In addition, this study highlights through simulations the interest in switching off cells when necessary to optimize heterogeneous network energy consumption

Keywords: energy efficiency, energy saving, sleep mode, ON/OFF mechanisms, power amplifier

Document Revision History

Version	Date	Author	Summary of main changes
V0	27/10/14	Yves Louet (SUP)	ToC
V1	08/12/14	Yves Louet (SUP)	Merge of all propositions
V2	30/01/15	Yves Louet (SUP)	Stable version
V3	03/03/15	Yves Louet (SUP)	Second stable version and final

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This deliverable deals with energy saving mechanisms in heterogeneous networks from both Radio Resource Management (RRM) and base station equipment point of view. Regarding the RRM, smart ON/OFF switching mechanisms are proposed and regarding the base station equipment, solutions to mitigate the power amplifier consumption are put forward.

The first part of this deliverable considers various energy saving mechanisms based on RRM reconfiguration, in particular mechanisms where base stations are switched on and off on the basis of traffic needs.

One of the problems related to the state-of-the-art eNodeBs is that they consume a considerable amount of power even when they are idle, i.e. not serving any users. As concluded by the EARTH project [1], an efficient way to reduce this idle eNodeB base-line power consumption is to apply some forms of fast cell DTX, e.g. micro sleep or antenna muting. The main principle of micro sleep is to switch off the eNodeB power amplifier when the eNodeB has nothing to transmit. A eNodeB sleep mode mechanism is proposed, where the pico-eNodeBs are switched to sleep mode when they have been idle for a given period of time. In this initial scheme the decision to enter the sleep mode is made by the pico-eNodeB itself, without any leverage of the neighboring eNodeBs. When it comes to the reactivation of a sleeping pico-eNodeB, a number of different schemes can be implemented, based either on the umbrella macro cell load or on the level of the received uplink interference, measured by the sleeping pico-eNodeB itself. Furthermore, a combination of these two schemes is possible.

In a second step, we consider a novel approach in which base stations are opportunistically switched on and off to improve the energy efficiency in wireless small cell networks. The proposed approach enables small cell base stations to optimize their downlink performance while balancing the load among each other and satisfying users' quality-of-service requirements. The problem is formulated as a non-cooperative game among the base stations that seek to minimize a cost function which captures a trade-off between energy expenditure and load. To solve this game, a distributed learning algorithm is proposed where the base stations autonomously choose their optimal transmission strategies.

In this context of sleep modes, the compensation of coverage by nearby base stations is a key issue. Then, an update on compensation based ON/OFF energy saving is presented where the coverage compensating base stations are determined through dominant sets derived from the neighborhood graph of a group of uncoordinated base stations (campus deployment scenario). The first results show that the proposed ON/OFF energy saving technique is beneficial for both energy savings and optimization of the overall system performance.

All the aforementioned studies are confirmed by system-level simulations using realistic radio propagation models. The results highlight the interest in switching off small-cells outside busy hours to optimize the network energy consumption. A particular scenario shows that a 29% gain in energy consumption can be achieved, without compromising the user QoS, when switching off small-cells having a traffic load lower than 3%. This strategy leads however to a residual service outage observed at peak hours which is partly due to the Almost Blank Subframe duty cycle limitation suggesting the use of Self-Organizing Network (SON) algorithms to achieve a time (of day) adaptive macro to small-cell offloading. Peak hours service outage also results from the mismatch between the uniform small-cell layer deployment and the non-uniform active user density. A better strategy would certainly be to deploy a non-uniform small-cell layer that matches the traffic hotspots distribution. Modeling the switch on mechanism is also of interest to assess the impact on service outage and energy consumption. In addition, measurements of energy savings are derived to validate smart base station ON/OFF switching mechanisms in periods where the total amount of traffic served by the network is such that it can be satisfied with just a subset of active base stations.

The second part of this deliverable deals with energy saving mechanisms that can be implemented in base station equipment. These proposed mechanism are focused on the

power amplifier, which is responsible of the majority of the overall power consumption on the transmitter side.

The first study proposes to smartly merge the peak factor reduction (also known as Peak to Average Power Ratio reduction) and the linearization steps prior to power amplification. Indeed, the smaller the peak factor, the larger the power amplifier efficiency will be. This peak factor reduction is directly associated to the linearization steps prior to power amplification, which are designed to compensate the non-linear characteristics of the power amplifier. Results show that, by jointly tuning linearity and peak factor reduction, the transmission performance can be improved leading to a reduction of power amplifier (and thus overall) energy consumption.

The last study proposes to adapt the operating point of the power amplifier. Indeed, this operating point is usually designed for maximum load, providing the highest energy efficiency at maximum RF output power. Nevertheless when the traffic load decreases, lower RF output power levels are required and the energy efficiency gets worse due to the power amplifier characteristics. Therefore, a reconfigurable power amplifier is proposed defining different operating points which can be optimized for different power levels, providing energy savings at medium and low traffic load in the base station radio equipment. Based on the obtained results, the PA efficiency could be improved by 20%, if the proposed dynamically reconfigurable solution is applied instead of the unique operating point solution implemented in conventional power amplifier. Furthermore, as the power amplifier could represent up to 64% of power consumption in macrocell scenarios, and 47% in microcell scenarios, the energy efficiency improvement in an urban area could be around 13% and 9% respectively.

1   INTRODUCTION

Currently, 3% of the world-wide energy is consumed by the Information and Communication Technologies (ICT) infrastructure which generates about 2% of the world-wide CO2 emissions, comparable to the world-wide CO2 emissions by all commercial airplanes or one quarter of the world-wide CO2 emissions by all vehicles. The ICT sector's carbon foot print is expected to quickly grow to 1.4 Giga ton CO_2 equivalents by 2020, nearly 2.7% of the overall carbon footprint from all human activities [29].

Therefore, lowering energy consumption of future communication systems and networks meanwhile increasing its total energy-efficiency is demanding greater attention not only within government, industry and standardization bodies but also within international research communities. Indeed, the Internet is built on the best effort principle, which is the use of the available bandwidth. This is what makes the IP protocols suite work over a 300 kbps modem as well as a 1 GBps Ethernet link. Following the advances in the communication technologies, and the difficulty to set up Quality of Service (QoS) when traffic crosses over multiple networks, the Internet has always increased the bandwidth capacity. This exponential growth of the bandwidth comes at the cost of an increase in energy consumption, especially due to the power consumption of routers needed to support packet forwarding at high rate [30].

The network layer gathers the functionalities to enable a set of devices to communicate together as a network. It mainly provides an addressing scheme, location and routing. Usually, the main functions are implemented in the core of the network, by dedicated routers. However, in specific networks such as sensor networks, end devices may also provide routing features and help in routing data packets towards a destination. Energy consumption of a global network can be divided into energy consumption of the end hosts or the terminals on the one hand, and on the other hand the core network, or the infrastructure. We can distinguish two different models of networks where the ratio of energy consumption is very different between the infrastructure and the terminals [31]. Cellular networks evolve from 50%-50% energy consumption between the infrastructure and the terminals in the GSM network to 90% of energy consumption only by the infrastructure in the UMTS. In the Internet, we observe nowadays that the infrastructure and the end hosts consume the same amount of energy. This is due to the efforts made on energy efficiency on personal computers, by better using stand-by modes for example. The energy star alliance project [32] was also an important initiative to reduce energy consumption of personal computers. However, only recent works have proposed to study what could be made in the infrastructure to reduce the energy consumption. Until now, routing protocols (and core devices) were not designed with any energy constraints in mind. In addition, the rapid development of cloud-based services has provoked the proliferation of data centers with huge energy requirements.

Many works are currently addressing the energy consumption in telecommunication networks (cf. state of the art below) with many different approaches such as web-content delivery services, or delay tolerant network (DTN).

There are a lot of initiatives in the ICT domain to decrease carbon emissions. We will not describe here all these initiatives. We only focus on some of them which are recognized as very pertinent.

At the worldwide level, Greentouch [33] is a consortium of leading ICT industry, academic and non-governmental research experts dedicated to fundamentally transforming communications and data networks, including the Internet and significantly reducing the carbon footprint of ICT devices, platforms and networks. Its claimed objective is to deliver, by 2015, the architecture, specifications and roadmap — and demonstrate key components — needed to increase network energy efficiency by a factor of 1000 from current levels. The results of the SHARING project may contribute to this initiative.

At the European level, some international projects exist within the EC 7[th]Framework Program (FP7), aiming to study green communications, but in general merely from the perspective of energy consumption and without resorting to the Green Cognitive Radio (GCR) approach

based on intelligence advances like the Cognitive Radio (CR) paradigm. For example, European Commission (EC) research Project EARTH (Energy-Aware Radio and NeTwork Technologies - European Integrated Project funded in EC FP7 ICT Call 4), has been focusing on energy saving techniques in mobile broadband systems, which is positioned as the first significant European project directly addressing power efficient wireless communications and networking. Within the EARTH project, energy efficiency in radio access networks is concentrated, which covers a wide span of research activities (e.g., energy-saving components, theoretical limits, etc.), the goal is to find good solutions and concepts that can reduce the energy consumption of mobile broadband systems by 50 percent [34]. We can also mention the FP7 European project TREND [35] aiming to design energy-efficient networks. The TREND research activities focus on collecting power consumption data, identifying architectural solutions for energy-friendly networks and defining network design criteria.

Restrepo *et al*. [36] classifies the energy consumption reduction potential of an infrastructure with three levels: circuit, equipment and network. At the circuit level, specific components using dynamic voltage and dynamic frequency scaling can significantly reduce the energy consumption. Electrical equipment may be replaced by optical devices, and specific actions can be taken regarding the heating dissipation. Another level is the link level which needs low bit rate to approach the lower energy consumption [37]. Indeed, the best effort principle introduced above impacts all layers of the OSI. In particular, the medium access control (MAC) and physical (PHY) layers ensure feasible access channel within multiple access network and with waveforms adapted to the medium of transmission. The best effort principle induces high bit rates with high modulation orders at the PHY level which is responsible for important energy consumption. Conventional energy efficiency is measured as the required energy to send one bit and spectral-energy efficiency tradeoff characterizes the physical layer performances [38]. Packet retransmission is a necessary protocol to increase the successful transmission probability. However, it induces delay in the network and the spectral-energy-delay efficiency tradeoff could be the last network performance taken into account.

The network level also offers solutions to reduce energy consumption. Chabarek *et al*. [39] study the energy consumption of different routers, and investigate the proportion of this energy required per chassis / line card in these routers. This study confirms the intuition that the number of chassis per router must be minimized, and the line cards per chassis must be maximized. The energy consumption - without traffic - of a base system varies from 210 to 430 W depending on the router model. Each additional chassis that is powered on consumes between 25 and 70 W. The effect of traffic is less important, and "only" increases the energy consumption by 10% (which is around 200 W for a fully loaded 12008 chassis). In the experiments (where the chassis were not fully loaded), without traffic the router consumes 755 W against 775 W when there is traffic. We can conclude that the impact of bandwidth is not as important as the impact of the number of chassis powered on. However, there is still a non-negligible consumption linked to the bandwidth. Finally, we can also say that the 755 W consumed by an idle router (without traffic) opens opportunities for hibernation techniques.

Following the numbers given in the previous paragraph, Restrepo *et al*. [40] present the potential for energy saving if a dynamic routing that take into account the energy constraint can be implemented. The idea consists in distributing the traffic to a subset of routers / routers' line card such as to adapt the capacity of the network to the current needed capacity. Redirecting the traffic to a subset of routers would allow to power off or hibernate some routers and thus to save energy. Chiaraviglio *et al*. [41] discusses the Internet topology modeling with graph theory to deduce energy property of the network. They conclude that 80% of the transport nodes can be removed from the graph, still guaranteeing the forwarding, considering both small world and power law distributions. Moreover, the network coding paradigm introduced in [42] specifies that bandwidth can be saved in the network if the information flows are not simply routed or replicated but coded with each others at some specific nodes. Since this pioneered work, other authors have studied the network coding paradigm and its fundamental performance [43]. The network coding can be seen as a mean to reduce energy consumption in the network. In [44], the authors considered multiple

unicast sessions into a superposition of multicast and unicast sessions with coding occurring only at each session. They proposed an energy-efficient backpressure algorithm exploiting the XOR network coding in each unicast session. In [45], the authors showed that significant energy improvements can be achieved using network coding paradigm for broadcasting scenario. From another point of view, network coding for energy efficiency in wireless body area network (WBAN) has been investigated in [46]. They showed that the gains increase as the number of nodes in the network increase and the channels seen by different sensor nodes become more asymmetric.

In cellular networks, energy efficiency is currently becoming a hot topic. Some studies have for instance proposed to perform wired backhauling rather than wireless backhauling in order to become greener [47]. The solution proposed in [48], which consists in using wired line backhaul in femto cell scenario, is an example of such solutions. In [49], the authors consider the downlink scenario and they investigated the energy consumption when some base stations are switched off. They derived the optimal base station switch off strategy in order to keep the average SINR unchanged in the network. They studied the case where the network topology is a line and where it is a plane. While their work is interesting, the results presented are obtained with very basic PHY layer and propagation conditions. Indeed, the authors considered a full frequency reuse network and only path loss propagation condition is considered. The shadowing process is responsible of large variations of the radio link performances and hence has a major impact on the overall network performance. This kind of effect needs to be taken into account in order to evaluate the global energy efficiency of the network. In another paper [50], the authors studied the tradeoff between the energy saving by switching off the base stations and the increase of the uplink power induced. The authors considered a line or a plane network topology and path loss or an exponential attenuation for the signal propagation conditions. Their first results show that the average uplink power (per terminal) is decreasing w.r.t. the number of base stations turned on when the impact of interference is ignored. However, the conclusion is the opposite when the impact of interference is taken into account. The scenarios considered are relatively simple but remain interesting for a first set of trends in green networking.

On the other hand, advances in Cognitive Radio (CR) techniques provide room for improvement concerning the energy-efficiency of the future networks. Papers [51] and [52], pointed out, in 2005 and 2008, respectively, that intelligence is one key word to deal with Green Communications problem (energy-efficiency, radio spectrum pollution, etc.); they proposed to use the CR as an enabling technology for providing green cognitive radio (GCR) communications. GCR is a cognitive radio, which is aware of sustainable development issues as an additional constraint in the decision making function of the cognitive cycle. GCR uses the sensors distributed throughout the different layers of the CR model for taking the best decision, in order to better contribute to sustainable development [53]. GCR is an increasing subject in the community (GREENTIC WS, ISCIT conference, within COST 0902 action,…) .

Wireless communication standards impose stringent requirements on linearity performance of Power Amplifiers (PA). In addition, since the PA consumes most of the energy in telecommunication equipments (almost 60% for macro base stations), its power-efficiency becomes a primary concern. In fact, with the growth in voice and data communication usage, a higher PA efficiency will contribute to the reduction of both the user equipment energy consumption resulting in a longer battery lifetime, and the base station energy consumption resulting in power savings and reduced environmental pollution. However, these two requirements, linearity and power-efficiency, tend to be mutually exclusive in traditional PA design, so that any increase of the PA linearity by amplifying in linear region is usually achieved at the expense of the efficiency and conversely. This PA issue is a primary concern when considering multi-carrier modulations like Orthogonal Frequency Division Multiplexing (OFDM) which are prone to high Peak-to-Average Power Ratio (PAPR) resulting in low power-efficiency.

In such conditions, techniques including linearization and PAPR reduction have been proposed separately in the literature in order to improve the performance of the transmitters, including PA. The linearization ensures high linearity of the PA in order to avoid carrier inter-modulations and to respect the power mask. In wireless communication systems, the most promising and cost-effective linearization method is predistortion which guarantees an acceptable linearity level of the PA over its intended power range. In case of the PAPR reduction, the dynamic range of the signal is reduced, which allows the PA to operate closer to the saturation power more efficiently. PAPR reduction includes some techniques like clipping and filtering, coding and tone reservation. Since PAPR reduction is associated to linearization and improves its effectiveness, the methodology of PA design focuses on a trade-off between linearity and power efficiency represented by linearization and PAPR reduction respectively.

Taking all these aforementioned issues and facts into account, SHARING WP4/Task 4.3 targets to investigate ways to save energy in heterogeneous networks. Precisely speaking, currently deployed base stations need to continuously signal their presence and availability, and to listen to the radio environment to detect incoming users even during the low-traffic periods (e.g. nights). At the same time, the massive large scale deployment of small cells can be ecologically worthwhile only if armed with smart energy efficiency and power saving mechanisms. Motivated by this, energy-efficient sleep mode strategies for small cell base stations are required in order to reduce the power consumption of cellular networks. These strategies allow the hardware components in the base station to be smartly switched off in idle conditions, which means that the total cell, and further the total network energy consumption will become highly load-dependent. Furthermore, in case of deployments with a large number of overlaid cells, serving cell selection can be a key mechanism to reduce the number of activated cells. Indeed, in this case, the classical approach where a UE is connected to the strongest Access Point (AP) may no longer be an optimal strategy in terms of interference and energy efficiency. Hence, there is a need to investigate new ways to select the serving cell for each UE. One solution lies in enabling a dynamic activation of local APs limiting the load of the worst (for example from the energy or interference point of view) APs. A possible improvement is to leverage on predictive Radio Resource Management (RRM) aspects, based on the estimation of traffic evolution, offering the required amount of capacity in both time and space.

As said, the scope of SHARING Task 4.3 is to investigate the performance of various types of energy saving mechanisms both at the network and the base station hardware level. At the network level, the proposed solutions will be based on load balancing and on RRM reconfiguration. More specifically, the common factor is to propose and evaluate the performance of (self-optimizing) energy saving mechanisms based on the use of fast cell DTX (e.g. micro sleep), eNodeB sleep mode and/or dynamical cell switch-off. The investigations will consider both the QoS of the users and the overall network performance (e.g. capacity).

Regarding the base station hardware, the innovations will focus on enhancing the energy efficiency of the PA. Studies will investigate how PAPR standards compliant methods can help to reduce the overall power consumption of the network. Furthermore, real-time tuning and optimization of Radio Frequency (RF) modules will be investigated. Through external parameters such as global and local load demand, different power levels can be defined and radiated power can be adjusted. For different radiated power values, the PA can be in several optimal working points. These optimal working points can even go to low power modes of operation and contribute to energy saving.

2   ENERGY SAVING MECHANISMS VIA RRM RECONFIGURATION

2.1   Utilizing eNodeB fast cell DTX and sleep mode to save energy

One of the problems related to the state-of-the-art eNodeBs is that they consume a considerable amount of power even when they are idle, i.e. not serving any users. As concluded by the EARTH project [1], an efficient way to reduce this kind of fixed power consumption of idle eNodeBs is to apply some form of fast cell DTX, e.g. micro sleep or antenna muting. The main principle of micro sleep is to switch off the eNodeB power amplifier for the time periods when the eNodeB has nothing to transmit. In practice this means the time periods in between the cell-specific reference signals (CRS), broadcast (PBCH) and synchronization symbols (PSS/SSS), see Figure 1. The main advantage of micro sleep is that it can exploit the very short idle periods of the eNodeB. Another benefit of micro sleep is that the cell stays accessible, and hence, the overall system performance will not be affected.
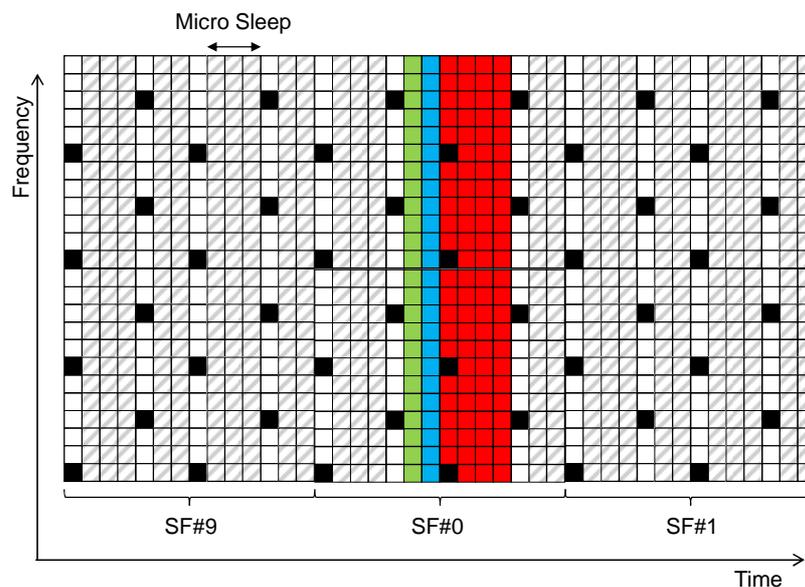

Figure 1 In case of micro sleep the eNodeB power amplifier is switched off in between the downlink control signals.

In case of antenna muting, all transmit antenna branches except of the first one are switched off to reduce the power consumption of an idle eNodeB, see Figure 2. The active antenna branch is needed to keep the cell accessible by transmitting the required downlink control signals (CRS, PBCH and PSS/SSS). The muted antenna branches are then re-activated when needed to serve the traffic. A more detailed discussion of the antenna muting mechanism can be found for example in [1] and [3].
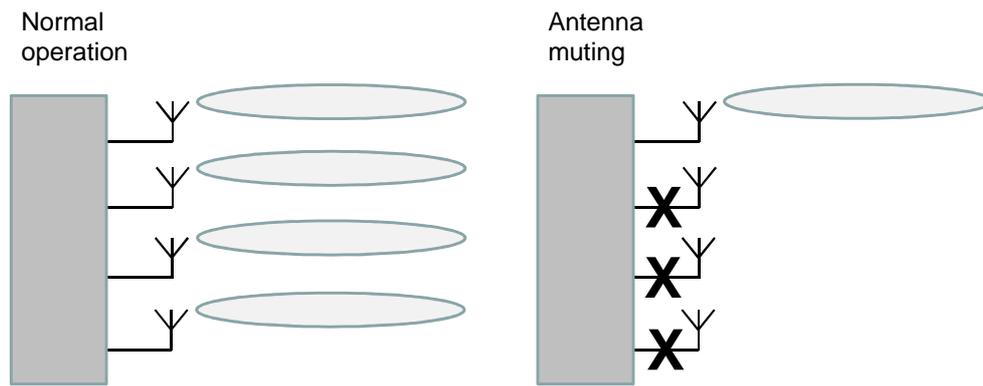
Figure 2 : In case of antenna muting, all transmit
antenna branches except the first one are switched off.

Since the fast cell DTX mechanisms do not affect the accessibility of a cell, they are applicable for all cells, i.e. also the ones required to guarantee the desired system coverage area.

In addition to micro sleep, the pico eNodeBs can be subject to a sleep mode mechanism, where not only the power amplifier, but also the RF and baseband (BB) components of the eNodeB are switched off. In case of the sleep mode the inactive periods are expected to be in the order of seconds. Since all the eNodeB transmissions are deactivated during the sleep mode, the cell will not be accessible to the users, and hence, the maximum achievable system performance will potentially be affected. That is also the reason why sleep mode is not applicable for cells which are needed to provide the overall system coverage.

When it comes to the sleep mode, the main challenge is to maximize the sleep periods, without affecting the system performance too much. Hence, the eNodeB should be deactivated as soon as it is not needed to serve the offered traffic. Similarly, a sleeping eNodeB should be reactivated as soon as it is needed to serve close-by users.

Here, an eNodeB sleep mode mechanism is proposed, where the pico eNodeB is switched to sleep mode as soon as it has been idle for δ_{idle} seconds. In this initial scheme the decision to enter the sleep mode is made by the pico eNodeB itself, without the control of the neighboring eNodeBs.

When it comes to the reactivation of a sleeping pico eNodeB, a number of different schemes can be implemented. For example, the reactivation can be based on the monitored macrocell load [4] (reactivation scheme A in Figure 3). Alternatively, the reactivation can be based on the level of the received uplink interference, measured by the sleeping pico eNodeB itself (reactivation scheme B in Figure 3). Furthermore, a combination of scheme A and scheme B is possible.

In case of reactivation scheme A, macro eNodeB reactivates the sleeping eNodeB, if the monitored macrocell load, averaged over δ_{macro} seconds, is higher than the predefined threshold (λ_{macro}). Suitable measures for the macrocell load can for example consist of the cell throughput or the average utilization of the cell resources. Reactivation scheme A assumes that the macro eNodeB is able to transmit the reactivation command to the desired pico eNodeBs. However, since there are no strict delay or bandwidth requirements for this message, it can be transmitted over the normal X2-interface.
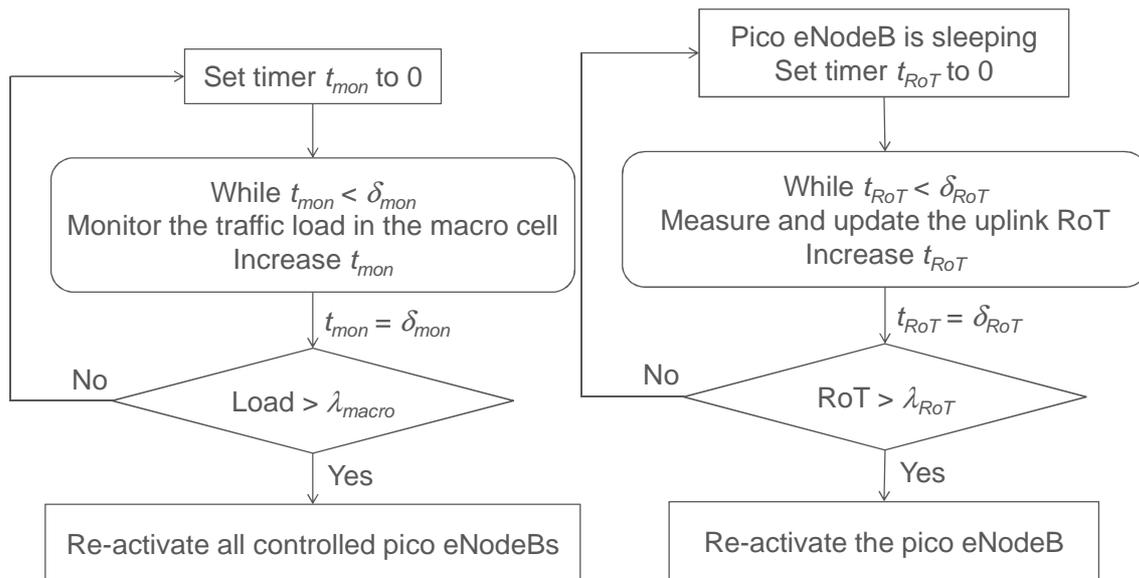
Figure 3 : Two cell reactivation schemes are proposed: a centralized reactivation scheme A (left) and a distributed reactivation scheme B (right).

As mentioned, reactivation scheme B is based on the uplink measurements performed by the sleeping eNodeB. Hence, in that case it is assumed that either the pico eNodeB includes a separate network listening module, capable of performing uplink measurements, or that the pico eNodeB uplink receiver is left active when entering the sleep mode. Now, the principle is that if the measured uplink power, averaged over δ_{RoT} seconds, is greater than the predefined threshold (λ_{RoT}), the sleeping pico eNodeB is reactivated. This also means that there is no need to transmit any additional control signalling between the eNodeBs.

One possible way to combine scheme A and scheme B is to wake up a sleeping pico eNodeB only when $RoT > \lambda_{RoT}$ and at the same time $Load > \lambda_{RoT}$ for the controlling macro eNodeB.

The aggressiveness of the proposed sleep mode algorithms can be adjusted via the timer (δ_{idle}) and the load thresholds (λ_{macro} and λ_{RoT}). For example, the algorithm can be made more aggressive by reducing the value of δ_{idle}. Similarly, the algorithm becomes more aggressive if higher values of λ_{macro} or λ_{RoT} are applied. By increasing the aggressiveness of the sleep mode algorithm, the total network power consumption can be reduced. However, at the same time the risk for an unacceptable reduction of the system performance will increase.

The performance evaluation of the proposed energy saving mechanism is based on Scenario 2.3.3 described in [5]. Hence, the performance of a heterogeneous LTE network deployment consisting of 21 macrocells and 66 outdoor picocells is simulated over a period of 24 hours. During that period both the average hourly area traffic volume and the geographical traffic distribution are assumed to vary. For example, the lowest area traffic volume, 28 Mbps/km^2, is experienced during the early morning hours, while the peak traffic equal to 200 Mbps/km^2 is experienced around 09pm-10pm. Similarly, during the office hours most of the traffic is concentrated within the commercial areas, while during evenings, nights and morning hours most of the traffic is generated within the residential areas.

A dynamical system-level simulator is used, and for each of the 24 hours the system is simulated for 10 minutes (small files) or 30 minutes (large files). The output of the system simulation consists of the observed user performance and the total network power consumption measured as kW/km^2.
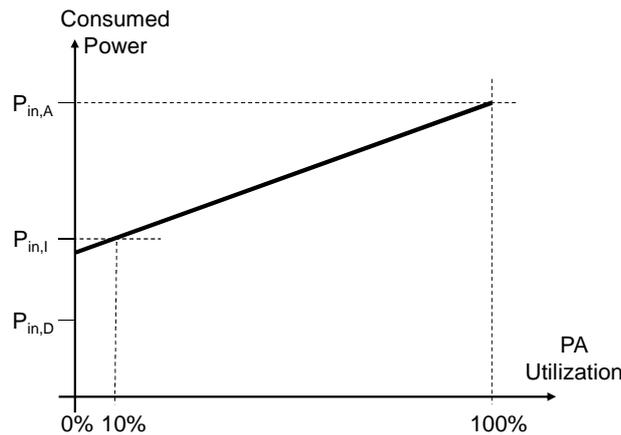
Figure 4 : Assumed model for the eNodeB power consumption.

The impact of fast cell DTX is modelled by defining different eNodeB power consumption values for the different activity states, see Figure 4. In general, an eNodeB is assumed to be either in active state or in idle state. In case of an active state, i.e. when the eNodeB is serving at least one user, the eNodeB is transmitting at the maximum output power and the power consumption is equal to $P_{in,A}$. In case of an idle state, the eNodeB is transmitting only the downlink control signals, and the average power consumption is assumed to be equal to $P_{in,I}$ or $P_{in,D}$ depending on whether fast cell DTX is applied or not. The assumed P_{in} values have been derived from [6] and are listed in Table 1. Yet another assumption is that the power consumption during the sleep mode ($P_{in,S}$) is equal to 10 W.

Table 1 : Assumed eNodeB power consumption values (per cell).

Type of eNodeB	Maximum Output Power	$P_{in,A}$	$P_{in,I}$	$P_{in,D}$	$P_{in,S}$
Macro	40 W	280 W	174 W	144 W	N/A
Pico	1 W	97.7 W	96.3 W	78.1 W	10 W

Evaluation results are presented for two different sleep mode parameter value combinations: *aggressive sleep mode* ($\delta_{idle} = 1$ second and $\lambda_{macro} = 10$ Mbps) and *guaranteed performance* ($\delta_{idle} = 4$ seconds and $\lambda_{macro} = 6$ Mbps). While the aggressive sleep mode results in a low energy consumption throughout the day, the experienced user performance during the peak traffic hours becomes unacceptable, see Figure 5. By relaxing the sleep mode algorithm parameters, the impact on user performance can be reduced and the desired user performance can be "guaranteed" throughout the day. Unfortunately, at the same time the total network energy consumption becomes higher.

In addition to the two sleep mode setups, the evaluations consider scenarios both with and without fast cell DTX. Furthermore, to illustrate the potential of the sleep mode algorithm, results are shown for an ideal sleep mode scheme, where all idle capacity cells are automatically assumed to be in sleep mode [9].
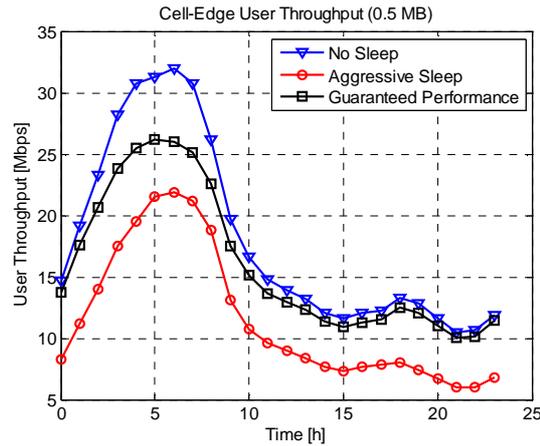
Figure 5 : Observed cell-edge user throughput for the different sleep mode algorithm setups.
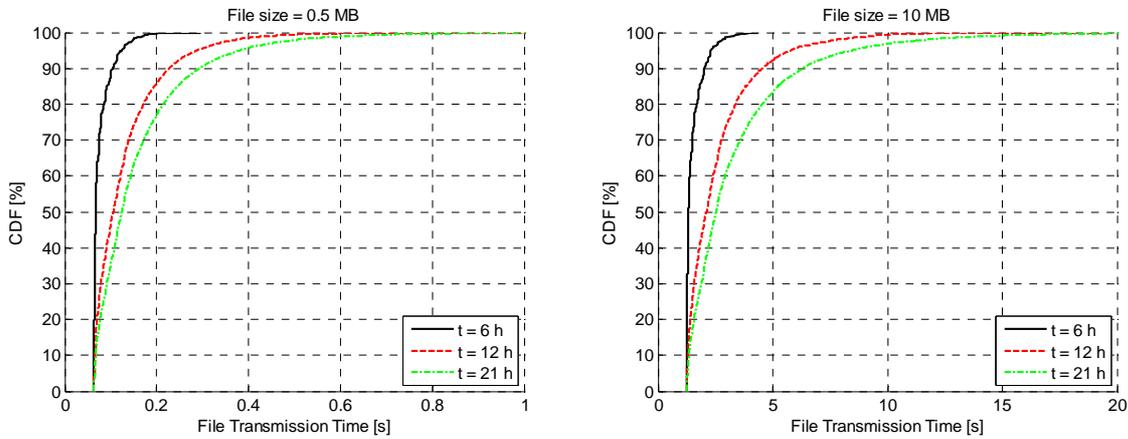

Figure 6 : Observed file transmission times for small (left) and large (right) file sizes.

As described in [5], the evaluation assumes that each user downloads a file of a fixed size, and leaves the system, when the file has been successfully transmitted. To evaluate the impact of file size (i.e. the file transmission time or the user "lifetime") on the performance of the proposed energy saving mechanism, two different file sizes are assumed (0.5 MB and 10 MB), see Figure 6.

Evaluation results for reactivation scheme A are presented in the following figures. As indicated by the results in Figure 7. The proposed sleep mode algorithm is quite effective during the low-traffic hours, in particular with an aggressive parameter setup. Unfortunately, the algorithm is much less effective during the peak-traffic hours, especially if the transmitted files are small. The reason for this is that even though the pico eNodeBs are idle in 95% of the time, the idle periods are so short that the sleep mode algorithm is not able to utilize them. At the same time the lifetime of a user is very short due to the assumed file transmission traffic model. In practice this means that when the macrocell is serving multiple simultaneously active users, the cell load exceeds the pre-defined threshold and the pico eNodeB reactivation is triggered. However, at the time when the pico eNodeBs have been reactivated, the users have already finished their file transmissions, and have left the system. Hence, the algorithm can only attempt to adapt to the longer-term traffic variations between the different areas. In case of larger files, the algorithm can in many cases wake up the pico eNodeB already during the ongoing file transmission.
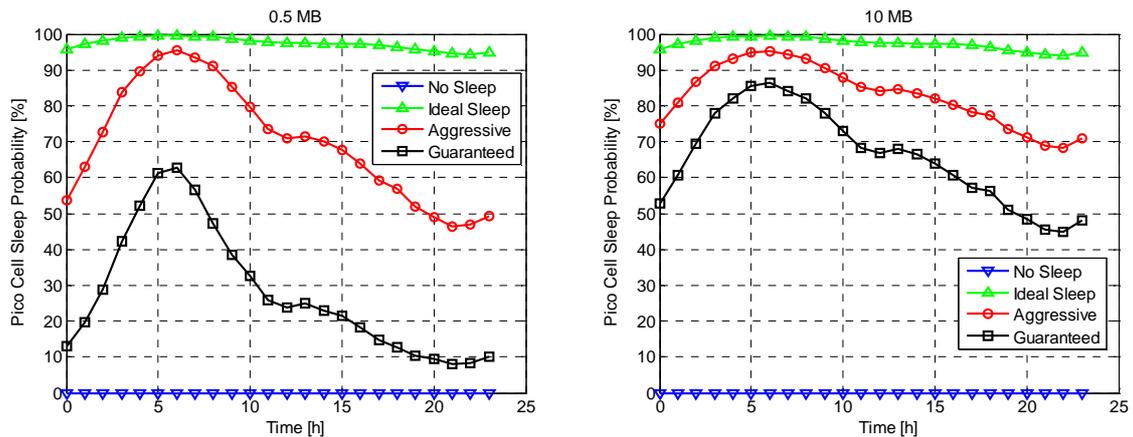
Figure 7 : Picocell sleep probability as a function of time with small (left) and large (right) file sizes. Pico eNodeB reactivation scheme A.

Results for the total network power consumption as a function of time are shown in Figure 8. Looking at the results, the impact of file size becomes evident, in particular for the peak-traffic hours. In case of the low-traffic hours, the performance of the aggressive sleep mode is roughly the same for both file sizes. However, the performance of the less aggressive sleep mode becomes clearly more effective, when larger files are assumed. In all, the gains achieved by the sleep mode are bigger than the gains achieved by the fast cell DTX alone. In case of the peak-traffic hours, the efficiency of the less aggressive sleep mode algorithm is dramatically increased with larger files. The efficiency of the aggressive sleep mode algorithm is improved as well, but as mentioned, the observed user performance is at an unacceptable level. Now, the sleep mode algorithm is less effective than fast cell DTX when the files are small, but the situation changes with large files.
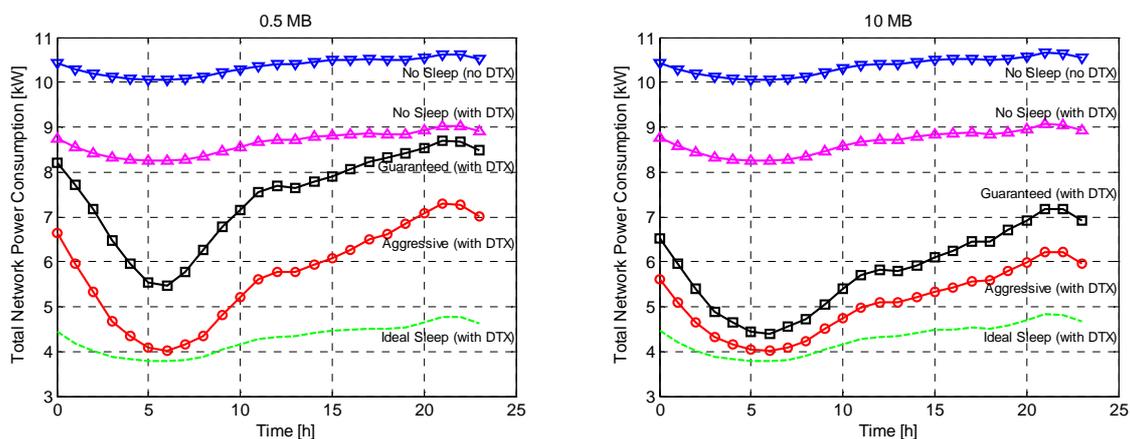

Figure 8 : Total network power consumption as a function of time with small (left) and large (right) file sizes. Pico eNodeB reactivation scheme A.

Finally, results for the normalized daily energy consumption are shown in Figure 9. The results have been normalized with respect to the daily energy consumption of a system without any energy saving mechanisms. Again, it is clearly visible that the proposed sleep mode algorithm is much more efficient, when large files are assumed. For both cases, the gain of fast cell DTX alone is approximately equal to 16%. With the less aggressive sleep

mode setup, but without fast cell DTX, the daily energy consumption becomes approximately 15% (small files) or 35% (large files) lower than without any energy saving mechanisms. If the fast cell DTX is combined with the less aggressive sleep mode algorithm, the daily energy consumption becomes approximately 28% (small files) or 44% (large files) lower than without any energy saving mechanisms.
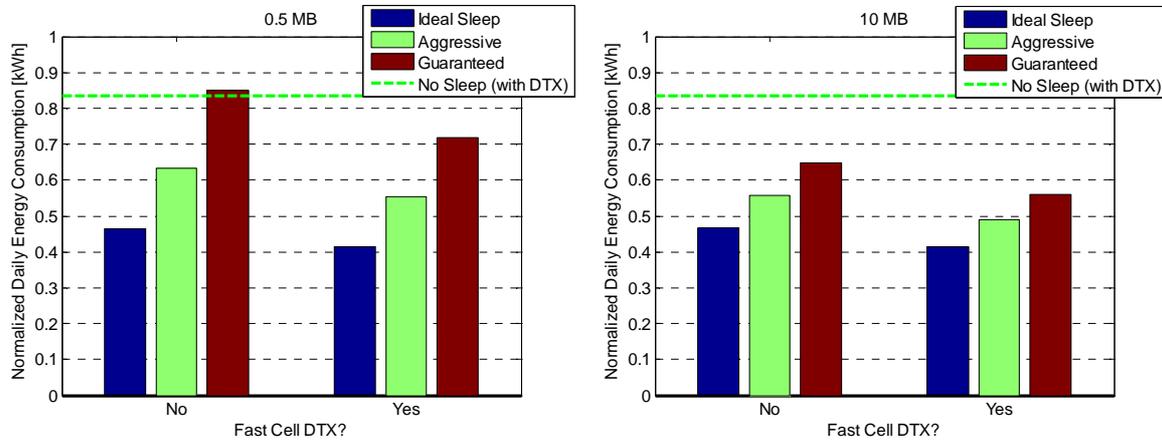

Figure 9 : Normalized daily energy consumption with small (left) and large (right) file sizes. Pico eNodeB reactivation scheme A.

Evaluation results for reactivation scheme B are presented in the following figures. Similar to reactivation scheme A, also scheme B is quite inefficient during the peak-traffic hours when the transmitted files are small. However, when the transmitted files are large, the performance of scheme B is dramatically improved. In fact, the performance of scheme B becomes clearly better than the performance of scheme A. The reason for this is that due to the shorter measurement filtering period, scheme B can better react on the ongoing file transmissions.

Looking at the results for the total network power consumption as a function of time, see Figure 11, the performances of scheme A and scheme B are quite similar when the transmitted files are small. However, when the transmitted files are large, scheme B results in clearly lower network power consumption throughout the day. Furthermore, it becomes evident that the gains achieved by the proposed sleep mode algorithm are larger than the gains achieved by the fast cell DTX.
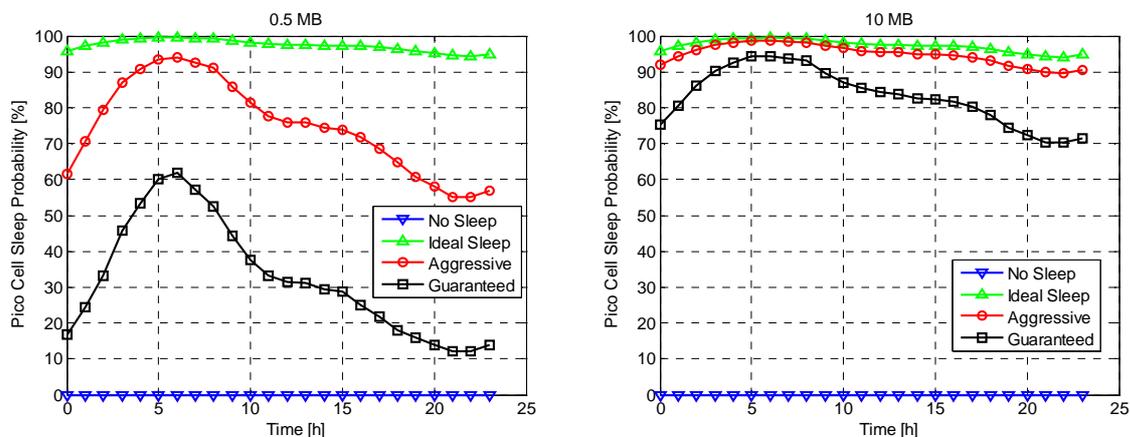

Figure 10 : Picocell sleep probability as a function of time with small (left) and large (right) file sizes. Pico eNodeB reactivation scheme B.
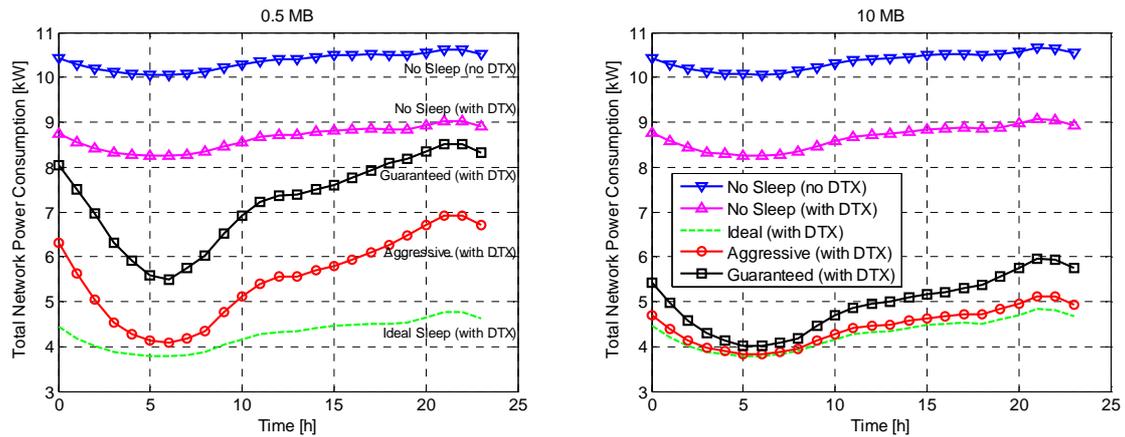
Figure 11 : Total network power consumption as a function of time with small (left) and large (right) file sizes. Pico eNodeB reactivation scheme B.

Results for the normalized daily energy consumption are shown in Figure 12. Again, it is clearly visible that cell reactivation scheme B is much more efficient than scheme A, when large files are assumed. Now, with the less aggressive sleep mode setup, but without fast cell DTX, the daily energy consumption becomes approximately 17% (small files) or 45% (large files) lower than without any energy saving mechanisms. Finally, if the fast cell DTX is combined with the less aggressive sleep mode algorithm, the daily energy consumption becomes approximately 30% (small files) or 52% (large files) lower than without any energy saving mechanisms.
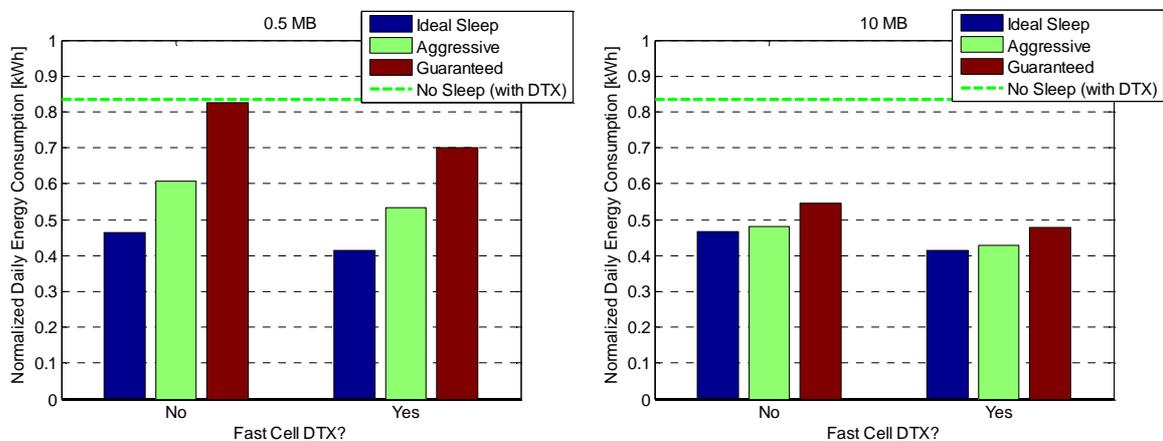

Figure 12 : Normalized daily energy consumption with small (left) and large (right) file sizes. Pico eNodeB reactivation scheme B.

Based on the obtained evaluation results, the distributed pico eNodeB reactivation scheme (scheme B) utilizing uplink measurements offers slightly better performance compared to the centralized pico eNodeB reactivation scheme (scheme A) based on the measured macrocell downlink load. However, both the proposed reactivation schemes are inefficient, when the transmitted files are small, i.e. when the user activity is at a low level. One possible way to improve the performance of the sleep mode algorithm could be to combine scheme A and scheme B into a scheme, where the pico eNodeB is reactivated only when the measured uplink interference is above the threshold, and at the same time the load within the controlling macrocell is sufficiently high. Yet another possible way to improve the sleep mode performance could be to implement a SON-based scheme, which dynamically adjusts the

aggressiveness of the sleep mode algorithm: making the algorithm more aggressive during the low-traffic hours, and less aggressive during the peak-traffic hours.

A more detailed design of the algorithm enhancements and the corresponding performance evaluations are left for further study and are not part of this deliverable.

2.2 Centralized techniques for ON/OFF energy saving in HetNet campus scenario

Compensation-based ON/OFF energy saving is an important energy saving technique that is currently discussed in 3GPP[7]. In compensation-based energy saving, a compensation base station set is determined by the network planning such as the base stations that are not allowed to be turned OFF are in the compensation set. Two compensation-based energy saving strategies are foreseen in the current status of the standardization discussions as a means for energy saving in interference limited scenarios discussed in [7]:

- Single compensating base station energy saving: where one base station compensates for the coverage of neighboring base stations.

- Multiple compensating base stations energy saving: where multiple base stations compensate for the coverage of neighboring base stations.

Figure 13 illustrates the single base station compensation and multiple base stations compensation use cases.

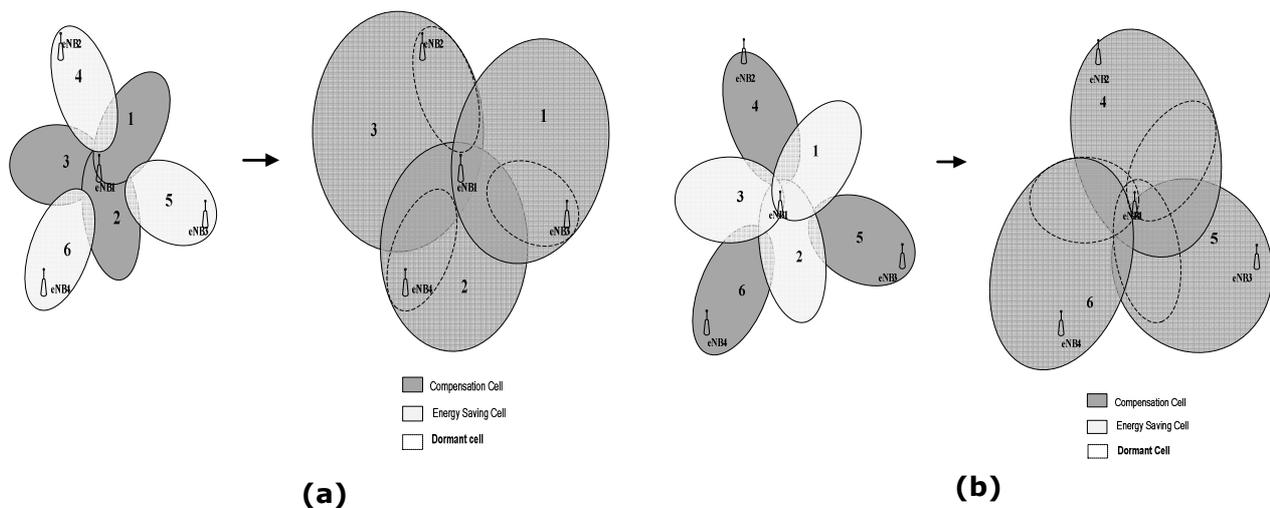(a)   (b)

Figure 13 : Compensation-based energy saving
(a) single compensating eNodeB, (b) multiple compensating eNodeBs.

When triggering the system state into energy saving state, the compensating base stations are transmitting on the maximum power and the non-compensating (energy saving base stations) are set OFF. So, the determination of the compensating base stations is an important task when considering such energy saving mechanisms. Currently two level network planning is foreseen as a means for determining the compensating base station set. The first network planning level is considered when the network is not in energy saving, then a lower target capacity is set for another network planning of the network when it is in energy saving state. This network planning effort may be relevant for homogeneous macro base station network but is clearly irrelevant for heterogeneous network deployment and especially for clusters of randomly deployed heterogeneous networks that are used to cover shopping

malls or campus areas (campus scenario). This scenario, presented in [5] as scenarios 2.2.2 and 2.3.10 will be the benchmark scenario for the techniques we propose in this study for the determination of the compensating base stations and the associated evaluations. In this study we propose methods for finding the compensating base station set for cluster of base stations randomly deployed in a coverage region C and connected to a central node by using the graph representation of its coverage as shown in Figure 14-(a) for the campus deployment and Figure 14-(b) for the graph representation of the campus.

The graph representation of the campus deployment is defined as follows: the nodes of the graph represent the nodes of the campus (eNB, HeNBs, etc). The edges of the graph representation are present if the neighboring base station is reported with sufficiently high Signal-to-Interference-and-Noise Ratio (SINR)(above a predetermined threshold) to the current, attachment base station of a given active user terminal UE in the scenario. For example two nodes i and j of the graph representation are connected by an edge if the SINR of at least one UE attached to the base station i, of base station j is above -10dB and the SINR reported by at least one UE attached to the base station j of the base station i is above -10dB. Another possible method for the graph construction is to link the two nodes i and j by an edge if the downlink interference level seen in the two base stations is above a predetermined threshold. Intuitively, this graph representation is used to find the compensation set as the set of nodes that is a neighbor to every node in the graph representation, so the SINR of the UEs deployed randomly in the coverage region of the campus are optimized even when the energy saving base stations are set OFF. So, the compensation base station set is equivalent to a dominant set in the campus.

(a)

(b)
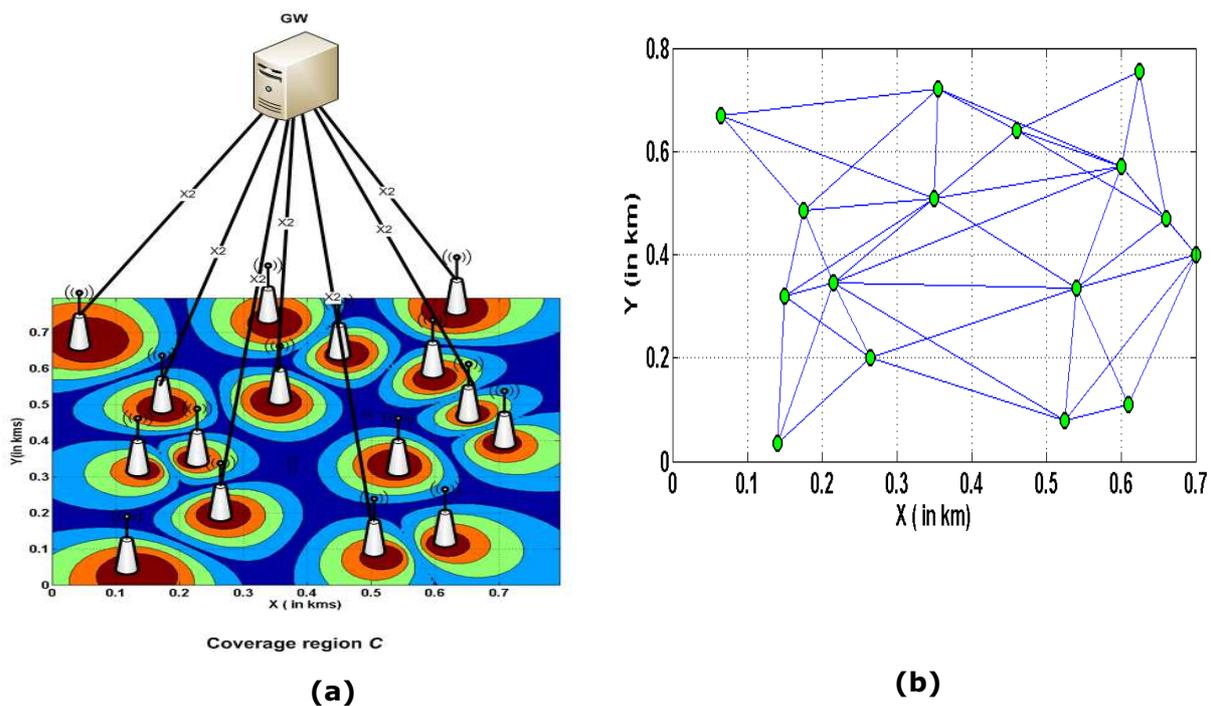
Figure 14 : Campus deployment scenario and coverage map (a) and corresponding graph representation (b)

The dominant set may be either a set of neighboring nodes in the graph or a set of independent, non-neighboring nodes such that every node in the graph is reachable from the dominating set nodes. This definition leads to the following dominant set classes:

1. Minimum connected dominant set (CDS): where the nodes of the dominant set are allowed to be connected by an edge in the graph, i.e. can be neighbors.

2. Maximum independent set (MIS): where the nodes of the dominant set are not connected in the graph i.e. the nodes of MIS set are not allowed to be neighbors. .

In Figure 15, examples of the two dominant sets are shown for the graph of Figure 14 above: the blue nodes are CDS nodes and the red nodes are MIS compensating nodes. CDS and MIS based compensation sets correspond to the single and multiple compensating base stations sets.
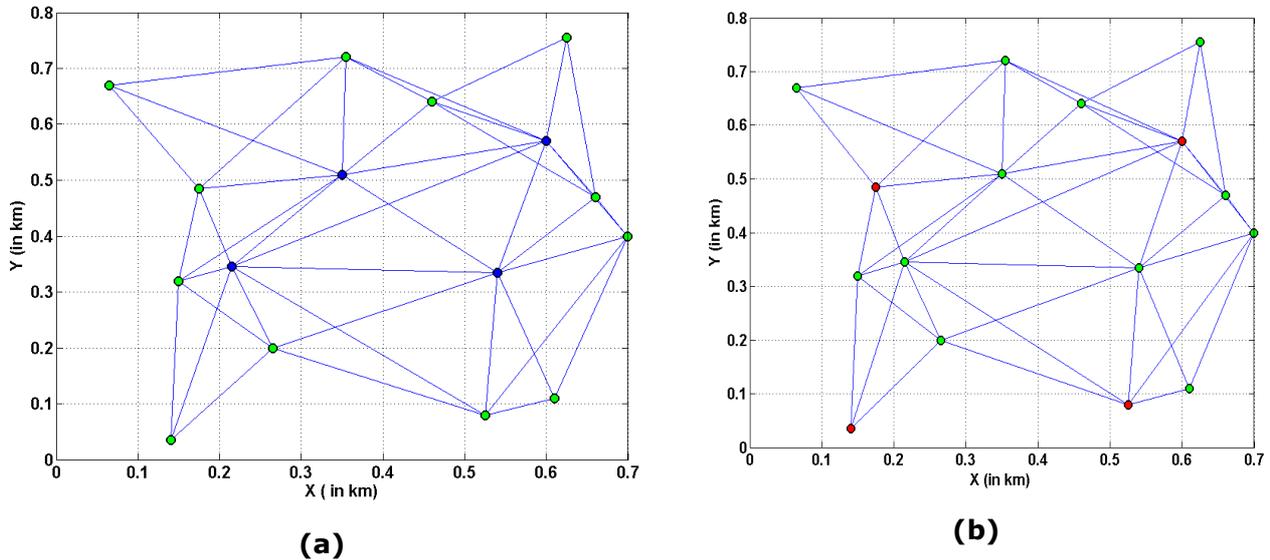

(a)

(b)

Figure 15 : Dominant set based compensating base stations (a) connected dominating set compensation (b) maximum independent sets compensation.

We have presented in the Sharing deliverable [11] initial spectrum efficiency results and sum power savings for campus of Pico base stations, showing basically that compensation based on MIS sets improves both performance and sum radiated energy in the average for the campus scenario deployment. In this report we will present energy saving and throughput gains evaluations for homogeneous campus of pico base stations, i.e. all the base stations of the campus are assumed to be pico base stations. The power consumption of the base stations is given by the power consumption model from the EARTH FP7 project [1][2]. The simulation assumptions are described further in the section below.

2.2.1   Simulation assumptions

2.2.1.1   *Campus simulation assumptions*

The campus scenario model used in the validation of the compensation based energy saving is a square region of 300x300m, 600x600m and 1000x1000m where a group of 4, 10 and 14 pico base stations are randomly deployed. The radio parameters of the pico base stations are obtained from [8] and are shown in Table 2 :

Table 2 : Radio parameters of the pico base station campus.

Parameter	Value
Maximum Tx power	30dbm
Bandwidth of the transmission	10MHz
Antenna gain	5dbi
Pathloss parameter	a=140.7
	b=36.7
Noise density	-174dbm/Hz
Noise factor	5dB
Shadowing	NA

The scheduling used in the simulations is round robin scheduling. Furthermore, 50 physical resource elements are assumed per pico base station with elementary resource bandwidth of 180 kHz. The scheduling and the elementary resource bandwidth will be used for the calculations of the system throughput and the load of the pico base stations within the campus. The pico base stations are assumed to be deployed outdoors and they are assumed to be accessible for all users. Finally, the signal-to-interference ratio for the user terminal j at base station i is calculated by taking into account the neighboring base stations loads.

The power consumption model of the pico base stations is given in [1], [2]. The total power consumed by the pico base station P_c is equal to $P_0+\Delta_P P_r$ if $0<P_r<P_{max}$. P_r is the radiated power and P_{max} is the maximum transmit power of the pico base station. If $P_r=0$, then $P_c=P_{sleep}$.

The rest of the parameters of the model are listed in Table 3 for the different types of base stations.

Table 3 : Power consumption model parameters

BS type	P_{max} [W]	P_0 [W]	Δ_p	P_{sleep} [W]
Macro	20	130	4.7	75
Micro	6.3	56	2.6	39
Pico	0.13	6.8	4	4.3
Femto	0.05	4.8	8	3.9

2.2.1.2   Dominating set calculation assumptions

Finding the dominating set for general graph topologies is an NP-hard combinatory optimization problem, i.e. it is difficult to find polynomial time algorithm for solving the compensation based optimization problem. In this study we focus on centralized greedy strategy based algorithms for the determination of the dominating set. The greedy strategy consists of determining recursively the nodes with the highest number of neighbors (degree) in the graph, adding them to the dominating set and excluding their direct neighbors from the search in the next iteration.

The algorithm is centralized and is performed in the campus gateway (GW), based on the available measurements and topology information already used for the construction of the graph.

2.2.2   Simulation Results

We first present examples of throughput Cumulative Distribution Function (CDF) for a campus of 10 pico base stations, deployed in a region of 600x600m. User terminals are randomly and uniformly dropped in the coverage region until the base stations of the campus achieve an average load of either 30% or 60%.

Figure 16 shows the obtained throughput CDF where the performance of the campus of pico base stations in energy saving state is compared with the nominal performance of the campus, i.e.,when the pico base stations are transmitting at their maximum transmit power.

(a)

(b)
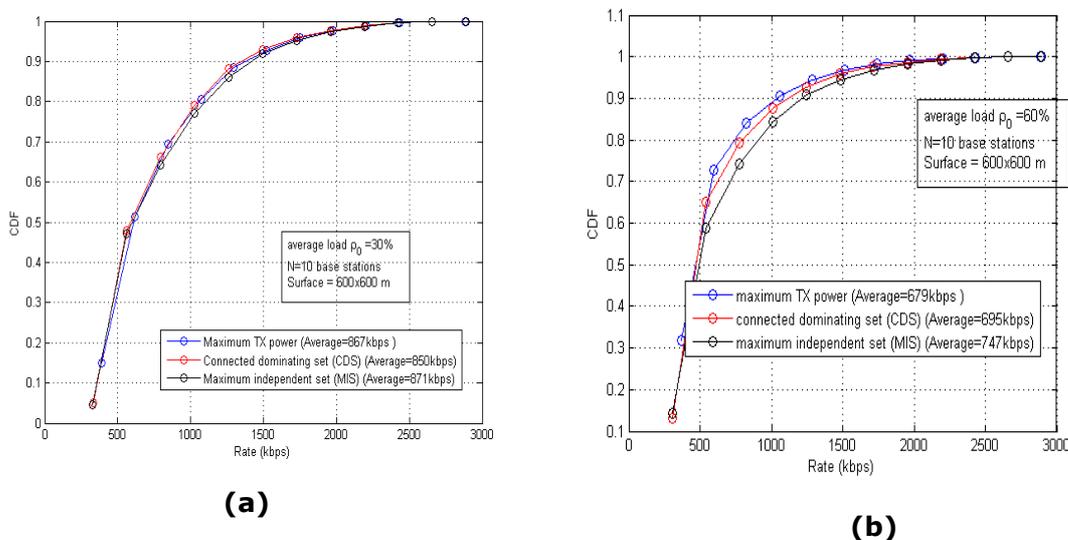
Figure 16 : Throughput CDF for 30% average load (a) for 60% average load (b).

It is seen from the results that throughput performance for the campus in energy saving state is similar to the nominal system performance for low average load. When the campus is highly loaded, it is seen that MIS-based campus performance improves the nominal performance by 21%. CDS-based campus improves the nominal throughput performance by 8%.

Figure 17 shows the CDF of the average radiated power gain, i.e. the average number of base stations that were set OFF for each campus configuration (pico base stations positions), for MIS- and CDS-based policies and for 1000 configuration trials.
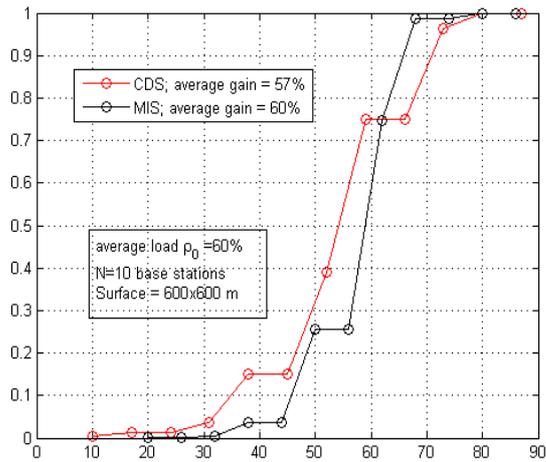
(a)
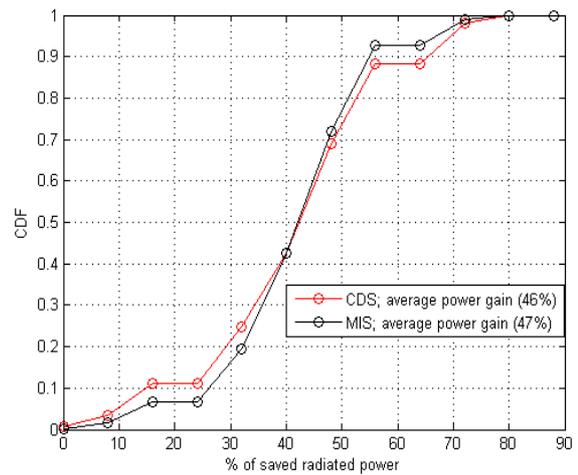
(b)

Figure 17 : (a) Radiated power gain CDF for 30% average load. (b) radiated power gain for 60% average load.

Results reported above show that, on the average, the MIS based compensation based energy saving has the best performance for both high load and low load scenarios. This maximum average radiated power gain is 47% for low load and 60% for medium loaded campus situations.

Figure 18 illustrates the average throughput and the average radiated power gain as functions of the average loads, for 4 base stations and for different campus sizes, i.e. (300x300m, 600x600m and 1000x1000m).
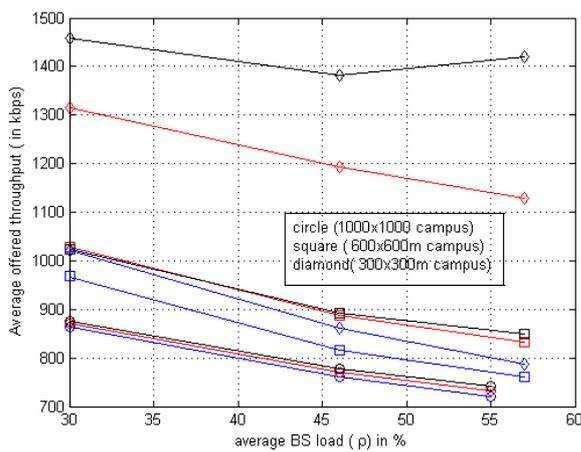
(a)
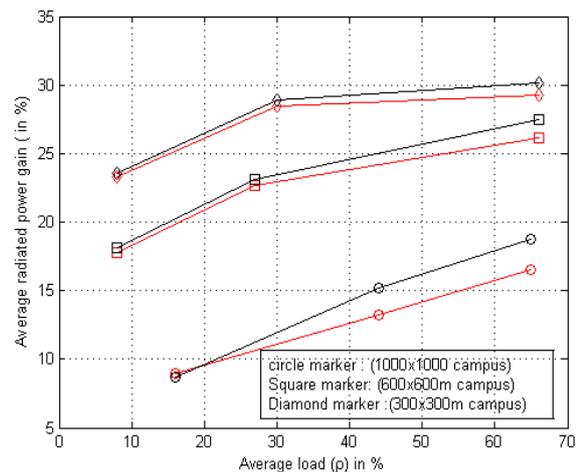
(b)

Figure 18 : Average throughput as function of the average load (a) Average radiated power gain as function of the average load (b).

The results show that the average throughput performance is similar for both nominal systems (all the nodes transmitting at maximum power) for CDS energy saving and MIS based energy saving for large campus (1000x1000m). For medium and small campus deployment, it is shown that the throughput of the campus in energy state is higher than the nominal system performance. This is because the deployment in this case is interference limited. For large campus, the maximum average radiated gain is around 18%. For interference limited campus, it is seen also that the maximum radiated power gain is around 27% for 600x600m and at maximum 30% for the 300x300m campus.

The results show that for reasonable campus setting (600x600m), it is possible to achieve 18% of radiated power gain through MIS based compensation strategy. For interference limited campus scenarios, both throughput improvement and radiated power gains are possible with at maximum 30% radiated power gain.

Starting from the results of Figure 18 and using EARTH's power consumption model, it is possible to evaluate the average power consumption gain of compensation based techniques. In order to evaluate the energy saving gain, traffic evolution through time was considered, i.e. evolution of the average base stations loads in time.

In Figure 19 we present the evolution of the traffic evolution over 24 hours for a typical campus deployment.
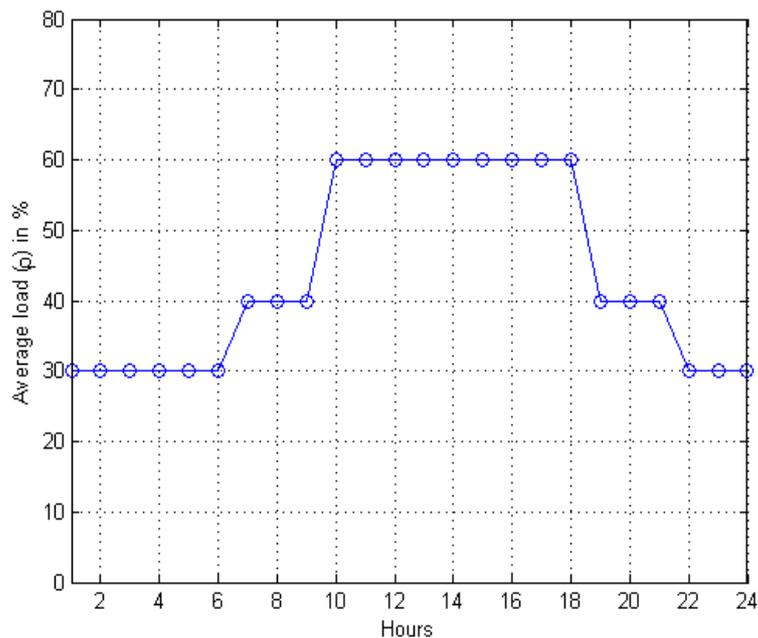
Figure 19 : Average campus load evolution over time.

This average load profile is only a test profile used for the evaluation of the energy saving gain, more realistic load evolution profiles will be considered in future work as well as average energy saving over various load profiles. Figure 19 shows the average throughput and energy saving for the traffic profile of Figure 19 for small, medium and large campuses, plotted as a function of the number of the base stations in the campus.

(a)   **(b)**
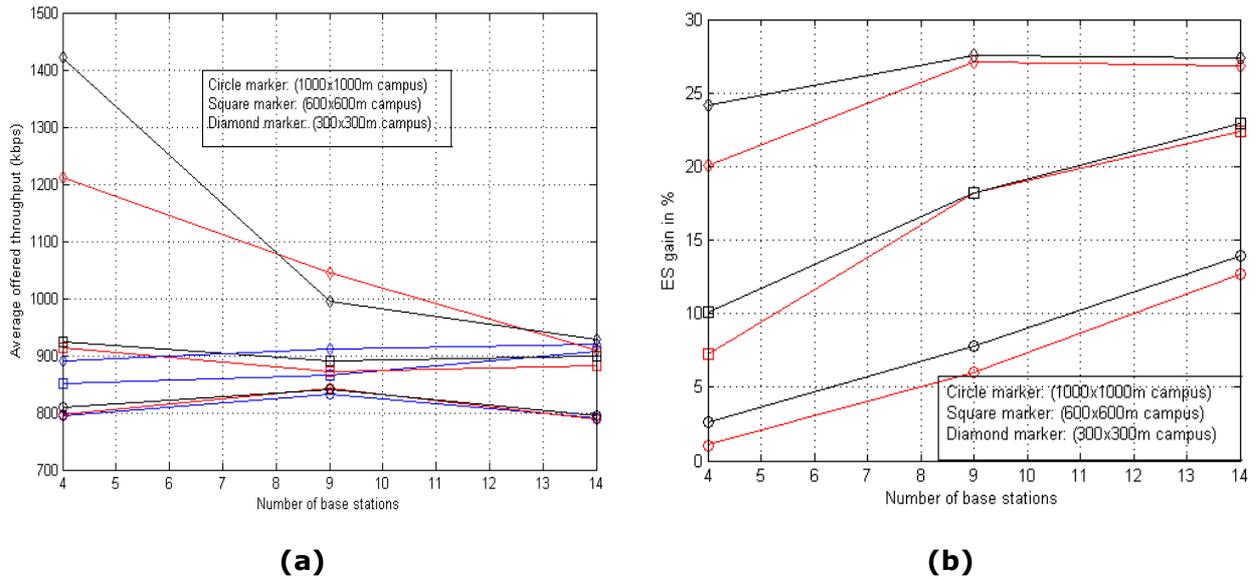
Figure 20 : (a) Average throughput as function of the number of base stations (b) Average radiated power gain as function of the number of base stations.

The results show that the average energy saving gain for the different techniques is increasing with the number of the bases stations in the campus. For a large campus, it is seen that the energy saving gain ranges from 3% to around 15% for MIS based ON/OFF energy saving technique while CDS technique achieves lower energy savings (1% to 13%). The average throughput for the large campus scenario is similar for the different techniques and is very close to nominal system performance. For medium and small campuses, the average throughput may be improved. For highly interference limited campus scenario (300x300m) it is seen that this improvement is at maximum 55%. The maximum corresponding energy saving gain is around 28%.

2.2.3   Conclusion and future work

In this section, we presented an update on compensation based ON/OFF energy saving where the compensating base stations are determined from dominant sets calculated over a graph representation of the campus. The performance of compensation based ON/OFF energy saving, based on two dominating set constructions is compared with the performance of the nominal system where all the base stations of the campus are transmitting at their maximum power. The energy saving gain for large campus is at maximum 15% for large campus scenario and MIS based compensating base stations is showing the best performance results.

For medium and small campus scenarios, it is shown that the compensation based ON/OFF state improves the average throughput by at maximum 55% and the maximum energy saving gain is around 28% for MIS based technique.

The results show that the proposed ON/OFF energy saving technique based on dominant sets constructed from the campus graph is effective and may improve for the cases of highly interference limited scenario the average throughput offered to the user terminals. Another important result from the study is that the maximum independent set based energy saving technique achieves the best energy saving efficiency as well as performance improvement.

As future work, we are planning to study the impact on the load of the compensating base stations so to discuss and investigate combined ON/OFF energy saving and load balancing for the homogeneous campus scenario.

The results will be then extended in the HetNet campus deployment use case.

2.3   Dynamic cell ON/OFF power saving

This section is in charge of measuring the theoretical power saving derived to perform smart BS ON/OFF switching in those periods where the total amount of traffic served by the network is such that it can be satisfied with just a subset of active BSs. The main scenario of application of this method is in dense urban areas, where a great amount of cells are deployed so as to be able to handle peak hour constraints. It is applicable to the case where just macro cells are used, and it has special interest in areas where small WiFi cells are used to cover hot points where users tend to concentrate in certain periods of the day. An optimum management of cell status may result in important energy savings, while satisfying user demands and maintaining user QoS levels. Depending on the scenario used (just macro cells or macro cell with WiFi small cells), the following assumptions are made:

- In case just macro cells are considered, the assumption considered is that all macro cells are of the same size (a constant ISD). When a savings layout is applied, just the central macro cell remains active, switching off the surrounding cells (for instance, in a 1/7 scheme, the central cell is kept active and the 6 surrounding BS are deactivated). It is assumed that the central cell that is still active has the capability to increase its emitted powers so as to cover the whole area of all deactivated macro cells.

- On the other hand, the second scenario covers one big macro cell running in background and several WiFi small cells spotted inside the macro cell area deployed to cover hot points. In that case, whenever the traffic covered by one small cell can be absorbed by the macro cell, it is deactivated.

The way these ON/OFF orders are managed can be studied from two opposed perspectives: centralized and decentralized control of the network. In the former case, the network is managed from a central entity which is aware of all cell status and traffic handled at each moment. Depending on the particular situation of a predefined area, the system is able to send ON/OFF orders. This centralized mode is especially appropriate in case of a scenario where a macro cell coincides in the same area where a number of WiFi small cells are deployed. Fixed clusters of cells are pre-defined in the network. The central cell of the cluster will remain active and the adjacent cells switched off in case the network detects low traffic loads. In the latter case (decentralized), each cell is aware of the status of all neighboring cells. In case a cell estimates that it is able to handle not only its traffic but also the traffics of nearby BS, it can perform directly (without a central cluster or network entity doing this in its behalf) the layout change. This management scheme is very useful in just macro cell scenarios, although this just macro cell scenario can be also modelled with the centralized method.

The next important topic to address when defining the scenario covered by the envisaged simulations is how to approximate real world cellular deployments. For this study, hexagonal and Manhattan cell layouts are considered. Depending on the layout used, the antenna assumptions (omni directional or trisectorial coverage) and the actual scenario (just macro cells or small cells inside a macro), only a subset of layouts will be available to use. We denote by X/Y the ON/OFF scenario where only X cells can be kept active while the rest surrounding Y cells are switched off. In an attempt to reduce the simulation parameters to the most common real world layouts, the following ones are considered in this study:

- 1/2 and 1/3 layouts for Manhattan deployments (Figure 21)

- 1/4, 1/7 and 1/19 layouts for hexagonal cells and omni directional antennas (Figure 22)

- 1/9 layout for hexagonal cells and trisectorial antennas (Figure 22)

The algorithm programmed and run in the simulations will use at least one of these layouts. It is also possible to use combination of several layouts (for instance 1/7 and

1/19 in omni directional hexagonal cells) to run dynamic simulations in a two-step approach, using the first layout for periods where traffic levels are low but not near the minimum and the second layout for very low traffic intervals (mainly during the night).
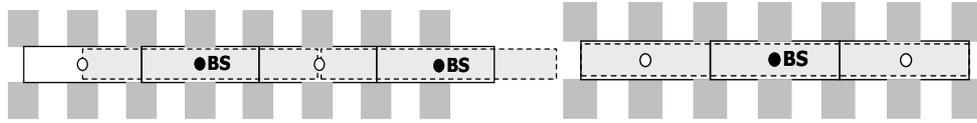

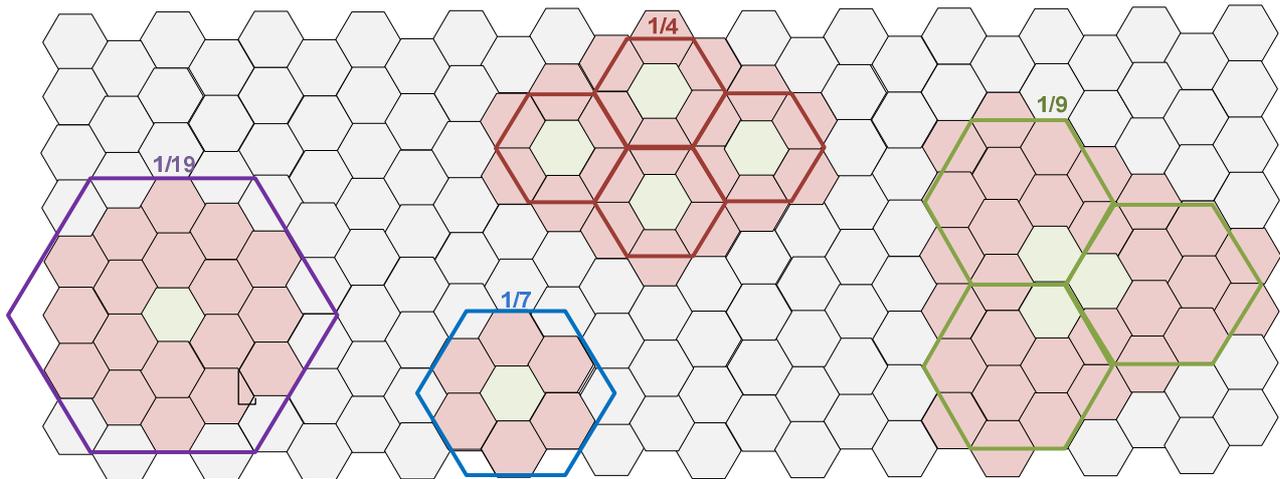Figure 21 : Layouts for Manhattan deployments, 1/2 (left) and 1/3 (right).

Figure 22 : Layouts for hexagonal cells (from left to right 1/19, 1/7, 1/4 and 1/9).

At this point, the scenario and layouts to be applied are fixed. There is an important parameter related to the layout used which needs to be carefully set up so as to enable efficient performance of the method: the threshold used to trigger the ON/OFF change. As the traffic in real cells is quite bursty (deep traffic valleys or peaks may occur unexpectedly) it is important to fix thresholds so as to avoid multiple ON and OFF orders when the traffic is near the threshold. The solution used to avoid these situations is forcing thresholds to be applied 20% above (for OFF orders) or below (ON orders) of the theoretical value. For example, when using 1/2 layouts, the algorithm will not switch off cells when the overall traffic is below 50%, but when it goes under $0.8 \cdot 0.5 = 40\%$. In addition, linear predictors are considered so as to forecast the network behavior and determine if the traffic is likely to remain under the threshold in the next 30 minutes or it may rise, forcing us to re-activate the cells in the near future. This way inefficient ON/OFF orders can be avoided, preventing the network from being unstable.

With all aforementioned considerations, the scenarios in which the ON/OFF algorithm will be applied are fully characterized. There is just one parameter left to be defined in the system, which is the traffic model used to simulate the daily load of the network. Three different approaches are used:

- Theoretical curves with linear approximation. Flat maximum zone for peak hours and flat minimum region for valley hours with linear connections with slope "d". In this case, the amount of time where the cells can be deactivated is between τE and τM where $L(\tau E) = L(\tau M) = a$ (point where a base station can absorb the traffic from the ones that are switched off) with $\tau E < \tau M$ (Figure 23).
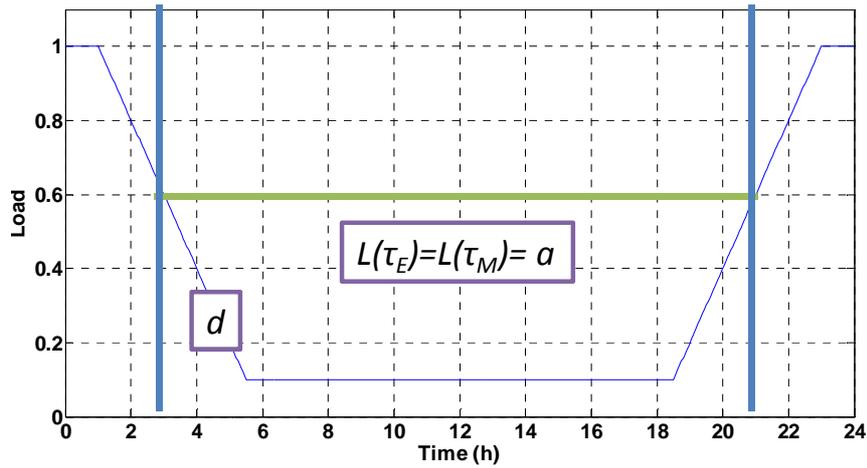
Figure 23 : Linear approximation.

- Theoretical curves with gaussian approximation. In this case, two peak hours are modeled (lunch and dinner times). The curve is obtained through the combination of two different gaussian with same mean and different standard deviation (being σ'=3σ, see Figure 24).
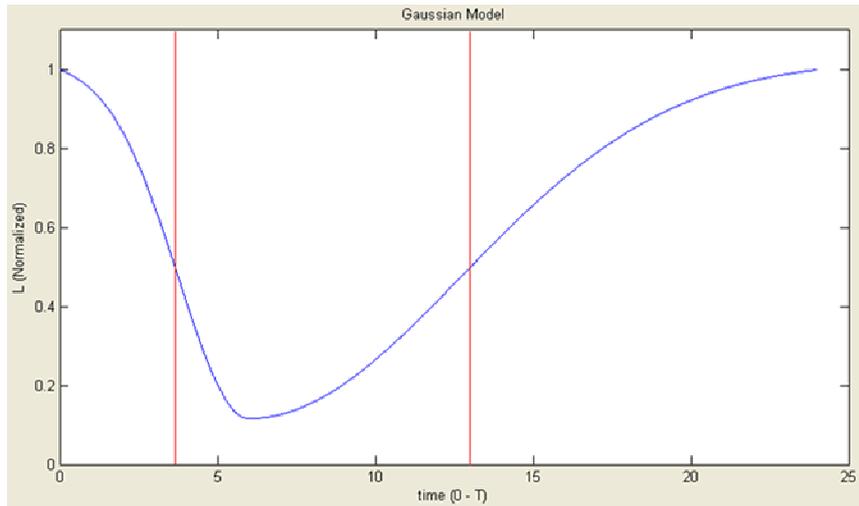
Figure 24 : Gaussian approximation.

- Real data using FP7 EARTH models [54]. Finally, as the most realistic approach, EARTH project mean and real data patterns are used. They consider several scenarios. In this case, macro, micro and pico models for urban scenarios are used together with macro rural models. As these curves are averaged, some additive noise has been introduced so as to model real daily load patterns, as seen in Figure 25.
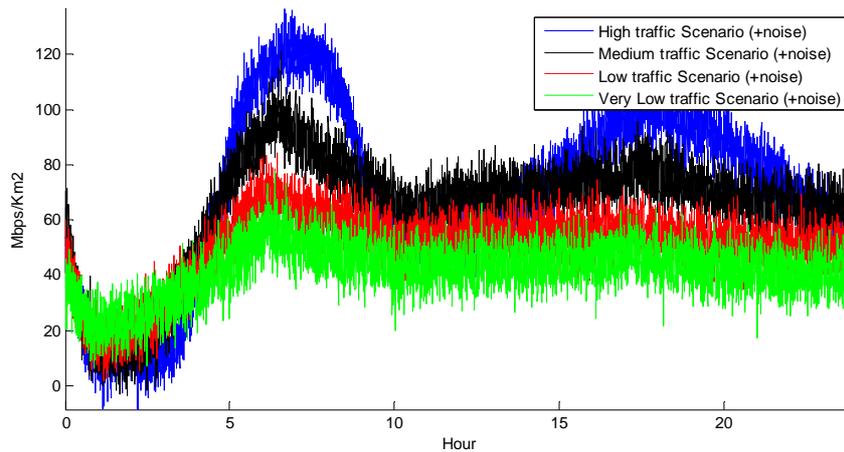
Figure 25 : EARTH traffic model and real curve introduced in the system.

Combining all potential situations based on the previously defined parameters (only macro cells or macro with small cell scenarios, centralized or decentralized management, 1/2, 1/3, 1/4, 1/7, 1/9 or 1/19 layout, use of traffic predictors and theoretical or real curves), Figure 26 depicts all results obtained. For each case analyzed, 1.000 simulations were run. The result presented is the mean value of energy savings (in percentage) among all simulations.
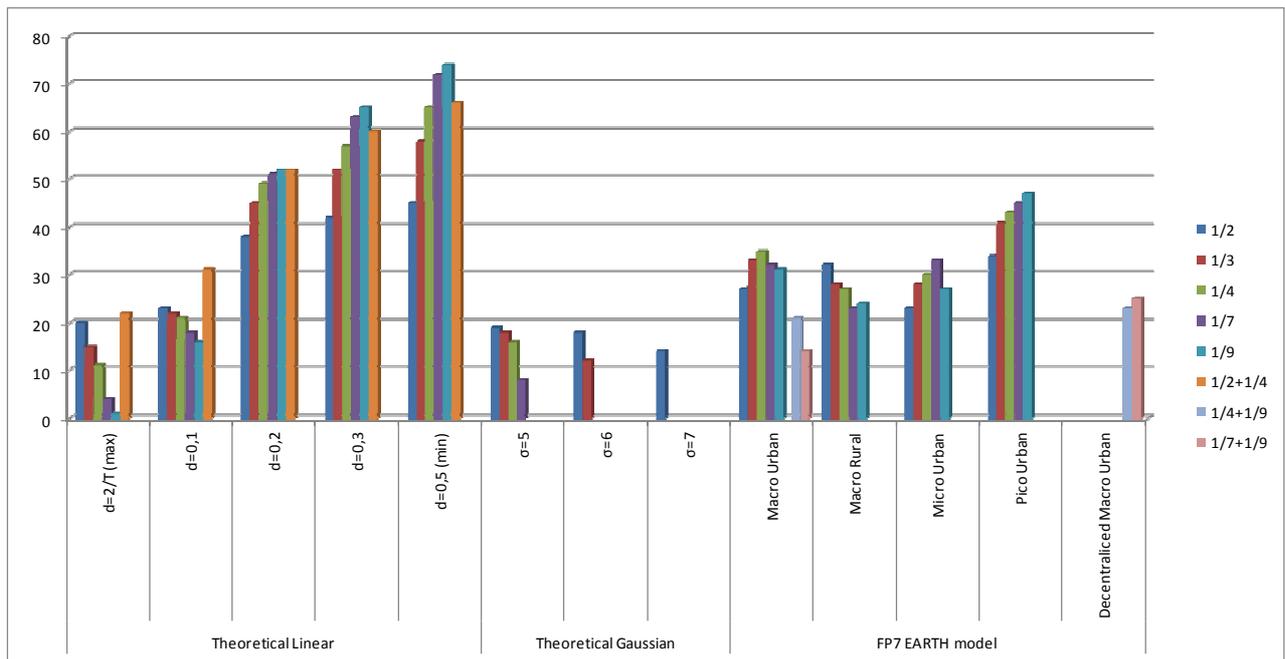
Figure 26 : Summary of simulation results depending on the method and scheme applied.

As a conclusion, it is possible to state that making use of ON/OFF schemes is always a good idea so as to save energy, with no degradation in user QoS in any case. As Figure 26 shows, theoretical simulations (linear or Gaussian) set the upper bounds of reachable gains when using the ON/OFF algorithms. With the modeled traffic curves, up to 40% savings can be achieved depending on the scenario used. Using decentralized techniques provides higher gains than centralized approaches. In the same sense, combining multiple layouts provide

higher gains, although, depending on the network deployment, it is not always possible to apply all of them together. Finally, the use of traffic predictors has been considered as not very useful, as the cell ON-OFF status changes can be performed in less than a minute, so, establishing proper thresholds make the use of predictors irrelevant.

2.4 Opportunistic sleep mode strategies for small cells

2.4.1   Uncoordinated case

The design of energy-efficient mechanisms is one of the key challenges in emerging wireless small cell networks. In this subsection, a novel approach for opportunistically switching base stations on and off to improve the energy efficiency in wireless small cell networks is proposed. The proposed approach enables the small cell base stations to optimize their downlink performance while balancing the load among each other and satisfying the users' QoS requirements. The problem is formulated as a non-cooperative game among the base stations that seek to minimize a cost function which captures the tradeoff between energy expenditure and load. To solve this game, a distributed learning algorithm is proposed where the base stations autonomously choose their optimal transmission strategies.

Consider the downlink transmission of a heterogeneous wireless network with a set of base stations $B = \{1, .., B\}$. The set B consists of a set of small cell base stations (SBSs) $B_S = \{1, …, B_S\}$ underlaid on a macro cellular network with a set of macro base stations (MBSs) $B_M = \{1, .., B_M\}$. Without loss of generality, we assume that the MBS 1 is located at the origin of the two-dimensional network layout and we let x be any location on the plane measured with respect to the origin. Moreover, let L_b be the coverage area of base station b such that any given user equipment (UE) at a given location x is served by base station b if $x \in L_b$. Let S_b be the transmission indicator of base station b such that $S_b = 1$ indicates the active state while $S_b = 0$ reflects the idle or sleep state. Here, we assume that, in active state, each base station will serve all UEs in its coverage area. From an energy saving perspective, some base stations might have an incentive to switch into sleep mode. Note that during the idle state, a base station consumes a non-zero power to sense the UEs in its vicinity. Consider that base station b uses a Cell Range Expansion Bias (CREB) ζ_b to absorb additional UEs (expand its coverage area L_b) along the transmission power P_b. The concept of CREB has been proposed in small cell networks due to the disparate cell sizes between different base stations. Since the CREB concept is used only by SBSs, we let $\zeta_b = 0$. Moreover, we assume that all base stations transmit on the same frequency spectrum (i.e. co-channel deployment). Therefore, the received SINR from base station b at location $x \in L_b$ is given by:

$$\gamma_b(x) = \frac{P_b S_b h_b(x)}{\sum_{\forall b' \in B/b} P_{b'} S_{b'} h_{b'}(x) + N_0} \quad (1)$$

Further, the data rate at a given location x from base station b is given by:

$$R_b(x) = \omega \log_2\big(1 + \gamma_b(x)\big) \quad (2)$$

We assume that the UEs connected to base station b are heterogeneous in nature such that each UE has a different QoS requirement based on its individual packet arrival rate. In this respect, let $\lambda(x)$ and $1/\mu(x)$ be the packet arrival rate and the mean packet size of any UE at location $x \in L_b$. The data rate offered to the UE at location x from base station b is $R_b(x)$ and thus, the load density of base station b becomes

$$\varrho_{\mathbf{b}}(\mathbf{x}) = \frac{\lambda_{\mathbf{b}}(\mathbf{x})}{\mu_{\mathbf{b}}(\mathbf{x})R_{\mathbf{b}}(\mathbf{x})} \quad (3)$$

Consequently, the base station load ρ_b of base station b is given by:

$$\rho_{\mathbf{b}} = \int_{\mathbf{x} \in \mathcal{L}_{\mathbf{b}}} \varrho_{\mathbf{b}}(\mathbf{x})\mathbf{dx} \quad (4)$$

Here, for each base station $b \in B$, we define a cost function that captures both energy consumption and load, as follows:

$$\underset{\rho}{\text{minimize}} \sum_{\forall b \in B} \Gamma_b(\rho)$$

$$\text{subject to} \quad 0 \leq \rho_b \leq 1, \forall b \in B$$

$$P_b^{Total}(\rho) \leq P_b^{Max}, \forall b \in B \quad (5)$$

where $P_b^{Total}(\rho) = 1/v_b \times P_b^{work} + P_b^{Base}$ and $P_b^{work} = \rho_b P_b$. $\Gamma_b(\rho)$ is the cost function that captures both energy consumption and load. At time instant t, each base station b advertises its estimated load $\rho_b(t)$ via a broadcast control message along with its transmission power $P_b(t)$ and CREB $\zeta_b(t)$. Considering both the received signal strength and load, at time t the UE at location x connects to base station $b(x,t), x \in L_b(x,t)$, according to the following UE association rule:

$$b(x,t) = \arg\max_{b \in B}\{(\widehat{\rho}_b(t) + \epsilon_b)^{-\delta}P_b^{Rx}(t)\} \quad (6)$$

$\hat{\rho}_b(t)$ is the estimated load of BS b, $P_b^{Rx}(t)$ is the received power from BS b. Due to fact that the base stations need to estimate their loads beforehand, the estimations must accurately reflect the actual load. In order to obtain an accurate estimation for the load of the base station b, we compute the load estimation $\hat{\rho}_b(t)$ at time t based on history as follows

$$\widehat{\rho}_b(t) = \widehat{\rho}_b(t-1) + v(t)\big(\rho_b(t-1) - \widehat{\rho}_b(t-1)\big) \quad (7a)$$

where $v(t)$ is the learning rate of the load estimation. Leveraging different time-scales, we assume that $v(t)$ is selected such that the load estimation procedure is much slower than the UE association process.

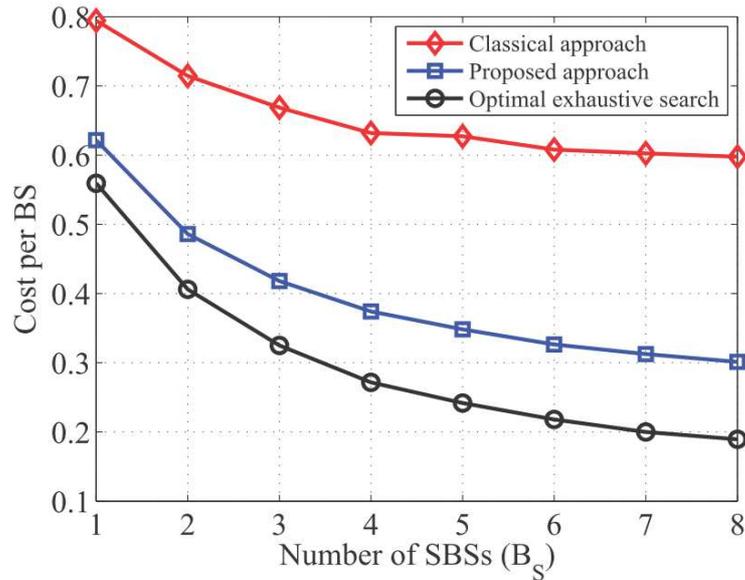The innovation presented above is related to scenario 2.3.8 defined in[5].

Figure 27 : Variation of the cost per base station with respect to the number of SBSs. The number of UEs is fixed to 100.

Figure 27 shows the average cost achieved per base station as the number of SBSs varies. The cost captures the tradeoff between load and energy consumption. As the number of base stations increases, the total energy consumption of the network increases. However, the load from a fixed number of UEs is distributed among the various base stations. Therefore, the energy required to handle the load decreases per base station resulting in a cost decrease per base station for all three approaches as seen in Figure 27. In the proposed approach, the base stations switch to a sleep state when there are no UEs in their vicinity. Consequently, Figure 27 shows that the proposed approach exhibits a considerable cost reduction compared to the classical model (where all base stations are ON). We also see that for a single SBS, the proposed approach exhibits a cost reduction of 21.8% and it reaches up to 49.5%, relative to the classical approach, with eight SBSs. Figure 27 also shows that the difference in the average performance between the proposed approach and the optimal exhaustive search solution reaches about 18.8% at $B_S = 8$ SBSs. However, the optimal solution requires an exhaustivecentralized search which yields significant overhead. Indeed, the gap between the exhaustive search and the proposed approach is a byproduct of the uncoordinated decision making processes and the sells behavior of the players (base stations).

2.4.2   Coordinated case

Unlike the selfish case, here, small cell base stations coordinate with other neighboring cells in order to efficiently offload their own traffic prior to turning off. This requires a decentralized mechanism to cluster small cells into judiciously well chosen groups, within which user association and dynamic switch ON/OFF is carried out. The dynamic clustering is based on the well-known spectral clustering technique typically used in graph clustering. Spectral clustering exploits the connectivity and compactness of the network graph composed of small cell base stations. First, the graph Laplacian matrix is formed as L=D-S where D is the diagonal matrix with b-th diagonal element given as $\sum_{b'}^B S_{bb'}$ and the Gaussian similarity based on the distance between base stations b and b' is:

$$s_{bb'}^{d} = \begin{cases} \exp\left(\dfrac{-\| y_b - y_{b'} \|^2}{2\sigma_d^2}\right) if & \| \, y_b - y_{b'} \| \leq \varepsilon_d \\ 0 \end{cases}$$

$$(7b)$$

where y_b represent coordinates of the vertex (or base station) $b \in B$ in the Euclidean space. σ_d^2 controls the impact of neighborhood size. Furthermore, the distance-based similarity matrix S^d is formed using $s_{bb'}$ as the (b, b')-th entry. The base stations located far from each others have low similarities and as they come closer, similarities increase in which base stations are more likely to cooperate with each other. Furthermore, the distance-based similarity matrix Sd is formed using $S_{bb'}^{d}$ as the (b, b')-th entry. The number of clusters k is related to the eigenvalues of L and can be found as: $k = \arg \max_{i} (| \xi_{i+1} - \xi_i |), \qquad i = 1, ..., | B |$ where ς_i is the i-th smallest eigenvalue of L.

Unlike the static distance-based clustering, load-based clustering provides a more dynamic manner of grouping neighboring BSs in terms of traffic load. Moreover, the load-based similarity between BSs b and b' is given as follows:

$$s_{bb'}^{l} = \begin{cases} \exp\left(\dfrac{-\| \rho_b - \rho_{b'} \|^2}{2\sigma_l^2}\right) \\ 0 \end{cases}$$

$$(7c)$$

The final joint similarity S with $s_{bb'}$ as the (b, b')-th element is formulated as follows: $s_{bb'} = \left(s_{bb'}^{d}\right)^{\theta} . s_{bb'}^{l}{}^{(1-\theta)}$ with $0 \leq \theta \leq 1$ which controls the impact of the distance and the load similarities on the joint similarity. Once clusters are formed, the lightly loaded base stations within the cluster can be selected as the cluster head. The function of a cluster head is to coordinate the transmissions between the cluster members. Consequently, the entire load of the cluster is distributed between its members and such orthogonal resource allocation helps to mitigate intra-cluster interference. Due to the fact that base stations within a cluster have the ability to coordinate, the entire cluster can be seen as a single super base station which serves all the UEs within its vicinity. By enabling such redistribution of UEs between base stations within a given cluster, we can enable the base stations to switch OFF while ensuring UEs' QoS. The actual loads of each base station in a cluster is unknown before the downlink transmission takes place. Therefore, the UE offloading within the cluster is carried out with the objective of minimizing the estimated cluster load. This can eventually reduce the number of UEs served with low rates.
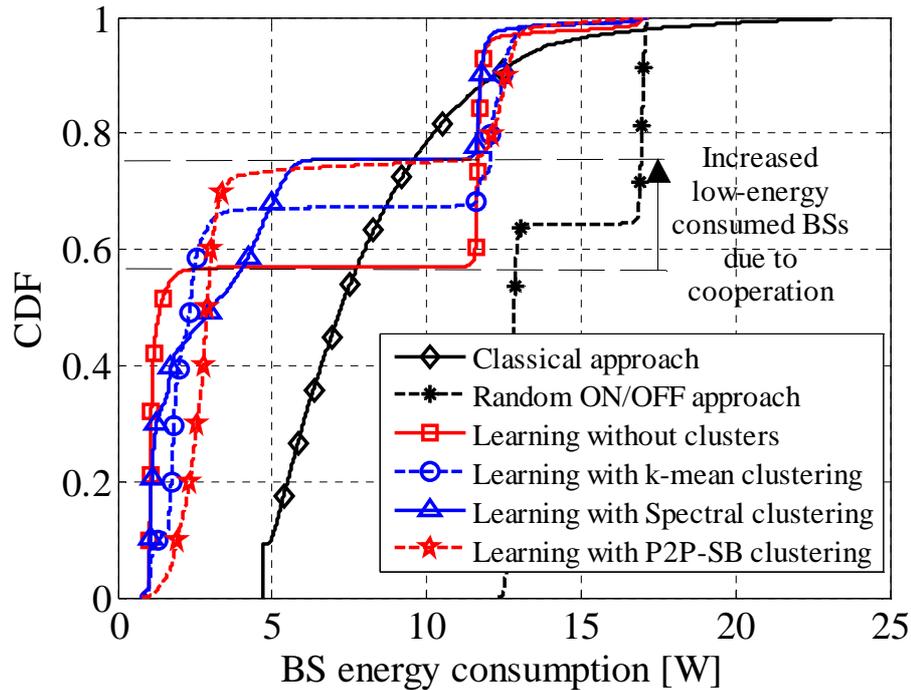
Figure 28 : CDF of the BS energy consumption for 10 SBSs, and 10-75 UEs.

In Figure 28, we plot the CDF of the base station energy consumption for 10 SBSs and 10–75 UEs. It can be shown that the random ON/OFF approach has the highest energy consumption. Although those base stations are allowed to switch OFF, the random switching OFF overloads the switched-ON base stations which results in a higher energy consumption for all the base stations and, a higher average energy consumption per base station is observed. Moreover, we can see that for the classical approach, the frequency of base stations having high energy consumption is much higher than in the cases with learning. Indeed, the proposed learning method allows lightly-loaded base stations to offload their traffic and switch OFF, thus yielding significant energy reductions. Coordination between clusters allows more base stations to switch OFF and, thus, the proposed learning approaches with clustering yield a larger number of base stations consuming less energy.

Figure 29 : Variation of the cost per base station with respect to the number of UEs for various clustering algorithms.

Figure 29 depicts the average cost per base station as the number of UEs varies. It can be shown that as the number of UEs increases, the load in the system as well as the base station energy consumption increases and thus, the average cost increases. However, it can be seen that the proposed learning approaches can reduce the cost by balancing the energy consumption and the handled load. Although the random ON/OFF approach manages to reduce the cost compared to the c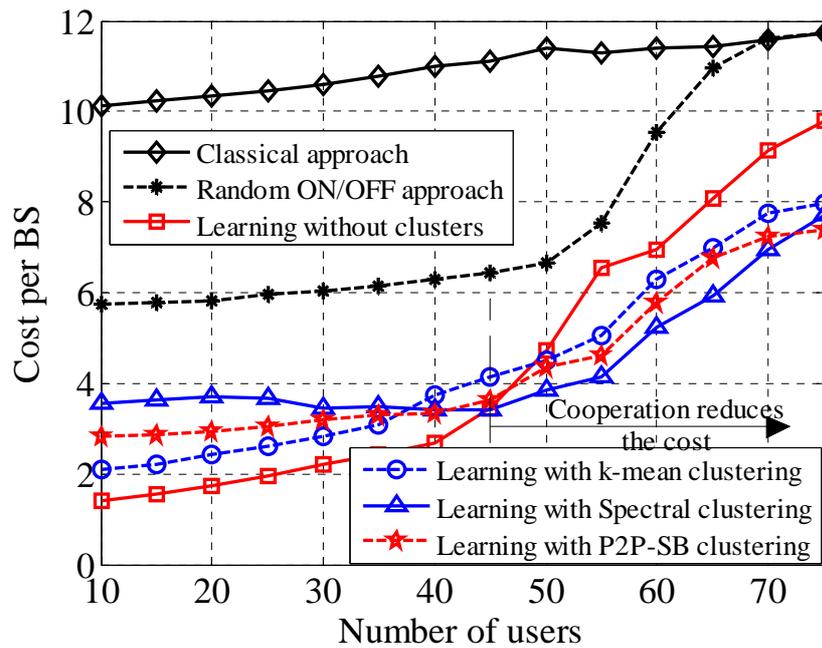lassical scenario, the cost reduction is not as high as compared to the learning approaches. The random ON/OFF behavior leads to overload the switched-ON base stations and thus, the energy saving at switched-OFF base stations has become insignificant due to the increased load and the energy consumption of overloaded base stations. As the number of UEs increases, the base stations have lesser opportunities to switch-OFF while satisfying the UEs' QoS and thus, the behavior becomes closer to the classical approach.

2.5 Energy efficiency of heterogeneous network using on/off small cells in real large scale environment

2.5.1 Introduction

As explained in the previous sections, the use of a dense small-cell layer to offload the macro layer is not justified all day long as the user traffic demand varies periodically between a minimum during the night and a maximum during busy hours [24]. However, the deployment of the small-cell scan only be permanent, unless making use of virtual small-cells (see SHARING deliverable D4.2[55]). One possible optimization is thus to switch off some of the small-cells outside busy hours while maintaining the user QoS. The aim of this study is to assess the possible energy gain under a simple small-cell switch-off mechanism depending on small-cell traffic load. The following sections describe the scenario under study, the traffic

model used to emulate the space- and time-variant user traffic demand, and finally provide simulation results and conclusions.

2.5.2   Scenario

The scenario relies on a typical macro network layout in the dense urban environment of Paris VII district, which is densified with co-channel small-cells (Figure 30). Macro eNode Bs (MeNB) are deployed over a larger area, on two rings (in orange and blue) around a central three-sector site (in red) in order to take into account realistic interference patterns. Small-cells are introduced into the network infrastructure to enhance the coverage and boost the capacity. They are uniformly deployed on urban furniture along the streets (at a moderate height of 6 m), leading to a quasi-constant Inter-Site Distance (ISD) of 100 m.

The framework integrates ICIC/eICIC techniques to mitigate interferences. In downlink, a static ICIC Fractional Frequency Reuse (FFR) scheme is considered for MeNB. It is complemented with a Time-Domain eICIC to enhance the cell-edge experience of small-cell users [25]. Main simulation parameters are summarized in Table 4.
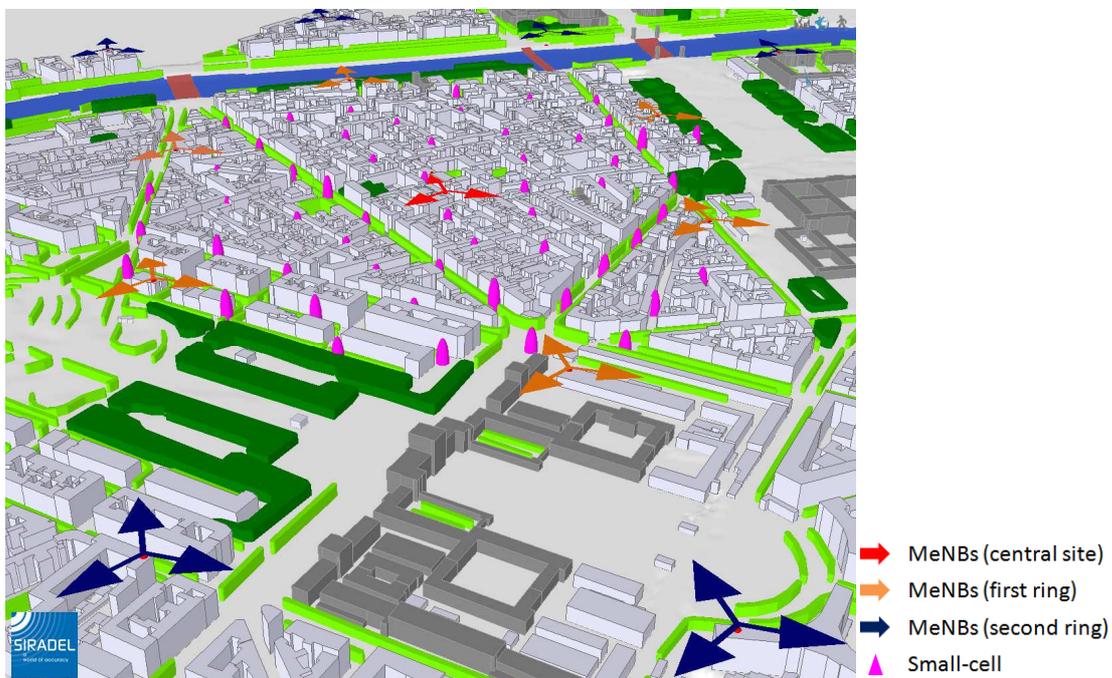

Figure 30 : Macro eNodeB and small-cell layer deployment in a real dense urban environment.

Table 4 Simulation parameters

System	LTE FDD 2x10 MHz. Central frequency: 2.6 GHz. Uplink/Downlink MIMO configuration: 2 x 2.
Macro-cell layout	Hexagonal site deployment: two rings around the central site, i.e. 19 sites corresponding to 57 cells Inter-site distance: 450 m. ICIC FFR scheme: 5% of total radio resources being allocated to each sub-band, re-use factor of 3. Average antenna height: 32 m above ground. Maximum total transmit power: 40 W. Antenna: directional, 14 dBi, 6° electric down-tilt, 32 m above ground.
Small-cell layout	Spectrum usage: co-channel. Uniform deployment of 55 SCs with average ISD of 100 m. eICIC parameters: Cell Range Extension (CRE) bias 9 dB, Almost Blank Subframe (ABS) duty cycle 12.5%. Maximum total transmit power: 5 W. Antenna: omnidirectional, 5 dBi, 6 m above ground.
User equipment	Antenna: omni-directional, 0 dBi, 1.5 m above ground. Downlink noise figure: 9 dB.

A square calculation zone of 1250 m side length and 5 m resolution is centered on the central macro site, representing UE locations at 1.5 m above ground. The analysis of the downlink LTE-A network performance relies on an extended version of the 3D coverage simulation tool presented in [26]. The user traffic is spatially distributed in 3D to take into account indoor users at different floors. Path-loss predictions are thus computed at different heights based on a real 3D environment representation and the Volcano technology (site-specific ray-based propagation models [27]. The metrics of interest in this study are:

- Macro eNodeB and small-cell traffic loads, linked to the space and time varying user demand,
- Service outage, which results from overloaded eNodeBs,
- Network power consumption, depending on eNodeB traffic loads and small-cell states (switched on or off).

Traffic loads are observed hourly in a reference scenario in which all small-cells are switched on. The small-cells to switch off are the ones having a traffic load below a given threshold. In this scenario this is triggered by a traffic load measurement done by each small-cell. As for the switch-on mechanism it is assumed as ideal, i.e. a small-cell switches on when it is about to support a traffic load above the given threshold. Switch-on mechanisms may be centralized or distributed and are not studied in this contribution.

The UEs located in the area covered by the switched off small-cells have to select either a neighbor small-cell or a macro eNodeB as the new serving base station. The new server logically observes a traffic load increase. The choice of the small-cell traffic load threshold is therefore important to avoid switching off too many small-cells and lose the small-cell layer offloading capacity. This is why the service outage is also observed hourly so as to check the user QoS under different thresholds. Finally the network power consumption is evaluated according to the model published in[28]with respect to eNodeBs maximum total transmit power and traffic load.

2.5.3   Traffic modeling

Observation of hourly eNodeB traffic loads is only relevant in conjunction of a space- and time-variant traffic modeling. The overall time variability is given by a time-variant mean active user density shown in Figure 31(left) that is typical of a daily traffic observed during week days [24]. The same downlink mean throughput demand of 130 kbps is assumed for all active users. Note that it is an average over all active users. Some may be only mailing, others downloading. This value is chosen so that the macro+SC network can absorb the peak hours traffic. The active user density is spatially distributed by exploiting Points of Interest (PoIs) provided by platforms such as Foursquare or Factual. PoIs together with a rule-of-thumb hourly repartition make it possible to distribute the population, and thus the user traffic demand, in space and time. For example, more people tend to be near transportation accesses at 8 AM whereas at 10 PM more people tend to be near bars and night clubs. Traffic hotspot locations thus undergo dynamic evolution during the day. The PoI-based distribution is complemented by a time-variant indoor ratio illustrated in Figure 31 (right), which models the proportion of indoor users compared to the total number of users. Indoor users are uniformly distributed in building floors. Figure 31 gives examples of active user density maps at 9 AM and 9 PM.
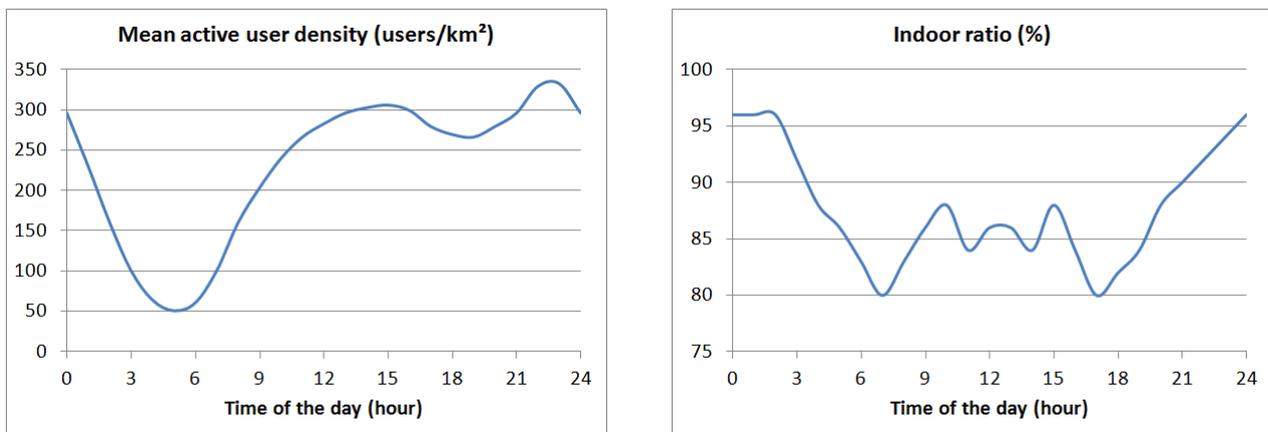
Figure 31 : Mean active user density (left) and indoor ratio (right).

2.5.4   Simulation results

In the following, we compare the results of different scenarios. The *Macro-only* scenario shows the capacity issue of such a network with the user traffic demand at busy hours. The *Macro+SC (ON)* scenario is the reference one with all small-cells switched on. In *Macro+SC (OFF p%)* scenarios, small-cells having less than p% of traffic load in the reference scenario are switched off, with p ranging from 1% to 5%.

Figure 32 summarizes this mechanism: the traffic load per small-cell is depicted hourly with respect to the different thresholds under study. We clearly observe that most of the small-cells will be switched off during night, but also that some of them supporting less traffic will be switched off for a longer time. Some small-cells may be switched off twice a day due to the traffic demand slightly decreasing between 3 PM and 7 PM, or because of traffic hotspots shifting.

Figure 33 shows precisely the active user density map and corresponding small-cells switched on or off with the 3% TL threshold at 9 AM (left) and 9 PM (right). We observe that small-cells are switched on in the vicinity of the traffic hotspots as expected.
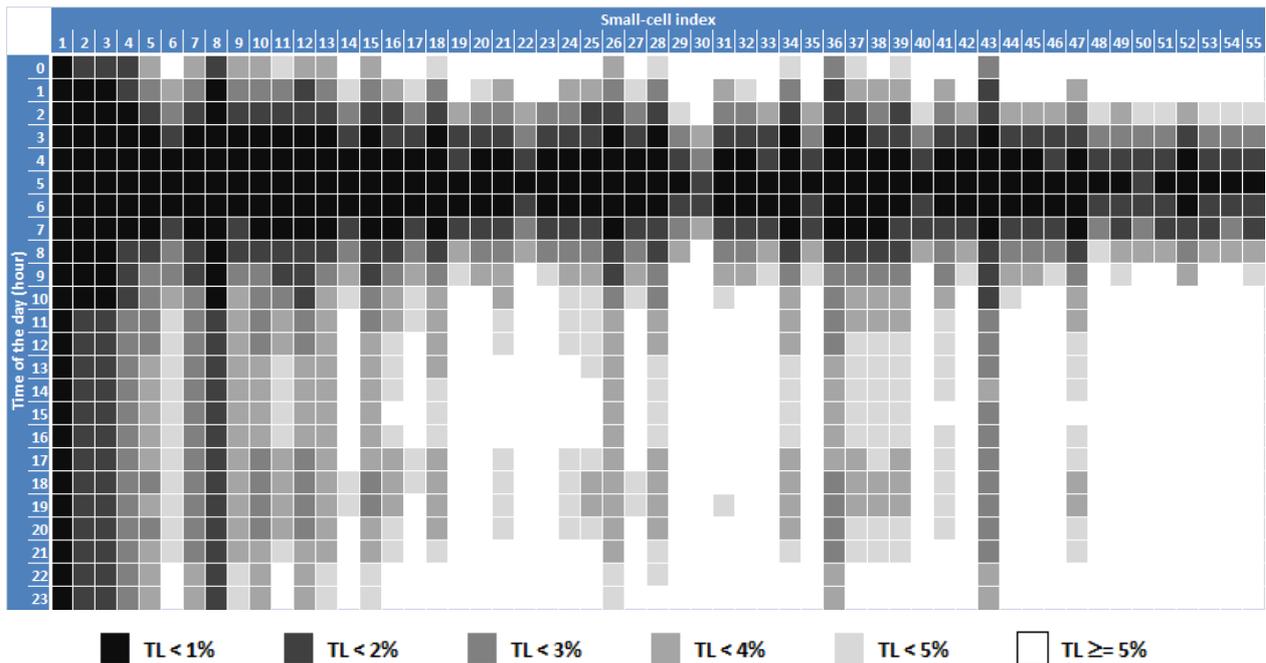
Figure 32 : Small-cell traffic loads as a function of time.
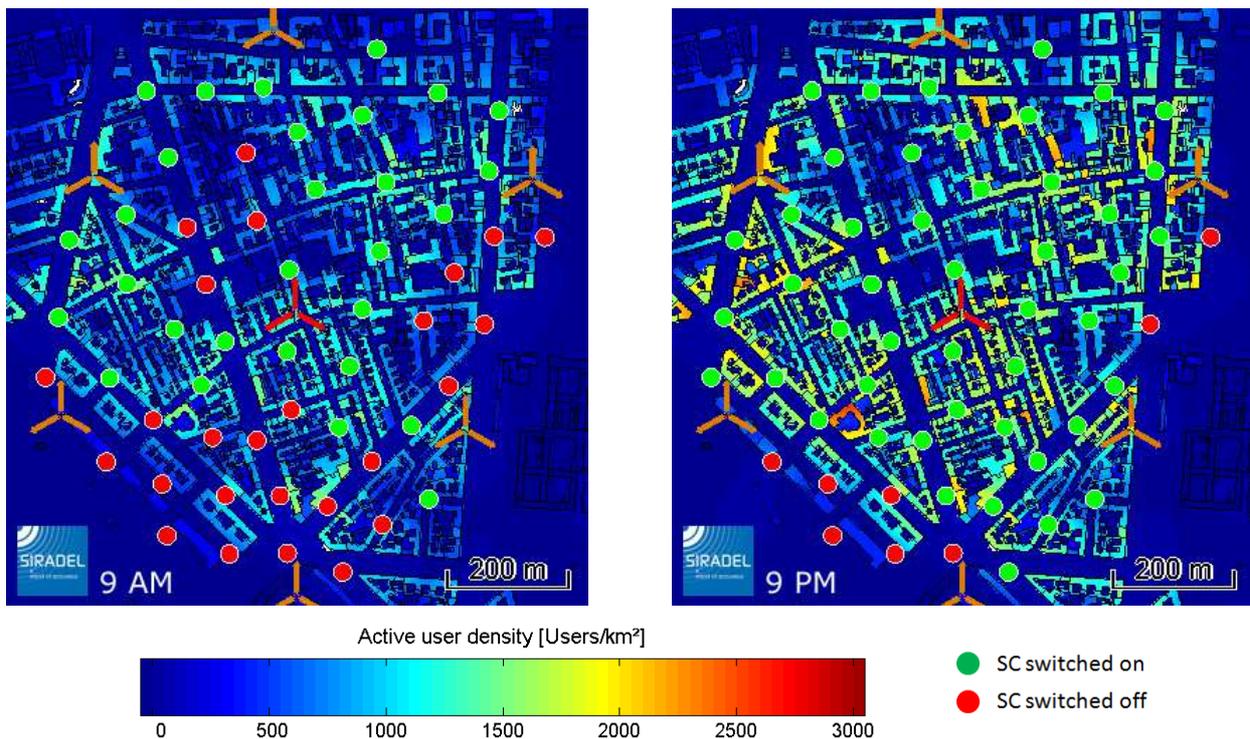
Figure 33 : Active user density maps and small-cells switched on or off at 9 AM (left) and 9 PM (right) for the 3% traffic load threshold.

Figure 34 (left) shows the average traffic load of the macro eNodeBs along the day. The offloading offered by the small-cell layer is obvious with a macro eNodeB average traffic load below 3% in the *Macro+SC (ON)* scenario. When switching off small-cells during night, the average macro eNodeB traffic load comes back to the one of the *Macro-only* scenario during

this period. Figure 34 (right) shows the average traffic load of the small-cells along the day. Line breaks indicate periods during which all small-cells are switched off. Remaining switched on small-cells with the highest traffic loads lead to an increase of the average small-cell traffic load with increasing threshold. We then observe that the average small-cell traffic load is higher and more constant, which is a good indicator that they are more efficiently used.
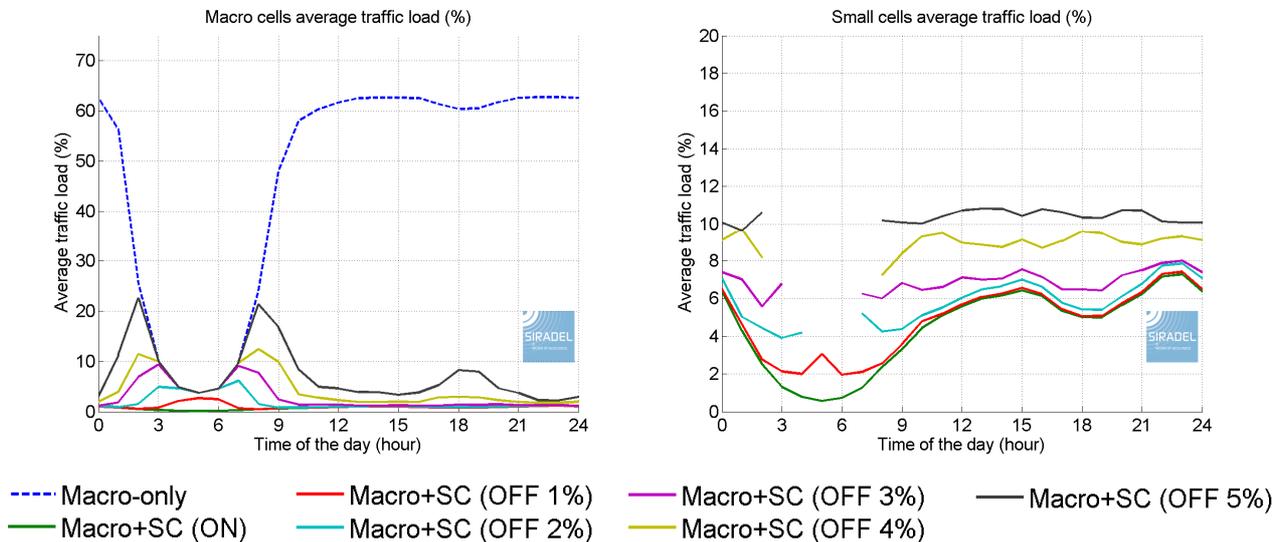
Figure 34 : Hourly average traffic loads of macro eNodeBs(left) and small-cells(right) in function of the small-cell traffic load threshold.

Figure 35(left) shows the service outage of the different scenarios along the day. The model is given at the end of the scenario section according to the model published in [28].Despite the deployment of a dense small-cell layer, the network cannot serve all users correctly at busy hours. The main reason is that the uniform small-cell deployment does not always match the traffic hotspots determined by the PoIs. This highlights the interest for mobile operators to design their small-cell layer with realistic traffic modeling. The indoor user ratio is also higher during these periods, leading to more cell-edge users on the small-cell layer. Their capacity to absorb cell-edge traffic is then limited by the ABS duty cycle. The service outage almost not changes when switching off small-cells with traffic load thresholds of 1, 2 or 3%, then quickly grows up with higher thresholds.

Figure 35(right) shows the network power consumption of the different scenarios along the day. Switched off small-cells are assumed to be in a stand-by mode waiting for a switch-on order, with a power consumption of 5 W. The interest in switching off small-cells is demonstrated by the power consumption at 5 AM which is 2.5 times higher with all small-cells switched on than the one of the *Macro-only* scenario. With the lowest small-cell traffic load threshold of 1%, it is already decreased to a factor of 1.2. With higher thresholds of 2 and 3%, we can observe that the daily power consumption comes closer to the one of the *Macro-only* scenario. With thresholds of 4 and 5%, the network even consumes less with a lower service outage than in the case of the *Macro-only* scenario. The power consumption follows then the same trend as the active user density presented in Figure 31 (left).

The daily energy consumption in kW/h is calculated with a simple integration and presented in Table 5. The addition of the small-cell layer to offload the macro eNodeBs leads to an increase of 39% in energy consumption if no switch-off mechanism is implemented. With the small-cell traffic load threshold of 3%, this energy consumption increase is reduced to 10% with similar QoS.
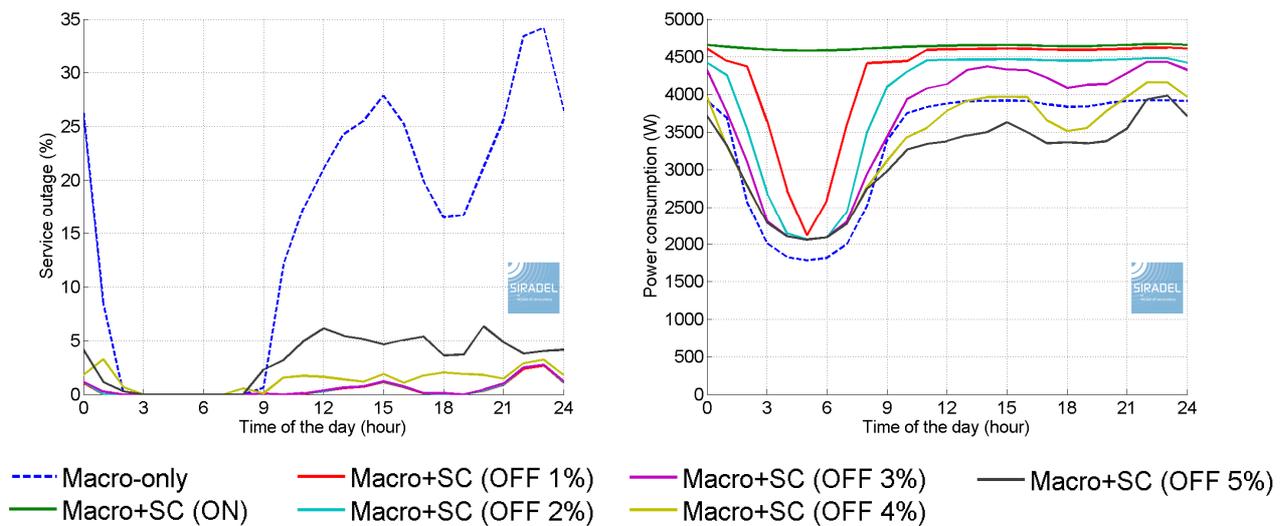
Figure 35 : Hourly service outage (left) and network power consumption (right) in function of the small-cell traffic load thresholds.

Table 5: Network energy consumption in function of the small-cell traffic load thresholds.

Scenario	Energy consumption
Macro-only	79.8 kW/h
Macro+SC (ON)	+39%
Macro+SC (OFF 1%)	+27%
Macro+SC (OFF 2%)	+17%
Macro+SC (OFF 3%)	+10%
Macro+SC (OFF 4%)	+0%
Macro+SC (OFF 5%)	-6%

2.5.5   Conclusion

This study demonstrates the interest in switching off small-cells outside busy hours to optimize the network energy consumption. Our scenario shows that a 29% gain can be achieved when switching off small-cells having less than 3% of traffic load, without compromising the user QoS. The service outage observed at peak hours is partly due to the ABS duty cycle limitation and suggests other Self-Organizing Network (SON) algorithms to achieve a finer macro to small-cell off loading depending on the time of the day. Another reason for the service outage is the unmatched between the uniform small-cell layer deployment and the non-uniform active user density. A perspective to this work is thus to deploy a non-uniform small-cell layer with regard to traffic hotspots. Modeling the switch-on mechanism is also of interest to assess the impact on service outage and energy consumption of switching on small-cells too late or too early.

3   ENERGY SAVING MECHANISMS IN THE DIGITAL AND ANALOG FRONT END

3.1   Efficiency and linearity enhancement in power amplifiers

Wireless communication standards impose stringent requirements on linearity performance of Power Amplifiers (PA). In addition, since the PA consumes most of the energy in telecommunication equipments (almost 60% for macro base stations), its power-efficiency becomes a primary concern. In fact, with the growth in voice and data communication usage, a high efficiency will contribute to the reduction of both the UE energy consumption resulting in a longer battery lifetime, and the base station energy consumption resulting in power savings and reduced environmental pollution. However, these two requirements, linearity and power efficiency, tend to be mutually exclusive in traditional PA design, so that any increase of the PA linearity by amplifying in linear region is usually achieved at the expense of the efficiency and conversely. This PA issue is a primary concern when considering multi-carrier modulations like Orthogonal Frequency Division Multiplexing (OFDM) which are prone to high Peak-to-Average Power Ratio (PAPR) resulting in low power efficiencies.

In such conditions, techniques including linearization and PAPR reduction have been proposed separately in the literature in order to improve the performance of the transmitters, including PA. The linearization ensures high linearity of the PA in order to avoid carrier inter-modulations and to respect the power mask. In wireless communication systems, the most promising and cost-effective linearization method is predistortion, which guarantees an acceptable linearity level of the PA over its intended power range. In case of the PAPR reduction, the dynamic range of the signal is reduced, which allows the PA to operate closer to the saturation power more efficiently. PAPR reduction includes some techniques like clipping and filtering, coding and tone reservation. Since PAPR reduction is associated to linearization and improves its effectiveness, the methodology of PA design focuses on a trade-off between linearity and power efficiency represented by linearization and PAPR reduction respectively.

Several works in the literature consider the association of PAPR reduction technique and digital predistortion as linearization method aiming to improve the performance of the PA in terms of linearity and efficiency. They can be classified in two approaches. The first approach consists of combining a PAPR reduction technique followed by predistortion; the second one takes into consideration the mutual effects of PAPR reduction and predistortion in order to propose an optimal combination. Obviously, the second approach will achieve better performance and a good trade-off between the PA linearity and the power efficiency.
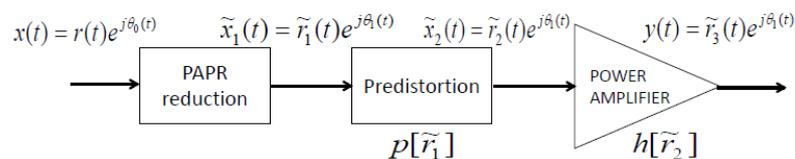
Figure 36 : PAPR and linearization processes.

We consider the simplified transmission chain presented in Figure 36. The OFDM signal $x(t)$ becomes $\tilde{x}_1(t)$ after PAPR reduction technique and $\tilde{x}_2(t)$ after predistortion. The amplified signal is $y(t)$. The PA is a memoryless Solid StatePower Amplifier (SSPA). Under the assumption that the power amplifier has no phase distortion, signals $\tilde{x}_1(t), \tilde{x}_2(t)$ and $y(t)$ have the same phase $\theta_1(t)$. Consequently AM/AM and AM/PM power amplifier characteristics are given by the following equations:

$$\mathbf{h(r)} = \frac{\mathbf{r}}{\left(1+\left(\frac{\mathbf{r}}{\mathbf{A}}\right)^{\mathbf{2b}}\right)^{\frac{1}{\mathbf{2b}}}} (8)$$

$$\mathbf{p(r)} = \frac{\mathbf{r}}{\left(1-\left(\frac{\mathbf{r}}{\mathbf{A}}\right)^{\mathbf{2a}}\right)^{\frac{1}{\mathbf{2a}}}}, \mathbf{r} \in [\mathbf{0}, \mathbf{A}[\quad (9)$$

$$\boldsymbol{\phi(r)} = \mathbf{0} (10)$$

Where a and b are the knee factors of the amplifier and predistorter respectively. We define a predistortion error $\epsilon(\tilde{r}_1)$ that compares the amplified signal $y(t)$ to the signal before predistortion $\tilde{x}_1(t)$. This error quantifies the performance of predistortion and is expressed by ($\tilde{r}_3$ is the modulus of the signal after the power amplifier):

$$\boldsymbol{\epsilon(\tilde{r}_1)} = |\boldsymbol{\tilde{r}_1} - \boldsymbol{\tilde{r}_3}| (11)$$

It can be shown that the upper bound of $\epsilon(\tilde{r}_1)$ is given by

$$\boldsymbol{\epsilon(\tilde{r}_1)} \leq \boldsymbol{\tilde{r}_1} \left| \mathbf{1} - \mathbf{2}^{\frac{\mathbf{b}-\mathbf{a}}{\mathbf{2ab}}} \right| (12)$$

In the rest of this section, this approximation will be considered to determine the second order moment of $\epsilon(\tilde{r}_1)$ depending on the distribution of the signal after the PAPR reduction technique.

We assume that the OFDM signal $x(t)$ is characterized by a complex stationary Gaussian process. Therefore, its amplitude $r(t)$ converges to a Rayleigh distribution. Our objective is to study the distribution of the signal after PAPR reduction in order to calculate the second order moment of the predistortion error firstly for probabilistic PAPR reduction methods and secondly for amplitude clipping methods.

PAPR reduction methods can be classified in two categories: the ones that modify the PAPR distribution and the ones that do not, i.e. where the Gaussian distribution remains Gaussian. In the case where the distribution of the signal $\tilde{x}_1(t)$ remains Gaussian (with Selective Mapping or Partial Transmit Sequences methods for instance) with mean power denoted $P_{\tilde{r}_1}$, the second order of the distortion error is given by

$$\tilde{m}_{2\_max}^{(prob)} = P_{\tilde{r}_1} \left(1 - 2^{\frac{b-a}{2ab}} \right)^2 \left(1 - \left(PAPR_{\tilde{r}_1} + 1 \right) e^{-PAPR_{\tilde{r}_1}} \right) (12b)$$

In the second case, we have considered the clipping method which is the most used due to its simplicity and its reduction gains. In this case, the second order of the predistortion error is given by

$$\hat{m}_{2\_max}^{(clip)} = \frac{P_{\tilde{r}_1}}{\gamma}\left(1 - 2^{\frac{b-a}{2ab}}\right)^2 \left(1 - e^{-\gamma PAPR_{\tilde{r}_1}}\right) \quad (12c)$$

The Error Vector Magnitude (EVM) of the amplified signal is defined as the ratio of the Root Mean Square (RMS) of the predistortion error to the root of the mean power of the signal after PAPR reduction. As a result, the EVM of $y(t)$ is given by eq. (13) for the probabilistic case and by eq. (14) for the clipping case. γ is the ratio of the signal power after and before PAPR reduction.

$$\mathbf{EVM_{max}^{(prob)}} = \left|\mathbf{1 - 2^{\frac{b-a}{2ab}}}\right|\sqrt{\mathbf{1 - \left(PAPR_{\tilde{r}_1} + 1\right)e^{-PAPR_{\tilde{r}_1}}}}(13)$$

$$EVM_{max}^{(clip)} = \left|\mathbf{1 - 2^{\frac{b-a}{2ab}}}\right|\sqrt{\left(\frac{1}{\gamma}\right)\left(1 - e^{-\gamma PAPR_{\tilde{r}_1}}\right)}(14)$$

Each simulation considers 5000 randomly generated OFDM symbols with 64 sub-carriers each 16-QAM modulated. The difference between theory and simulations comes from the aforementioned upper bound. The output PAPR has been fixed to 5 dB.
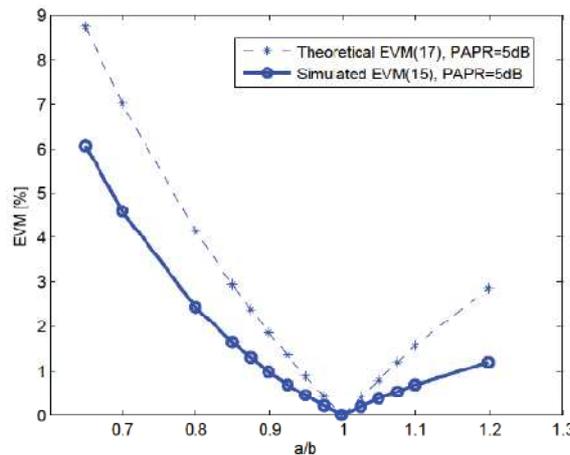

Figure 37 : EVM vs a/b ratio (for probabilistic PAPR reduction methods).

The main characteristic of the PA is that it dissipates power whatever the amplitude of the input signal. Figure 38 illustrates a basic PA power budget.
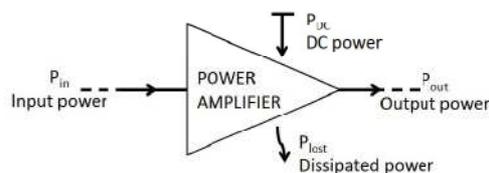

Figure 38 : Simplified PA power budget.

Referring to Figure 38, the PA efficiency can be defined as

$$\eta_{DC} = \frac{P_{out}}{P_{DC}} (15)$$

The maximum efficiency is achieved at the maximum amplitude of the linear output signal. This efficiency mainly depends on the input signal PAPR as shown in the next subsection but also on the PA class. η_{DC} cannot exceed 50% and 25% for class A and B power amplifiers respectively when operating in linear conditions.

As explained before, PA's linearity and its efficiency are mutually exclusive and cannot be achieved simultaneously. So the main challenge of designers is to achieve an optimal tradeoff by applying PAPR reduction and linearization techniques. Predistortion can be used to improve PA's linearity while PAPR reduction increases its efficiency. Nevertheless, linearization and PAPR reduction methods are studied and applied independently. After some derivations, it is shown that EVM is a function of the PA efficiency and the predistortion quality as well, as illustrated in

$$EVM_{max}^{(prob)} = \left| 1 - 2^{\frac{b-a}{2ab}} \right| \sqrt{1 - \left(\frac{\eta_{DC}}{G} \right)^{\frac{1}{g}} \left(1 - \frac{1}{g} \ln \left(\frac{\eta_{DC}}{G} \right) \right)} (16)$$

This dependency is illustrated in Figure 39 for different values of a/b ratios.
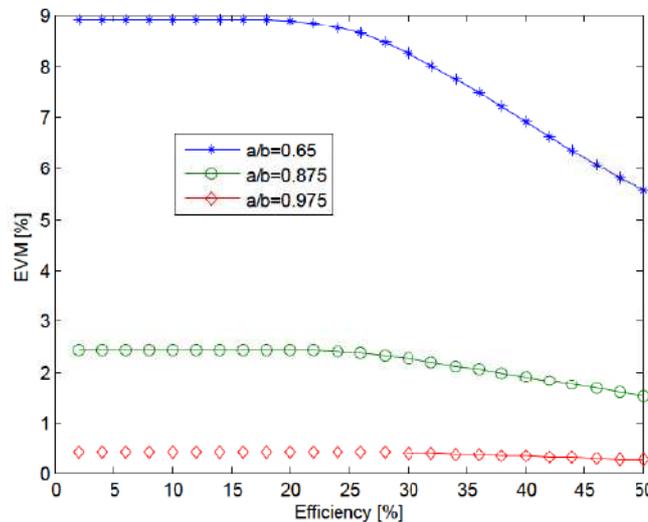
Figure 39 : EVM vs power efficiency.

Figure 40 shows the evolution of the EVM metric measured between the amplified signal and the PAPR reduced signal for different IBO values when predistorter is optimal ($a = 2$). To give an idea, we observe that the IBO = 5dB, corresponding to the targeted PAPR, could be a good trade-off between the linearity and the power efficiency. When IBO <5dB, the EVM increases from 0 to 17% for both Selective Mapping (SLM) and Amplitude Clipping. This means that the linearity is degraded while the power efficiency is improved. When IBO >5dB, the EVM is minimal, 0%, what means that the highest linearity is achieved while the efficiency is decreasing.
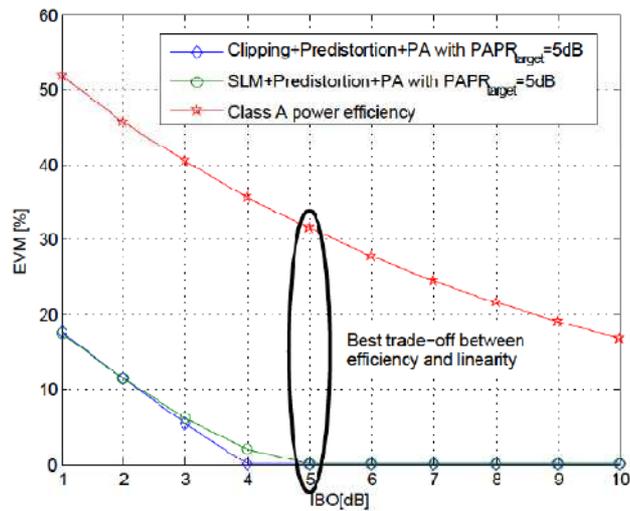
Figure 40 : Efficiency and linearity trade-off illustration.

This study, corresponding to scenario 2.9.9 in [5] has provided a theoretical analysis of the tradeoff between linearity measured by EVM metric and the efficiency in OFDM context. Analytical expressions of the EVM and the power efficiency for a memoryless SSPA have been formulated based on systems where PAPR reduction technique is followed by a predistortion before the nonlinear PA. The validity of the theoretical expressions has been verified through simulations for amplitude clipping and SLM techniques. Thus the performance of an OFDM system with nonlinear PA where a PAPR reduction technique is followed by a predistortion can be estimated theoretically without the need to perform extensive simulations. Some other configurations can be considered taking into account the memory effects of the PA or considering other linearity metrics like Adjacent Channel Power Ratio (ACPR). This will be the subject of our future work.

3.2   Joint optimization methods between PAPR reduction and linearization

Bussgang's theorem states that when a gaussian stationary process passes through a memoryless non-linear device, the cross-correlation function of input and output is proportional to the auto-correlation function of input[12]. Let f(.)be a non-linear function, x(t) and y(t) their input and output signals respectively. The theorem can be written as follows:

$$y(t) = f\big(x(t)\big) = \alpha x(t) + d(t) \qquad (17)$$

where

$$\alpha = \frac{R_{yx}(0)}{R_{xx}(0)}$$

Ryx and Rxx are the cross-correlation and the auto-correlation functions of the input signal and the output signal respectively. It is shown that the distortion term d(t) is uncorrelated with the input signal x(t).

Considering that OFDM signal can be modeled by a complex gaussian process[13]and using Bussgang's decomposition presented in equ.17, authors of [14]proved that all PAPR reduction techniques based on a nonlinear function (like clipping, companding, etc.) can be formulated as adding signal techniques and that the uncorrelated term d(t) is the useful signal for PAPR reduction. Generally speaking, the principle of adding signal techniques sets a peaks reducing signal c^{papr}(t) that, added to the input OFDM signal x(t), reduces its PAPR. Then, the signal y(t) whose PAPR is reduced is expressed as follows:

$$y(t) = f_{PAPR}\big(x(t)\big) = x(t) + c^{PAPR}(t) \qquad (18)$$

Where c^{papr}(t) = (α-1)x(t) + d(t) and f_{PAPR}(.) the corresponding non-linear PAPR reduction function.

By using null or dedicated sub-carriers of the OFDM frame to generate the peaks reducing signal c^{papr}(t), authors of [14] also proved that PAPR can be reduced without Bit Error Rate (BER) degradation. Let $\Re_u$ be the subset of useful data subcarriers and $\Re_r$ the one of reserved sub-carriers of a given standard. The set of all sub-carriers is given by $\Re_u \cup \Re_r$ and Ru ∩Rr = ∅. Then, the peaks reducing signal can be generated in order to keep BER unchanged.

3.2.1   Predistortion by adding signal

Most of the time predistortion consists in applying to an input signal, a non-linear function f_{dpd}(.) which is the inverse of the PA characteristic. As a result, the concatenation of this function and the PA is ideally equivalent to a linear function. In addition, if the memory effects of the PA are negligible, predistortion is therefore based on a nonlinear and memoryless function. As a consequence, it can be formulated as an adding signal technique thanks to Bussgang's theorem considering that OFDM signal is a complex gaussian process.Let y(t) be the input signal and c^{dpd}(t) the additional signal for predistortion. This predistortion signal c^{dpd}(t) is related to the signal for PA's non-linearities compensation. The output of the predistortion $\breve{y}$(t) is then expressed as follows:

$$\breve{y}(t) = f_{dpd}(y(t)) = y(t) + c^{dpd}(t) \qquad (19)$$

From equ. 17 and 19, the predistortion signal is expressed as cdpd(t) = (α-1)x(t) + d(t).We notice that the predistortion signal depends on the correlation factor αand the uncorrelated signal d(t) resulting from Bussgang's theorem applied to the predistortion process. More details about calculation of αand the predistortion signal c^{dpd}(t) can be found in [15].

3.2.2   Combined approach by adding signal

Figure 41 first illustrates a conventional system where PAPR reduction and predistortion are implemented separately.
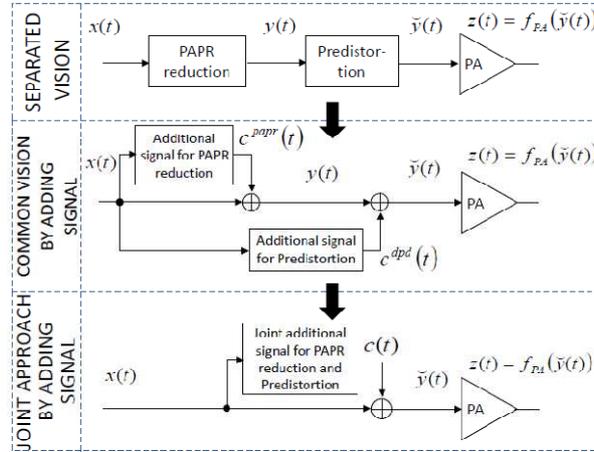
Figure 41 : Separate, common and joint approaches of PAPR reduction and predistortion.

Then, from OFDM input signal x(t), additional signals $c^{papr}(t)$ and $c^{dpd}(t)$ are generated for PAPR reduction and predistortion respectively, leading to the following expression:

$$\breve{y}(t) = f_{dpd}(y(t)) = f_{dpd}(x(t) + c^{papr}(t)) = x(t) + c^{papr}(t) + c^{dpd}(t) \quad (20)$$

where $\breve{y}(t)$ is the output signal after PAPR reduction and predistortion by adding signals. Finally, from equ.20, a new joint additional signal is defined:

$$c(t) = c^{papr}(t) + c^{dpd}(t) \quad (21)$$

for PAPR reduction and predistortion simultaneously. From Figure 41 and equ.21, it can be concluded that a single additional signal can be generated to perform at the same time PAPR reduction and predistortion[16].In the next section, we propose two algorithms to process the joint additional signal expressed in equ.21.

3.2.3   Proposed algorithms

After having proved thanks to the Bussgang's theorem that there exists a joint additional signal for PAPR reduction and predistortion, we move forward in this section to show that this signal can be generated by just solving an optimization problem. Two solutions are proposed.PA generally generates non-linearities in both amplitude and phase. Let's take $\breve{y}(t)$ and z(t) as the equivalent PA input and output signals respectively. z(t) is expressed as :

$$z(t) = f_{PA}(\breve{y}(t)) = H_{PA}(|\breve{y}(t)|e^{j(\Phi_{\breve{y}}(t) + \psi_{PA}(|\breve{y}(t)|)} \quad (22)$$

where $|\breve{y}(t)|$ and $\Phi\breve{y}$ (t) denote the amplitude and the phase of the input signal $\breve{y}(t)$ respectively. Functions $H_{PA}(.)$ and $\psi_{PA}(.)$ represent amplitude to amplitude (AM/AM) and amplitude-to-phase (AM/PM) non-linear transfer functions of the PA respectively. $f_{PA}(.)$ is the PA's transfer characteristic in its complex form. The objective of the joint approach by adding signal is to generate a joint additional signal c(t) that compensates PA non-linearities and also reduces the input signal's PAPR. As c(t) is a combination of $c^{papr}(t)$ and $c^{dpd}(t)$, the PA output is ideally expressed as follows:

$$z(t) = f_{PA}\big(\breve{y}(t)\big) = f_{PA}\left(x(t) + c^{papr}(t) + c^{dpd}(t)\right) = G_0(x(t) + c^{papr}(t)) \quad (23)$$

where G_0 is the gain of the power amplifier. We can notice from equ.23 that the PA output signal mainly depends on $c^{papr}(t)$. It is impossible to linearise the PA above its saturation power, so perfect linearity observed in equ.24 is achieved only for large IBO to avoid PA saturation. Thus, the joint additional signal c(t) is computed in order to minimize a mean square criteria defined by:

$$J(c) = E\{|f_{PA}(x(t) + c(t)) - G_0(x(t) + c^{papr}(t))|^2\} \quad (24)$$

Considering that the peaks reducing signal $c^{papr}(t)$ is calculated first, the joint additional signal c(t) is found out by solving the following problem:

$$\min_c J(c) \quad (25)$$

Mathematically speaking, many methods are possible for solving this optimization problem. Its convergence is obvious as the PA's characteristic $f_{PA}(.)$ is a convex function and since J is a convex objective. In the following, we propose two solutions. The first one is an algorithm based on a polynomial model of PA. The second one considers equ.25 in the frequency domain and then minimises J(c) using on an iterative error compensation algorithm.

3.2.3.1 Algorithm based on polynomials' roots finding

Without lost of generality, we assume in the following that PA is modelled as a memoryless polynomial with odd degree terms [17].Polynomial model is a simple way in terms of computation to model non-linear systems like PA compared to Volterra kernels. PA transfer characteristic is then given by equ.26 :

$$f_{PA}(x(t) + c(t)) = (x(t) + c(t)) \sum_{k=0}^{m-1} d_{2k+1}|x(t) + c(t)|^{2k} \quad (26)$$

where $\{d_1, d_3, ..., d_{2m-1}\}$ are the polynomial coefficients. These coefficients are obtained thanks to an identification step. The parameter m is the degree of the polynomial. Let us

introduce r_n as the amplitude of the digital signal to be amplified and $d0,n$ as the amplitude of the desired PA output signal in the digital domain :

$$r_n = |x_n + c_n| \qquad (27)$$

$$d_{0,n} = G_0\left|x_n + c_n^{papr}\right| \qquad (28)$$

where x_n, c_n and c_n^{papr} represent receptively x(t), c(t) and c^{papr}(t) in the digital domain. Considering the PA polynomial characteristic (equ.28), the cost function can be rewritten as follows:

$$\mathbf{J(c)} = \mathbf{E}\left\{\left(\textstyle\sum_{k=0}^{m-1} d_{2k+1}(r_n^{2k+1}) - d_{0,n}\right)^2\right\} \qquad (29)$$

Under these conditions, finding the optimal additional signal c_n by solving equ.25 is equivalent to find the signal r_n by minimizing equ.29. Then, the additional signal c_n is deduced from equ.27 as follows:

$$\mathbf{c_n = r_n e^{j\emptyset xn} - x_n} \qquad (30)$$

Where Φ_{xn} represents the phase of the desired PA output signal $x_n + c_n^{papr}$. Finding the solution which minimizes equ.29 can be turned into finding the roots of the following equation:

$$\mathbf{d_{2m-1} r_n^{2m-1} + d_{2m-3} r_n^{2m-3} + \cdots + d_1 r_n - d_{0,n}} \qquad (31)$$

The problem is now a polynomials' roots finding problem where each root represents $r_n = |x_n + c_n|$. We can notice that equ.31 depends on the discrete time variable n so each equation allows to calculate a sample of r_n and of x_n. Considering an OFDM signal that is constituted by NL samples with N the total number of sub-carriers and L the oversampling factor, the additional signal c = {c_0, c_1, ..., c_{NL-1}}is a solution of the system

$$\boldsymbol{R \times D = D_0} \qquad (32)$$

with

$$R = \begin{bmatrix} r_0 & \cdots & r_0^{2m-1} \\ \vdots & \ddots & \vdots \\ r_{NL-1} & \cdots & r_{NL-1}^{2m-1} \end{bmatrix}$$

$$D = \begin{bmatrix} d_1 \\ \dots \\ d_{2m-1} \end{bmatrix} \text{ and } D_0 = \begin{bmatrix} d_{0,0} \\ \dots \\ d_{0,NL-1} \end{bmatrix}$$

Notice that the nth line of the system of polynomials (equ.32) represents the equation equ.31 so solving equ.32 in order to generate the joint additional signal for one OFDM symbol is equivalent to solving equ.31 for n = 0, 1, ..., NL-1. Solving equ.32 is done into two steps:

1. generate the peaks reducing signal $c^{papr} = \{c_0^{papr}; c_1^{papr}, ..., c_{NL-1}^{papr}\}$ using a PAPR reduction technique [18]. Tone Reservation methods can be used to avoid BER degradation [19].

2. find c by solving equations equ.30 using a polynomial's roots finding method under the assumption that polynomial coefficients $\{d_1, d_3, ..., d_{2m-1}\}$ are known thanks to a preliminary step of PA characteristics estimation.

Determining roots of a polynomial is one of the oldest problems in mathematics for which many methods have been proposed in the past. The odd degree of the polynomial equ.30 ensures the existence of at least one real root. The eigenvalues method is used in this paper. It allows finding all roots with high precision. Eigenvalues method is an algorithm that computes the roots of a polynomial by finding the eigenvalues of the corresponding companion matrix. Indeed, in classical linear algebra, the eigenvalues of a matrix are defined as the roots of the characteristic polynomial. For this, we rewrite the polynomial equ.30 into a monic form $f_{dpd}(.)$ expressed as follows:

$$f_{dpd}(r_n) = r_n^{2m-1} + p_{2m-3}r_n^{2m-3} + \cdots + p_1 r_n - p_{0,n} \qquad (33)$$

with $p_{0,n} = d_{0,n}/d_{2m-1}$ and $p_{2k-1} = d_{2k-1}/d_{2m-1}$ for k=1,3, ..., m-1 and d_{2m-1} is different from zero. The companion matrix of this polynomial is derived as follows:

$$M_n = \begin{pmatrix} -p_{2m-2} & -p_{2m-3} & \cdots & -p_1 & p_{0,n} \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} \qquad (34)$$

Note that this matrix is square and has characteristic polynomial $P_{Mn}(r_n) = det\{r_n Id-M_n\} = f_{dpd}(r_n)$.

The last step of eigenvalues method is to calculate the eigenvalues of this matrix that correspond to the polynomial roots. Normally for one polynomial, several roots can be found. It is important to choose among them the smallest positive root in order to avoid phase distortions and to limit the mean power of the joint additional signal. After determining r_n, the joint additional signal c_n is calculated using equ.25. Considering an OFDM symbol and the system of polynomials described in equ.31, we can notice that only one element is changing

from companion matrix of one polynomial to another one. This element ($p_{0,n}$ in matrix M_n(equ.34)) is related to the desired PA output signal.

Consequently, it is possible to apply the eigenvalues method once for one polynomial and then deduce the roots of the other polynomials.
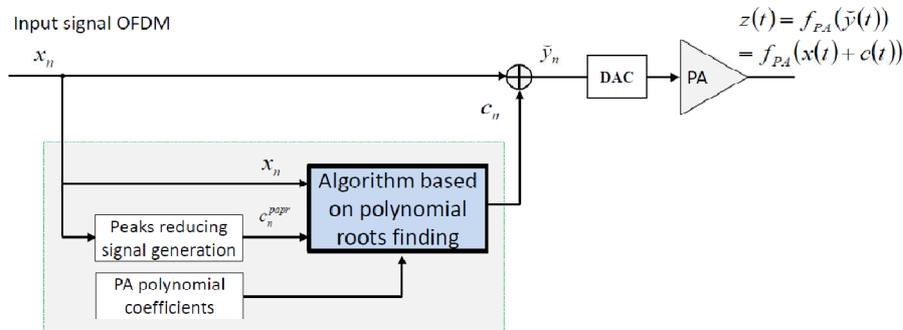

Figure 42 : Block diagram of the algorithm based on polynomials's roots finding for a joint approach of PAPR reduction predistortion by adding signal.

Figure 42 illustrates the blocks diagram of the proposed polynomials based algorithm for the joint approach of PAPR reduction and predistortion by adding signal. Contrary to the conventional separated vision, the proposed approach has the advantage of merging PAPR reduction and predistortion into a single block and tuning them simultaneously. Moreover, the computational complexity of the proposed algorithm can be estimated to $O(NLm^3)$ (N is the total number of useful carriers) due to the eigenvalues computation of m x m matrix; but this complexity can be reduced to $O(m^3)$ by considering the possibility to find out the roots of all polynomials in equ.31 from the roots computation of a single polynomial, for example equ.31 with n = 0. However, it is important to notice that a preliminary step of PA's characteristics estimation is mandatory for the proposed algorithm. The algorithm also requires the generation of a peaks reducing signal. This increases its complexity and its performance will depend on the PAPR reduction technique used. In the following, we propose another algorithm for the generation of the joint additional signal. This algorithm does not need a preliminary PA's characteristics estimation or a first generation of the peaks reducing signal; it is based on an error compensation algorithm.

3.2.3.2   Error compensation algorithm

As proved in equ.21, the joint additional signal c is a combination of a PAPR reduction signal c^{papr} and a predistortion signal c^{dpd}. By translating this in the frequency domain, we combine in the following, the PAPR reduction signal on the reserved or null sub-carriers and the predistortion signal on the useful sub-carriers as illustrated in Figure 43. The first reason of doing this is to avoid the BER degradation due to the PAPR reduction signal. Moreover, predistortion is more accurate as it is focused on data (useful) sub-carriers.
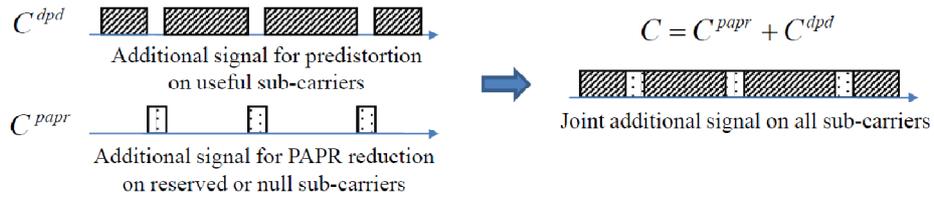
Figure 43 : Frequency vision of the targeted joint additional signal.

Let us take X and C as the initial OFDM's signal and the joint additional signal in the frequency domain respectively. The amplification of $\tilde{y}(t)=x(t) + c(t)$ will introduce some distortions on useful sub-carriers as well as on reserved sub-carriers. Let us consider $\hat{X}$ as the distorted data after amplification on the useful sub-carriers and $\hat{C}^{papr}$ the distorted signal after amplification on the reserved sub-carriers as illustrated in Figure 44. We define a frequency domain error vector Ek between the amplified signal on the useful sub-carriers $\hat{X}_k$ and the initial OFDM's data symbol X_k as follows:

$$E_k = X_k - \widehat{X_k} \text{ , } k \in \Re_u \qquad (35)$$
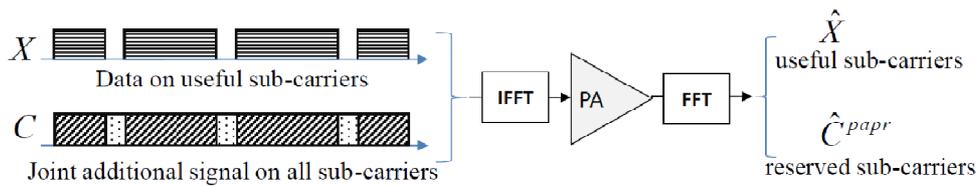
Figure 44 : Frequency vision of the amplification's process in the context of joint approach by adding signal.

where $\Re u$ represents the subset of the useful sub-carriers. The objective of the error compensation algorithm is to minimize the error E_k on the useful sub-carriers and to keep a peaks reducing signal on the reserved subcarriers. Indeed, depending on the power fluctuations of the input signal, the PA saturates and clips the signal. Distortions introduced are actually a peaks reducing signal (i.e clipping signal) that spreads on all sub-carriers including useful and reserved sub-carriers. The part on useful sub-carriers is removed by the error compensation algorithm that achieves PA predistortion. The part of the distortions introduced by the PA on the reserved subcarriers (i.e $\hat{C}_k^{papr}$ with $k \in \Re_u$) remains and used for PAPR reduction.
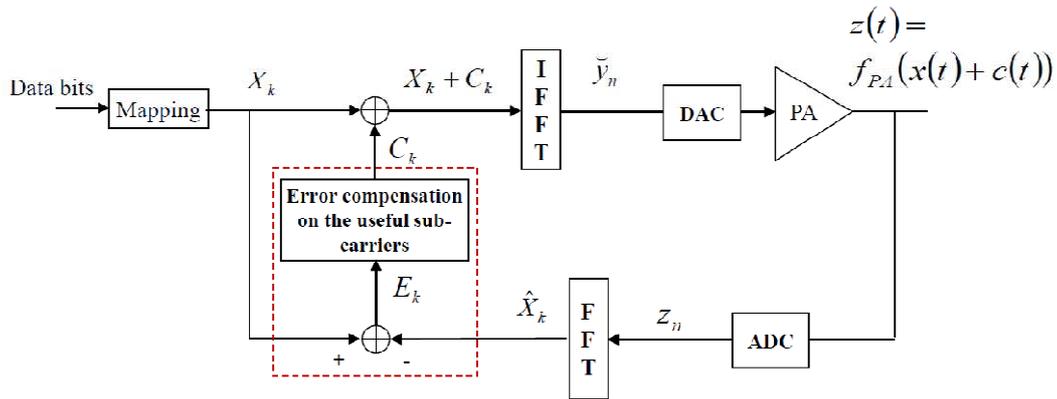
Figure 45 : Block diagram of the error compensation algorithm for the generation of the joint additional signal.

Figure 45 presents the block diagram of the error compensation algorithm for the generation of the joint additional signal. The PA characteristics are directly learned thanks to the feedback path in order to compensate the error (equ.34) on the useful sub-carriers. $\{C_k\}$, k =0, 1, ..., NL -1 are initialized to zero and then added to the initial OFDM's symbols $\{X_k\}$, k =0, 1, ..., NL -1. After amplification, the amplified signal $z(t) = f_{PA}(x(t) + c(t))$ is transformed to the frequency domain, the error vector (19) is computed and the joint additional signal is updated. The error compensation algorithm is based on gradient method and is detailed in Alg. 1. Steps 1 to 4 represent the initialization steps and are computed once. The joint additional signal is iteratively computed at the steps 9 and 10 depending on the error (step 8). The algorithm stops when this error falls below a given threshold (i.e less than ϵ) or after a maximum number of iterations. An appropriate rate μ is necessary to ensure the convergence of the algorithm. We can notice that each iteration of Alg.1 uses IFFT/FFT, so its asymptotic complexity can be estimated to O(R x NL log$_2$NL) with R the number of iterations. Alg.1 also uses multiple times the physical amplifier but a model based method can be envisaged in real systems where iterations will be performed over a model of the PA before transmitting the data and the additional signal over the physical PA in a last step.

Algorithm 1 Error compensation algorithm for the generation of the joint additional signal

Require:
1: X_k input OFDM symbol,
2: $\mu \Leftarrow$ convergence rate,
3: $i \Leftarrow 0$, initial index of iterations,
4: $C_k^{(i)} \Leftarrow 0, k \in \mathfrak{R}_u \cup \mathfrak{R}_r$, initial additional signal,
5: **repeat**
6: $\quad \breve{y}_n^{(i)} \Leftarrow \mathbf{IFFT}\left\{X_k + C_k^{(i)}\right\}$, predistorted and PAPR reduced signal,
7: $\quad \hat{X}_{k_u}^{(i)}, \hat{C}_{k_r}^{papr} \Leftarrow \mathbf{FFT}\left\{f_{PA}\left(\breve{y}_n^{(i)}\right)\right\}, k_u \in \mathfrak{R}_u, k_r \in \mathfrak{R}_r$, the feedback,
8: $\quad E_{k_u}^{(i)} \Leftarrow X_{k_u} - \hat{X}_{k_u}^{(i)}$, the error on the useful sub-carriers,
9: $\quad C_{k_u}^{(i+1)} \Leftarrow C_{k_u}^{(i)} + \mu E_{k_u}^{(i)}, k_u \in \mathfrak{R}_u$, the joint additional signal on useful sub-carriers,
10: $\quad C_{k_r}^{(i+1)} \Leftarrow C_{k_r}^{(i)} + \hat{C}_{k_r}^{papr}, k_r \in \mathfrak{R}_r$, the joint additional signal on reserved sub-carriers,
11: $\quad i \Leftarrow i + 1$, increment index of iterations,
12: **until** $i \geq$ maximum number of iterations or $\left|E_k^{(i)}\right| \leq \varepsilon$.

3.2.4 Proposed algorithms

Simulation is based on IEEE802.11a/g standards with an OFDM system of 64 sub-carriers including 52 useful subcarriers for the data and 12 null sub-carriers that can be used for PAPR reduction. The data is16-QAM modulated and the over-sampling factor is fixed to 4. We consider a Solid State Power Amplifier (SSPA) characterized by a Rapp model. The parameter m representing the degree of the PA's characteristics polynomial is fixed to 7. We first implement the joint approach by adding signal based on polynomials' roots finding with a Clipping & Filtering (C&F) as PAPR reduction technique. Then, we simulate the error compensation algorithm before comparing the two. Simulation results include PAPR, EVM, ACPR and Power Spectral Density (PSD) criteria measured after PA.

3.2.4.1 Algorithm based on polynomials' roots finding with clipping and filtering as PAPR reduction technique

We simulated here the first proposed algorithm. It generates a joint additional signal for PAPR reduction and predistortion by using a polynomials' roots finding method as explained before. The preliminary step of PA characteristics estimation is achieved by modeling the PA as a polynomial. For the PAPR reduction technique, we use clipping and filtering (C&F) with the clipping ratio (CR) fixed to 5.5dB. The performance of this polynomial based algorithm is compared first to that of a conventional system for which PAPR reduction technique (i.e C&F) is combined with an ideal predistortion (i.e the ideal inverse of the PA's characteristic) and secondly to the performance of a system without predistortion where only C&F is used before the PA.

Figure 46 : CCDF performance of the algorithm based on polynomials' roots finding method with clipping and filtering as PAPR reduction technique at IBO value equal to 8 dB.

Figure 46 shows the Complementary Cumulative Distribution Function (CCDF) of PAPR over 106 randomly generated OFDM signals. PA operates with IBO = 8dB. We notice that only with a PAPR reduction technique, the PA saturates and its non-linearities act as a PAPR reduction. This results in high PAPR reduction gain (ΔPAPR = 3dB at 10^{-2}probability) and consequently high efficiency but also strong non-linearities. Thanks to the joint approach by adding signal, the PAPR of the signal is reduced and the PA non-linearities are compensated. This reduces the PAPR reduction gain but also the PA nonlinearities (ΔPAPR = 2.2dB at 10^{-2}probability). We notice that the CCDF of the joint approach by adding signal is identical to that of the conventional system, which means that the joint additional signal generated by the proposed algorithm is efficient to compensate PA nonlinearities and to reduce the PAPR.

Figure 47 : EVM performance versus IBO of the algorithm based on polynomials' roots finding with clipping and filtering as PAPR reduction technique.

Figure 47 shows the linearity performance evaluated at the PA output using EVM criteria. EVM is measured in the frequency domain by comparing the amplified signal to the initial OFDM signal. It is simulated over 105 OFDM symbols and plotted versus the PA Input Back-Off (IBO). The EVM performance decreases proportionally to the IBO until its minimum value (3.6%) that corresponds to BER degradation due to the clipping technique. We notice that thanks to the joint approach by adding signal, the linearity is improved compared to the case without joint approach (with only PAPR reduction). EVM values when using joint approach are close to that of the conventional case but a slight difference (around 1%) is noticeable. This is due to the accuracy of the method used for polynomials'roots finding and the parameter m chosen.

Figure 48 represents the PSD performance of the joint approach by adding signal for an IBO = 8dB. The simulation is done over 105 OFDM symbols. It shows the capacity of the proposed algorithm to mitigate the out-of-band regrowths. When only PAPR reduction technique is used, the PA introduces non-linearities depending on PAPR reduction gain so some spectral regrowths are noticeable. Thanks to the joint approach, the PA non-linearities are compensated and the PSD performance meets that of the conventional system.
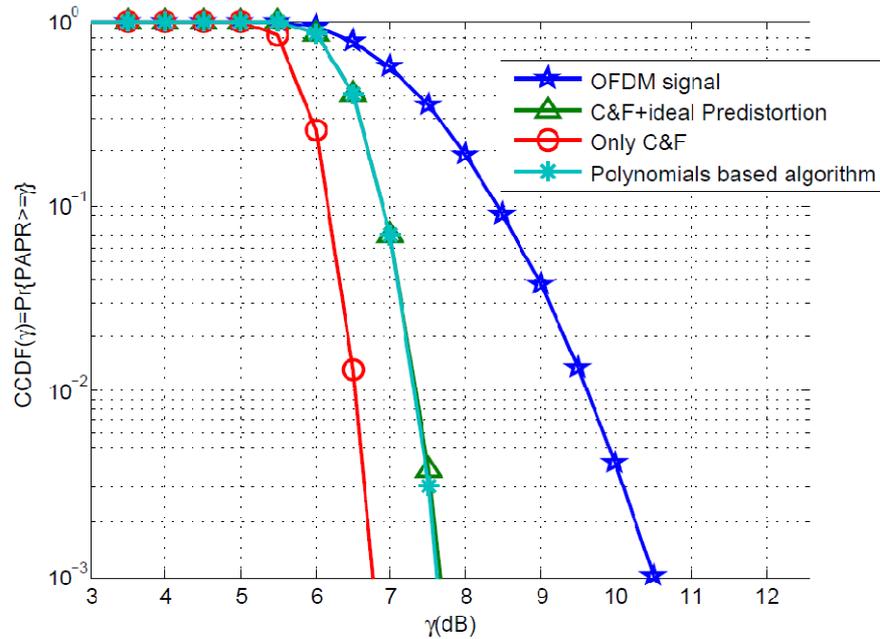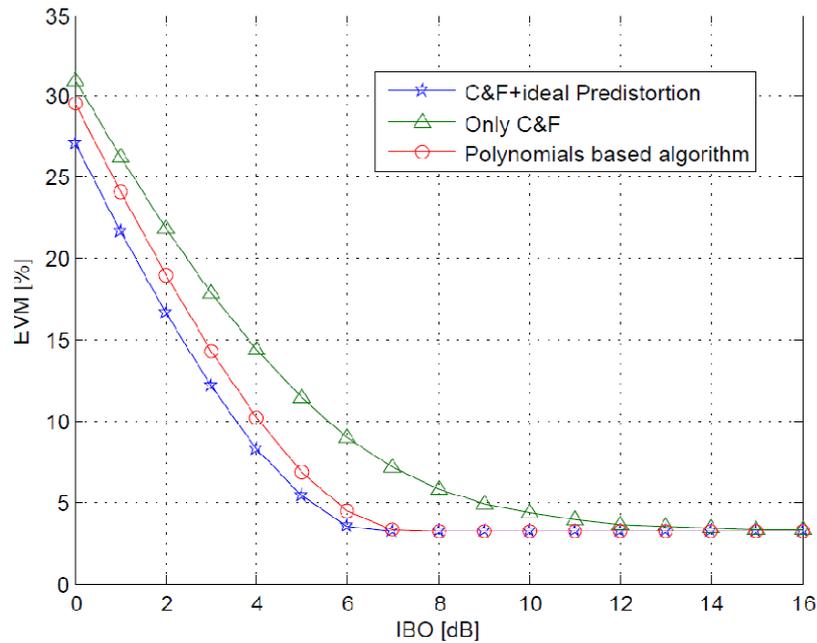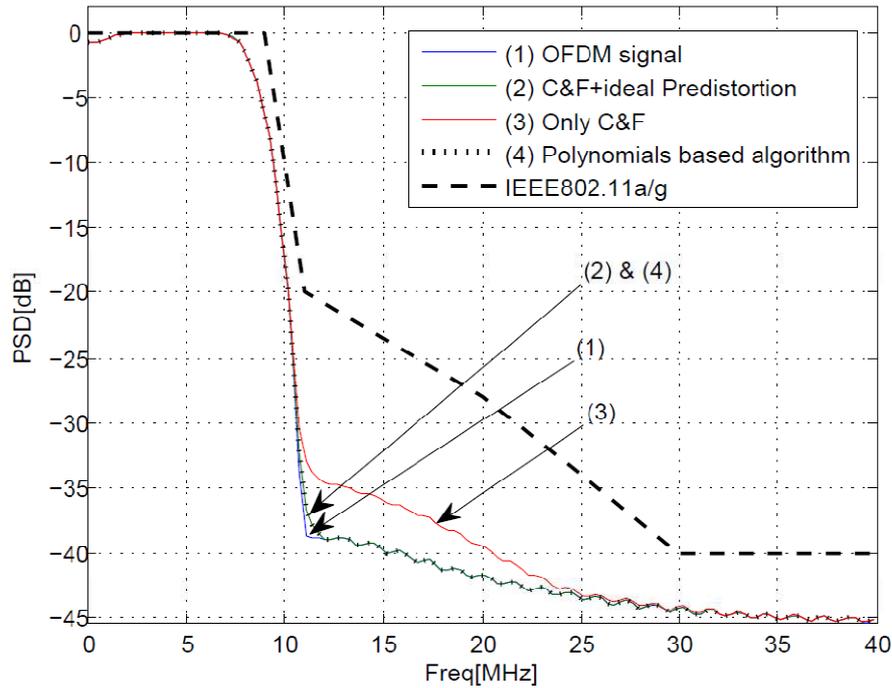
Figure 48 : PSD performance of the algorithm based on polynomials' roots finding with clipping and filtering as PAPR reduction technique (IBO = 8dB).

From the discussions above, we can conclude that the proposed algorithm based on polynomials' roots finding method is efficient to generate a joint additional signal. It has the same performance compared to a conventional system combining PAPR reduction and ideal predistortion. Indeed, they are mathematically equivalent. In the case of EVM, we noticed some difference due to the accuracy of the PA characteristics estimation and the polynomials' roots finding method. This means that the algorithm can be more accurate if the polynomial degree m is increased or if we choose another polynomials'roots finding method.

3.2.4.2   Algorithm based on error compensation on the useful sub-carriers

We simulate here the second proposed algorithm based on error compensation for a joint approach by adding signal. This algorithm is detailed in Algorithm 1. The joint additional signal for PAPR reduction and predistortion is generated iteratively and its performance increases proportionally to the number of iterations. We compare the performance of this algorithm to that of a combination of a clipping technique followed by an ideal predistortion. Indeed, both are theoretically equivalent when the clipping threshold is equal to the PA saturation (i.e CR=IBO). We first measure the evolution of the mean PAPR depending on the number iterations for two different IBO values. Then, we evaluate the EVM and the PSD at the PA output.
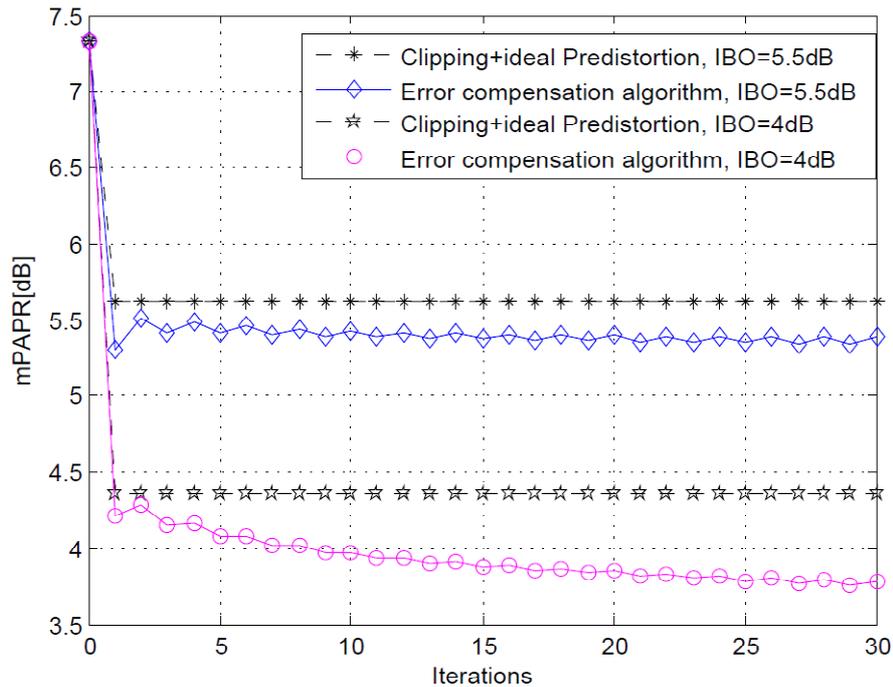
Figure 49 : PAPR reduction performance of the error compensation algorithm depending on the number of iterations.

Figure 49 plots the mean PAPR (mPAPR) of the amplified signal depending on the number of iterations of the error compensation algorithm for two IBO values, 4dB and 5.5dB. Simulation is done over 105 OFDM symbols. We notice that the PAPR reduction of the proposed algorithm depends on the number of iterations and the IBO. Already after one iteration, the PAPR reduction gain is higher than that of the ideal system where the OFDM signal is clipped to the saturation and linearly amplified (ideal predistortion). So, the proposed algorithm has fast convergence. For example with an IBO = 5.5dB and 10 iterations, the mean PAPR achieved is 5.4dB compared to 5.6dB in the ideal system. Compared to the 7.4dB mean PAPR of the input OFDM signal, this gives 2dB PAPR reduction gain for the proposed algorithm and 1.8dB for the ideal system. With an IBO = 4dB and 10 iterations, the PAPR reduction gain is 3.3dB and 3dB for the proposed algorithm and the ideal system respectively. When the number of iterations increases, for example to 30 iterations, the PAPR reduction gain increases accordingly until it reaches a maximum value that depends on the IBO. For an IBO = 5.5dB, this maximum value is equal to 2.1dB and for an IBO = 4dB, it is 3.6dB. Consequently, depending on the targeted mean PAPR, the number of iterations must be fixed according to the IBO. Figure 50 shows the EVM performance of the error compensation algorithm depending on the IBO and the number of iterations. EVM is measured after PA over 105 OFDM symbols. Obviously, it decreases proportionally to the IBO until it reaches the minimum value corresponding to linear system. In addition, we notice that it also depends on the number of iterations. After two iterations, the proposed algorithm achieves better EVM performance compared to the ideal system where the signal is clipped at the PA saturation. When the number of iterations increases, the EVM performance of the proposed algorithm converges to an asymptotic value that is simulated here with 100 iterations. This can be explained by the high PAPR reduction gain without BER degradation and the more accurate predistortion (due the PA estimation thanks to the feedback path) of the proposed algorithm compared to the ideal system which introduces in-band clipping noise.
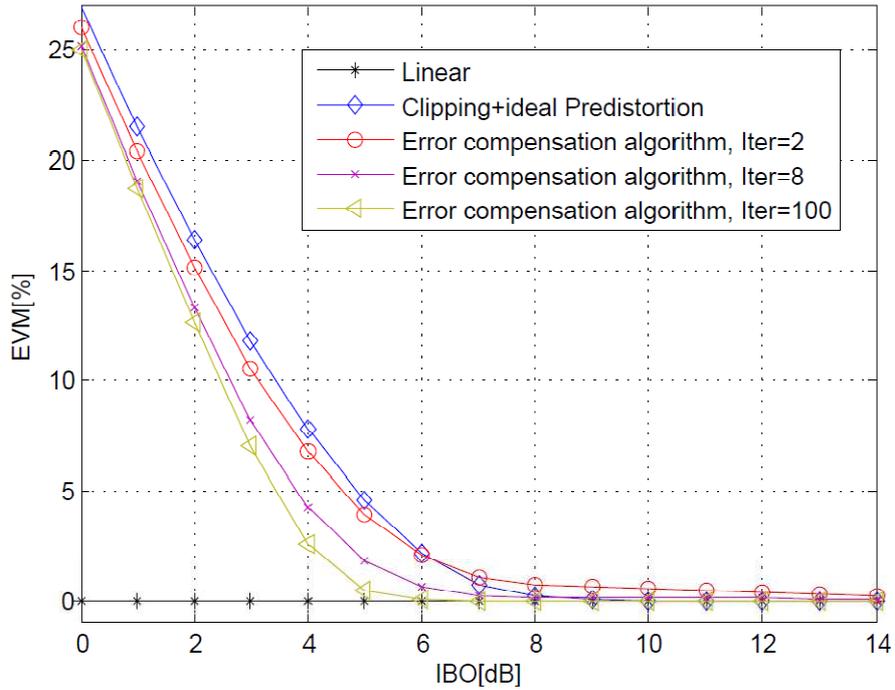
Figure 50 : EVM performance of the error compensation algorithm depending on the PA IBO.
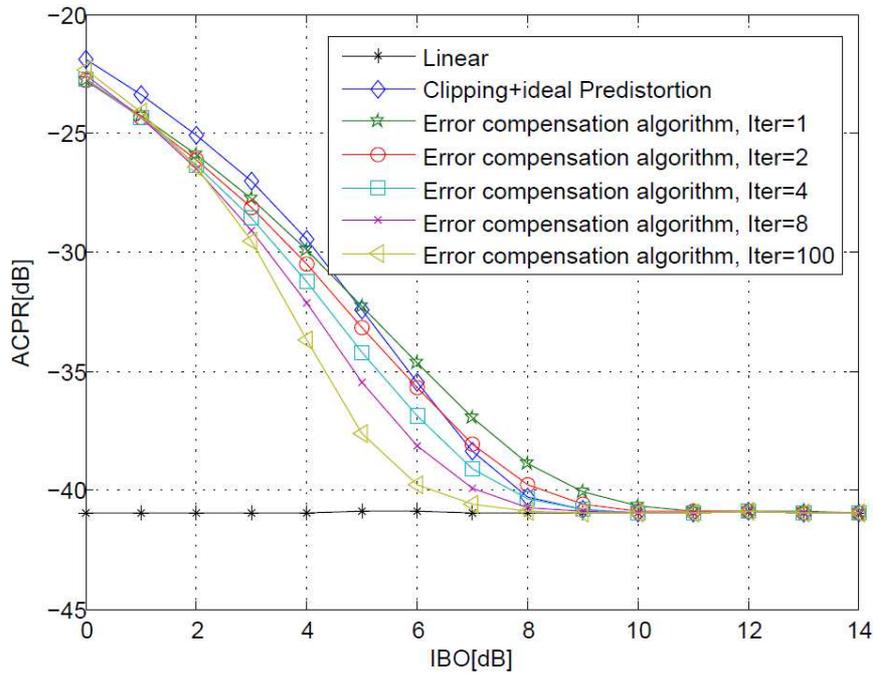
Figure 51 : ACPR performance of the error compensation algorithm depending on the PA IBO.
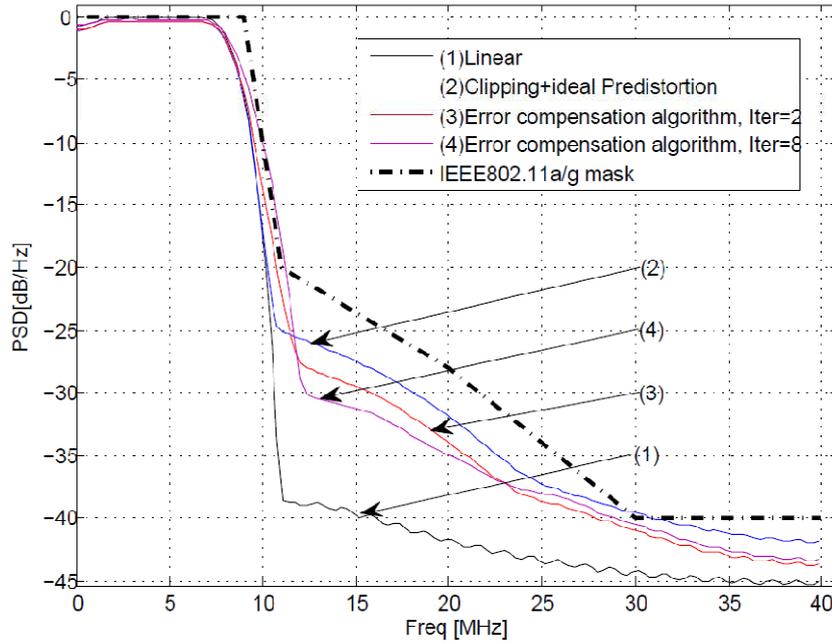
Figure 52 : PSD performance of the error compensation algorithm for an IBO value equal to 4dB.

Figure 51 and Figure 52 illustrate the ability of the error compensation algorithm to mitigate the out-of-band interference through ACPR and PSD simulations. The simulation is done over 105 randomly generated OFDM symbols. We notice that like in EVM case, the out-of-band interference mitigation improves proportionally to the number of iterations. For high number of iterations, an important PSD is noticeable around 11MHz due the part of the joint additional signal added on reserved sub-carriers. So, it is important to avoid high number of iterations and stay below the spectral mask of the standard. From the discussions above, we conclude that the error compensation algorithm is efficient with a small number of iterations for a joint approach of PAPR reduction and predistortion.

3.2.4.3   *Performance comparison of the two proposed algorithms*

Here, we conduct some simulations in order to compare the two proposed algorithms. We also compare both to a similar algorithm proposed in literature by R. N. Braithwaite in [20]. To do so, the polynomials-based algorithm is simulated with two different PAPR reduction techniques. The first one is C&F and the second one is Tone Reservation (TR)-Clipping [21]. TR-Clipping is an iterative clipping technique where the clipping noise is added on reserved sub-carriers as a tone reservation technique. So there is no BER degradation. For both clipping techniques, the clipping threshold is equal to the PA saturation.
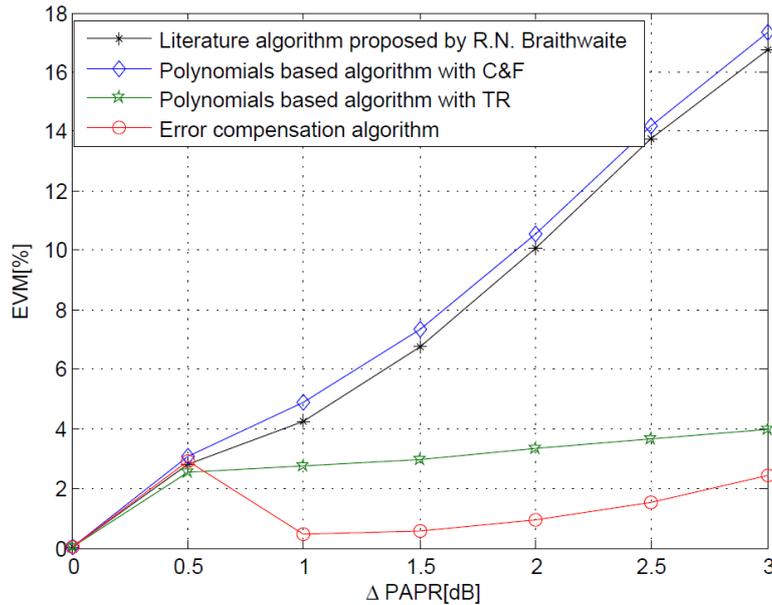
Figure 53 : EVM versus PAPR reduction gain (ΔPAPR) for the two proposed algorithms and the literature algorithm.

Figure 53 shows EVM performance depending on the PAPR reduction gain for an IBO = 4dB. Simulation is done over 104 randomly generated OFDM symbols. For a same PAPR reduction gain, we first notice that when using C&F for PAPR reduction, the EVM of the proposed polynomials-based algorithm is high compared to the case where TR-Clipping is used. This is quite obvious because of the noise of in-band clipping. When using TR-Clipping, there is no BER degradation due to the PAPR reduction. In this latter case, the EVM of the polynomials-based algorithm is low but it is not equal to zero because of the PA saturation and the accuracy of the polynomials-based algorithm. Compared to both C&F or TR-Clipping cases of the polynomials-based algorithm, the error compensation algorithm has better performance, its EVM is lower. This can be explained by the efficiency of the predistortion by adding signal on useful sub-carriers included in the error compensation algorithm. Moreover, thanks to the feedback path, the predistortion is more accurate compared to the case of the polynomials-based algorithm where the PA is modelled by a polynomial with a finite degree. Simulated in the same above conditions, we notice that the literature algorithm [20] has similar performance to the proposed polynomial algorithm with C&F. Table 6 compares the estimated complexity (i.e numbers of operations) of the algorithms. Literature algorithm is based on Least Mean Square(LMS) algorithms whose maximum complexity can be estimated to O(NL x m^3). Comparison results show that the error compensation has low complexity compared to that of proposed polynomial based algorithm and the literature one.

Thus, the error compensation algorithm is promising for a system implementing joint approach by adding signal. The PA IBO and the number of iterations must be judiciously chosen in order to keep the complexity low with good linearity and efficiency.

Table 6 : Complexity study

Algorithms	Estimated complexity	Complexity, number of operations $N=64, L=4, m=7, R=10$
Literature algorithm	$O(NL*m^3)$	$O(87\,808)$
Polynomials based algorithm	$O(NL*m^3)$	$O(87\,808)$
Error compensation algorithm	$O(R*NLlog_2NL)$	$O(20\,480)$

3.2.5   Conclusions

In this study, we first introduced PAPR reduction and predistortion by adding signal based on the Bussgang's theorem. This enables a common vision of PAPR reduction and predistortion and finally a joint approach by adding signal. Then, we proposed two algorithms for generating the joint additional signal. The first one expresses the PA characteristic as a polynomial and generates the joint additional signal using a polynomials' roots finding method. A preliminary step of PA characteristic estimation and the generation of a peaks reducing signal are necessary. The second algorithm requires neither PA estimation nor the generation of a peaks reducing signal. It is an error compensation algorithm. Simulations of CCDF, EVM, ACPR and PSD validate the performance of these two proposed algorithms. Both are efficient to simultaneously reduce the PAPR of the input signal and to increase the linearity of the PA. Compared to the polynomials-based algorithm, the error compensation algorithm has better performance and low complexity. In future work, its implementation is envisaged in a real system.

3.3   Energy efficiency strategies over different power levels in the power amplifier

This innovation corresponds to scenario 2.9.2 in SHARING deliverable D2.2 [5].

The PA is usually designed for maximum load, performing the highest energy efficiency at maximum RF output power. Nevertheless when the traffic load decreases, lower RF output power levels are required and the energy efficiency gets worse due to the PA power characteristics. Therefore, a reconfigurable PA is proposed defining different operating points which can be optimized to different power levels, providing energy savings at medium and low traffic load in the base station radio equipment.

This solution could be convenient for different base station types because it is applied at hardware level. It will be evaluated at different maximum output levels and different amplifier technologies. Concerning base station type, maximum mean output power per carrier in macrocell scenario is around 46 dBm, 38 dBm in microcell scenario, 24 dBm in picocell scenario and 20 dBm in femtocell scenario [10]. Regarding amplifier technologies, Laterally Diffused MOS (LDMOS) and Gallium Nitrate (GaN) technologies will be evaluated. LDMOS technology is the leading RF power technology for base station applications, since it offers the best values in terms of performance over cost ratio. However, GaN technology is presented as the best long term replacement of the LDMOS.

This solution presents promising energy efficiency improvements at base station radio equipment in macro and microcell scenarios, because the PA is really significant in DC power consumption breakdown [11]. PA could correspond to 64% DC power consumption breakdown in macrocell scenarios and 47% in microcell scenarios.

In SHARING deliverable D4.1 [11], a 24-hour traffic load profile for urban areas was analyzed, defining time ratios for different traffic load (see Figure 54). It can be easily observed that during most of the day there is medium and low traffic load, and consequently there is a significant potential to improve energy efficiency using a reconfigurable PA. During 29.9% of the time, the traffic load is lower than 40% of signal load (considered as low traffic load), and during 48.1% of the time, the traffic load is between 80% and 40% of signal load (considered as medium traffic load).
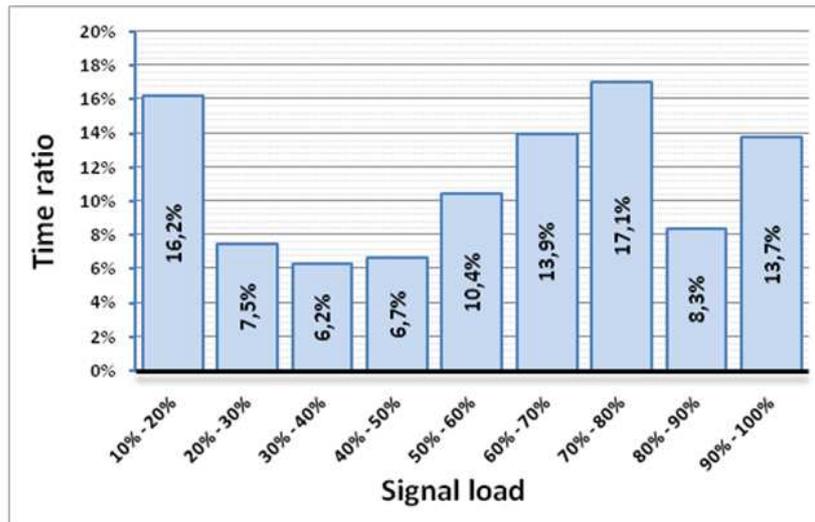
Figure 54 : Time ratio for different signal load ranges in an urban area during a day.

To evaluate PA reconfigurability, five different ranges were defined related to signal load during a day, as specified in Table 7.

Table 7 : Relation between signal load and time ratio for an urban area during a day.

Traffic load	Signal load (%)	Time ratio (%)
High	80 – 100 %	22 %
Medium	60 – 80 %	31 %
	40 – 60 %	17.1 %
Low	20 – 40 %	13.7 %
	0 – 20 %	16.2 %

The proposed signal load adaptive PA will work in different states to enable the required RF output power according to the traffic load. These intermediate states are performed applying different operating points to the PA instead of a unique one, reducing the energy consumption in medium and low traffic loads.

Table 8 presents a relation between the signal load ranges and the required RF output power. At maximum traffic load, the required RF output power ($Pout$) will reach the highest value, while the required RF $Pout$ will decrease when the signal load reduces.

Table 8 :Mean power level per carrier according to the signal load ranges.

Signal load	Required RF Pout
80 – 100 %	$Pout_{MAX}$
60 – 80 %	$Pout_{MAX} - 1\,dB$
40 – 60 %	$Pout_{MAX} - 2.2\,dB$
20 – 40 %	$Pout_{MAX} - 4\,dB$
0 – 20 %	$Pout_{MAX} - 7\,dB$

LTE downlink signals present high PAPR and demand high Output Back-Off (OBO) power levels at PA to be able to fulfill 3GPP specifications [10]. A typical OBO level due to PAPR restrictions could be estimated around 9 dB. This value was used in the analysis of proposed reconfigurable PA.

To provide the required RF output power according to signal load, different operating points will be applied in PA. The operating point adjustment could be performed modifying gate voltage and/or drain voltage at PA. However, all simulations were performed with drain voltage tuning, because gate voltage tuning doesn't offer such wide output power level adjustment as drain voltage.

Different commercial PAs were selected to perform non-linear simulations to evaluate PA performance applying different operating points. LTE band 7 was selected to perform all simulations and specifically at 2650MHz. The simulations were done over different transistor technologies, GaN and LDMOS, and different output power levels related to macrocell scenarios and microcell scenarios. The selected commercial PAs for the evaluation are listed in Table 9.

Table 9 :List of commercial PAs evaluated

Power amplifier	Transistor technology	Maximum output power (dBm)	Mean output power (dBm)	Scenario
CGHV27200 (Cree)	GaN	51 dBm	42 dBm	Macrocell
CGHV27030 (Cree)	GaN	43 dBm	34 dBm	Microcell
MRF7S27130HR3 (Freescale)	LDMOS	52 dBm	43 dBm	Macrocell
AFT20S015N (Freescale)	LDMOS	41 dBm	32 dBm	Microcell

For the evaluation, PA non-linear models were used which were provided by manufacturers. Using these models, non-linear simulations could be performed for evaluating power gain and Power-Added-Efficiency (PAE) for different operating points.

3.3.1   CGHV27200 evaluation

Firstly, CGHV27200 from Cree which delivers up to 51dBm was analyzed. The recommended operating point provided by manufacturer is 50 V drain voltage and 1 A current consumption (related to -2.6 V gate voltage). Several simulations at different operating points were performed to define which operating point is more appropriate regarding traffic load. The target of this study is to compare energy efficiency results between a conventional PA with a fixed operating point (recommended operating point) and the proposed reconfigurable PA with different selectable operating points.

Figure 55-a shows DC performance varying gate voltage and drain voltage. Meanwhile Figure 55b presents power gain performance modifying drain voltage from 50 V to 15 V. It is clearly observed that maximum RF output power level can be adjusted by varying only the drain voltage.

a)   b)
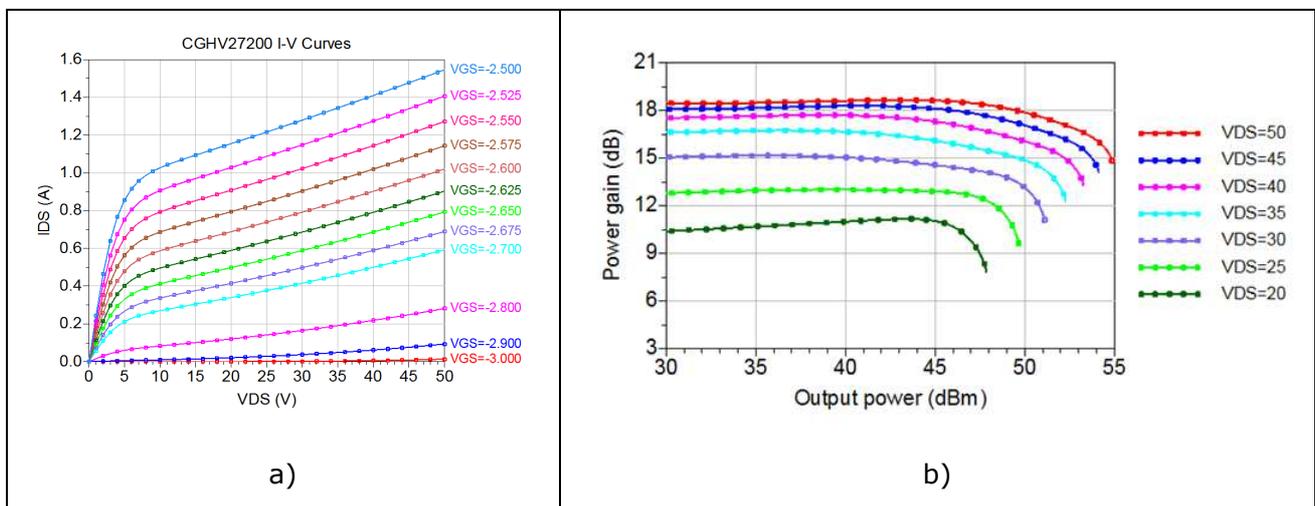
Figure 55 : a) DC performance in CGHV27200 using different operating points and b) Power gain performance in CGHV27200 using different operating points.

Figure 56 presents PAE results varying drain voltage from 50 V to 15 V. Lower drain voltages provide lower RF output power levels as well, but they present higher PAE performance compared to the recommended operating point implemented in conventional PAs. In case of low or medium traffic load demand, lower drain voltages could provide enough RF output power levels while improving energy efficiency.
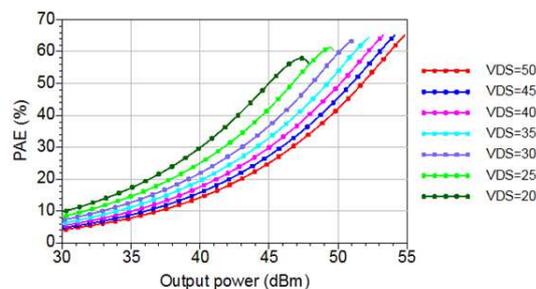
Figure 56 : PAE performance in CGHV27200 using different operating points.

Appropriate operating points in CGHV27200 according to traffic load were calculated. CGHV27200 is configured to deliver 51 dBm at 1dB compression output power for 50 V drain voltage, representing the conventional PA with the fixed operating point (recommended operating point). Then the maximum mean output power is 42 dBm (9dB OBO level) related to maximum signal load. Taking as reference Table 8 the required output power levels are defined. Table 10 summarizes the analysis of CGHV27200 using a single operating point (conventional PA) and different operating points (reconfigurable PA). Table 10 presents the potential PAE improvement related to traffic load during a day in an urban area using the proposed solution.

Table 10 :Analysis of different operating points in CGHV27200 related to traffic load.

Signal load (%)	Required RF mean Pout (dBm)		Fixed operating point		Reconfigurable operating point	
			Drain voltage	PAE (%)	Drain voltage (V)	PAE (%)
80 – 100 %	$Pout_{MAX}$	42 dBm	50 V	17.8 %	50 V	17.8 %
60 – 80 %	$Pout_{MAX} - 1\ dB$	41 dBm	50 V	16 %	45 V	17.7 %
40 – 60 %	$Pout_{MAX} - 2.2\ dB$	39.8 dBm	50 V	13.9 %	40 V	17.2 %
20 – 40 %	$Pout_{MAX} - 4\ dB$	38 dBm	50 V	11.3 %	30 V	17.7 %
0 – 20 %	$Pout_{MAX} - 7\ dB$	35 dBm	50 V	7.9 %	20 V	17.4 %

In Table 7, time ratio associated to signal load during a day in an urban area was presented. Now the next step is to link PAE improvements due to dynamic operating points with traffic load during a day. Table 11 summarizes PAE improvement at different signal load ranges and evaluates the total PAE improvement during a PA using the proposed solution versus the conventional one. The proposed dynamically reconfigurable PA could improve 20.1% PAE performance in an urban area.

Table 11 :PAE improvement in CGHV27200 related to signal load and time ratio for an urban area using reconfigurable operating points.

Traffic load	Signal load (%)	Time ratio (%)	PAE improvement (%)
High	80 – 100 %	22 %	0 %
Medium	60 – 80 %	31 %	9.6 %
	40 – 60 %	17.1 %	19.2 %
Low	20 – 40 %	13.7 %	36.1 %
	0 – 20 %	16.2 %	54.6 %
Total PAE improvement (%)			20.1 %

3.3.2   CGHV27030 evaluation

Secondly, commercial PA CGHV27030 from Cree was studied delivering up to 43 dBm (Figure 57 and Figure 58). Then the maximum mean output power for this component is 34 dBm, (applying 9 dB OBO) representing a microcell scenario. The recommended operating point provided by manufacturer is 28 V drain voltage and 150 mA current consumption (related to -2.98 V gate voltage). As in the previous analysis, suitable operating points were calculated related to traffic demand and results are summarized in Table 12 and Table 13.

Figure 57 : a) DC performance in CGHV27030 using different operating points and b) Power gain performance in CGHV27030 using different operating points.

Figure 58 : PAE performance in CGHV27030 using different operating points.

Drain voltage was modified from 28 V to 12 V to define suitable operating points for each signal load range.

Table 12 : Analysis of different operating points in CGHV27030 related to traffic load.

Signal load (%)	Required RF mean Pout (dBm)		Fixed operating point		Reconfigurable operating point	
			Drain voltage	PAE (%)	Drain voltage (v)	PAE (%)
80 – 100 %	$Pout_{MAX}$	34 dBm	28 V	17.9 %	28 V	17.9 %
60 – 80 %	$Pout_{MAX} - 1\,dB$	33 dBm	28 V	15.9 %	25 V	17.5 %
40 – 60 %	$Pout_{MAX} - 2.2\,dB$	31.8 dBm	28 V	13.8 %	22 V	17.1 %
20 – 40 %	$Pout_{MAX} - 4\,dB$	30 dBm	28 V	11 %	18 V	16.7 %
0 – 20 %	$Pout_{MAX} - 7\,dB$	27 dBm	28 V	7.3 %	14 V	14.5 %

Table 13 :PAE improvement in CGHV27030 related to signal load and time ratio for an urban area using reconfigurable operating points.

Traffic load	Signal load (%)	Time ratio (%)	PAE improvement (%)
High	80 – 100 %	22 %	0 %
Medium	60 – 80 %	31 %	9.1 %
	40 – 60 %	17.1 %	19.3 %
Low	20 – 40 %	13.7 %	34.1 %
	0 – 20 %	16.2 %	49.6 %
Total PAE improvement (%)			18.8 %

In this evaluation, PAE improvement using the proposed dynamically reconfigurable PA in an urban area during a day is 18.8 %.

3.3.3   MRF7S27130HR3 evaluation

Thirdly, commercial PA MRF7S27130HR3 from Freescale which delivers up to 52 dBm at 1 dB compression output power was analyzed. The recommended operating point is 28V drain voltage and 1500 mA current consumption (related to 2.7 V gate voltage). Figure 59 and Figure 60 present DC performance, power gain performance and PAE performance for MRF7S27130HR3 using different operating points.

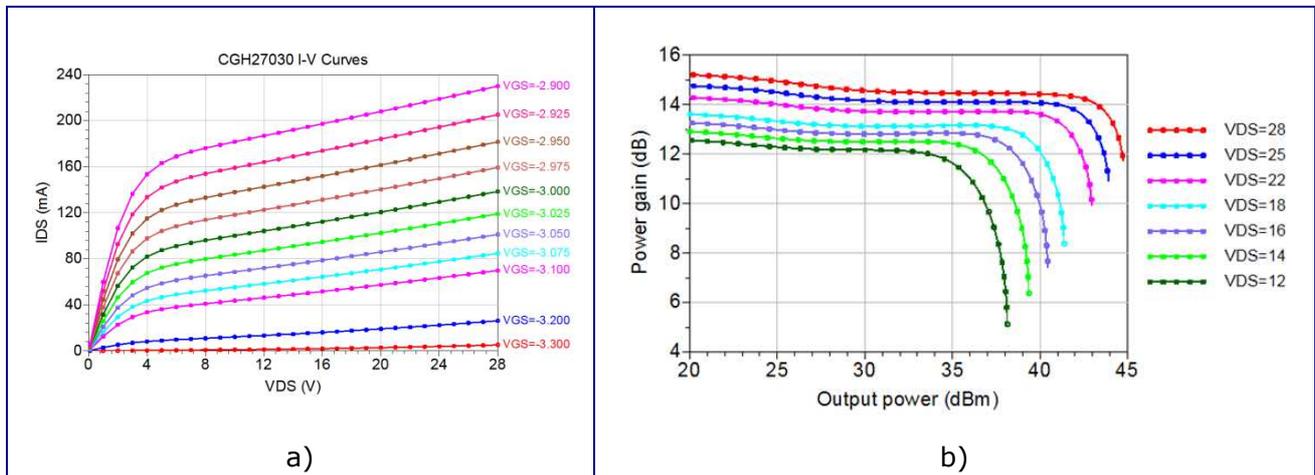a)   b)
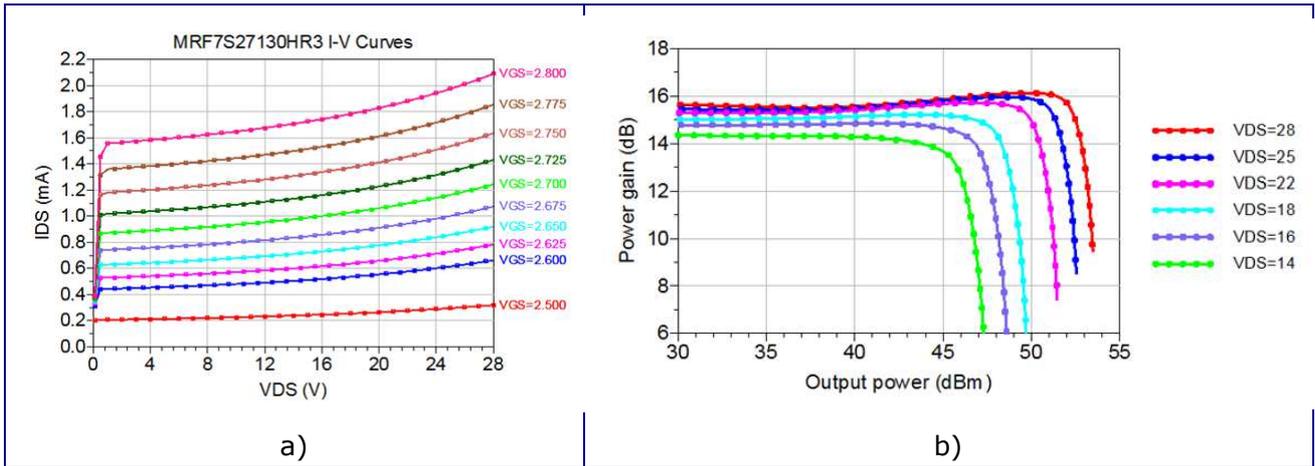
Figure 59 : a) DC performance in MRF7S27130HR3 using different operating points and b) Power gain performance in MRF7S27130HR3 using different operating points.
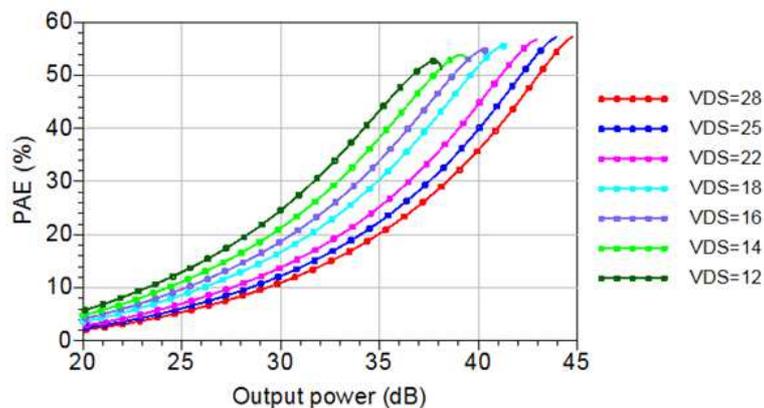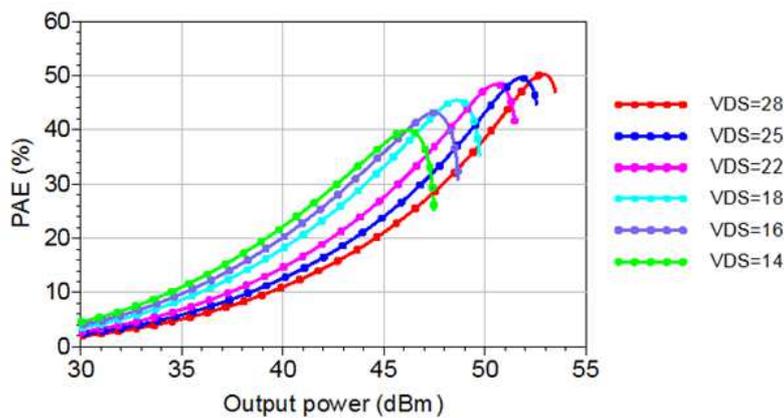

Figure 60 :PAE performance in MRF7S27130HR3 using different operating points.

Table 14 : Analysis of different operating points in MRF7S27130HR3 related to traffic load.

Signal load (%)	Required RF mean Pout (dBm)		Fixed operating point		Reconfigurable operating point	
			Drain voltage (V)	PAE (%)	Drain voltage (v)	PAE (%)
80 – 100 %	$Pout_{MAX}$	43 dBm	28 V	16.4 %	28 V	16.4 %
60 – 80 %	$Pout_{MAX} - 1\,dB$	42 dBm	28 V	14.6 %	25 V	16.4 %
40 – 60 %	$Pout_{MAX} - 2.2\,dB$	40.8 dBm	28 V	11.8 %	22 V	16.3 %
20 – 40 %	$Pout_{MAX} - 4\,dB$	39 dBm	28 V	9.5 %	18 V	15.9 %
0 – 20 %	$Pout_{MAX} - 7\,dB$	36 dBm	28 V	6 %	14 V	12.7 %

Table 15 : PAE improvement in MRF7S27130HR3 related to signal load and time ratio for an urban area using reconfigurable operating points

Traffic load	Signal load (%)	Time ratio (%)	PAE improvement (%)
High	80 – 100 %	22 %	0 %
Medium	60 – 80 %	31 %	10.9 %
	40 – 60 %	17.1 %	27.6 %
Low	20 – 40 %	13.7 %	40.2 %
	0 – 20 %	16.2 %	52.7 %
Total PAE improvement (%)			22.1 %

Appropriate operating points were defined between 28 V and 14 V drain voltage. As previously, the potential PAE improvement in an urban area during a day implementing the proposed solution was evaluated. Using MRF7S27130HR3, which is LDMOS transistor for macrocell scenarios, the expected PAE improvement is 22.1%.

3.3.4   AFT20S015N evaluation

Finally, LDMOS PA AFT20S015 from Freescale was evaluated. The recommended operating point provided by manufacturer is 28 V drain voltage and 130 mA current consumption (related to -2.88 V gate voltage). Figure 61 and Figure 62 show PA performance using different operating points.

a)   b)
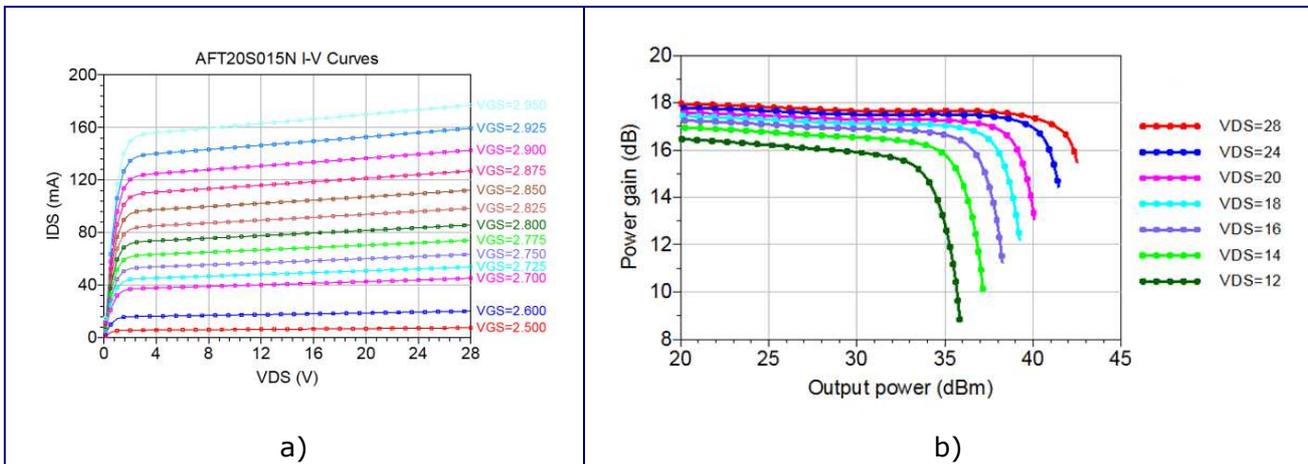
Figure 61 : a) DC performance in AFT20S015N using different operating points and b) Power performance in AFT20S015N using different operating points.
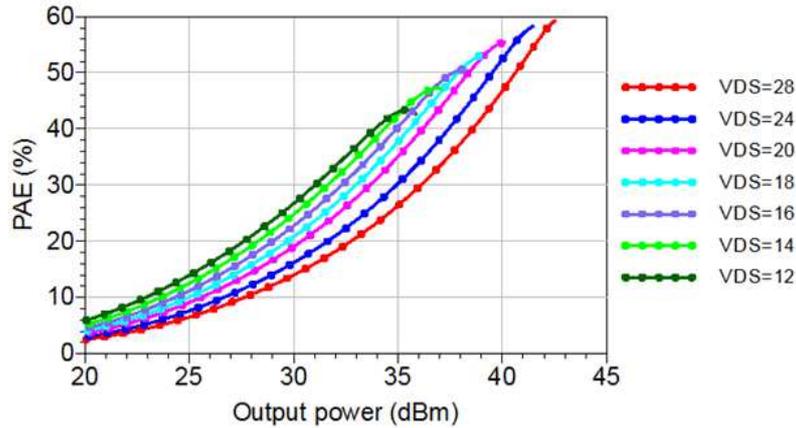
Figure 62 : PAE performance in AFT20S015N using different operating points.

Table 16 and Table 17 summarize the evaluation results in AFT20S015N using different operating points. In this evaluation, PAE improvement reaches 20.9% in an urban area implementing the proposed reconfigurable PA.

Table 16 : Analysis of different operating points in AFT20S015N related to traffic load.

Signal load (%)	Required RF mean Pout (dBm)		Fixed operating point		Reconfigurable operating point	
			Drain voltage	PAE (%)	Drain voltage (v)	PAE (%)
80 – 100 %	$Pout_{MAX}$	32 dBm	28 V	18.3 %	28 V	18.3 %
60 – 80 %	$Pout_{MAX} - 1\,dB$	31 dBm	28 V	16 %	24 V	18.4 %
40 – 60 %	$Pout_{MAX} - 2.2\,dB$	29.8 dBm	28 V	13.7 %	20 V	18.7 %
20 – 40 %	$Pout_{MAX} - 4\,dB$	28 dBm	28 V	10.6 %	18 V	15.8 %
0 – 20 %	$Pout_{MAX} - 7\,dB$	25 dBm	28 V	6.5 %	14 V	12.6 %

Table 17 : PAE improvement in AFT20S015N related to signal load and time ratio for an urban area using reconfigurable operating points.

Traffic load	Signal load (%)	Time ratio (%)	PAE improvement (%)
High	80 – 100 %	22 %	0 %
Medium	60 – 80 %	31 %	13 %
	40 – 60 %	17.1 %	26.7 %
Low	20 – 40 %	13.7 %	32.9 %
	0 – 20 %	16.2 %	48.4 %
Total PAE improvement (%)			20.9 %

3.3.5   Conclusions

After evaluating different commercial PAs through non-linear simulations Table 18 summarizes PAE improvement results. Different transistor technologies, GaN and LDMOS, and different output power levels related to macrocell scenarios and microcell scenarios were analyzed. This solution is applied at hardware level so it could be applied in different base stations.

Table 18 : Summary of evaluation results.

Power amplifier	Transistor technology	Maximum output power	Scenario	Total PAE improvement (%)
CGHV27200	GaN	200 W	Macrocell	20.1 %
CGHV27030	GaN	30 W	Microcell	18.8 %
MRF7S27130HR3	LDMOS	130 W	Macrocell	22.1 %
AFT20S015N	LDMOS	15 W	Microcell	20.9 %

PAE improvement could be estimated in 20% using the proposed dynamically reconfigurable solution versus a conventional PA.

In SHARING deliverable D4.1 [11], DC power consumption breakdown for different types of base stations at maximum load was revisited. As PA could represent 64% DC power consumption breakdown in macrocell scenarios, and 47% in microcell scenarios, the energy efficiency improvement diary in an urban area could be around 13% and 9% respectively.

4   CONCLUSION

This deliverable has provided results from SHARING Work Package 4, Task 4.3, on new opportunities, challenges and innovative concepts for energy saving in heterogeneous network deployments from both the RRM and the base station radio equipment points of view.

The first part of this deliverable covered RRM energy saving mechanisms and especially the base station ON/OFF mechanisms.

An eNodeB sleep mode mechanism has been first proposed, where the pico eNodeB is switched to sleep mode as soon as it has been idle a given time. In this scheme the decision to enter the sleep mode is made by the pico eNodeB itself, without the control of the neighboring eNodeBs. When it comes to the reactivation of a sleeping pico eNodeB, a number of different schemes can be implemented, based either on the monitored macrocell load or on the level of the received uplink interference, measured by the sleeping pico eNodeB itself. Furthermore, a combination of these two schemes is possible.

As the design of energy-efficient mechanisms is one of the key challenges in emerging wireless small cell networks, a novel approach (based on learning) for opportunistically switching base stations on and off to improve the energy efficiency in wireless small cell networks has been proposed. The proposed approach enables the small cell base stations to optimize their downlink performance while balancing the load among each other and satisfying the users' QoS requirements.

In the context of base station sleep mode, the compensation of coverage by nearby base stations is a key issue. Then, an update on compensation-based ON/OFF energy saving has been presented where the compensating base stations are determined through dominant sets calculated from the neighborhood graph of the campus. The results have shown that the proposed ON/OFF energy saving technique based on dominant sets is effective in both energy saving and the optimization of the overall system performance. Compensation based ON/OFF energy saving technique based on maximum independent sets of the graph is showing the best performance.

All the aforementioned studies have been confirmed by system-level simulations based on realistic radio propagation models, which highlight the interest in switching off small-cells outside busy hours to optimize the network energy consumption. A given scenario shows that a 29% gain can be achieved when switching off small-cells having less than 3% of traffic load, without compromising the user QoS. The service outage observed at peak hours is partly due to the Almost Blank Sub frame duty cycle limitation and suggests other SON algorithms to achieve a finer macro to small-cell offloading depending on the time of the day. In addition, theoretical measurements of the power saving are derived to validate smart base station ON/OFF switching principle in periods where the total amount of traffic served by the network is such that can be satisfied with just a subset of active base stations.

The second part of this document discussed proposals to save energy in the base station equipment. These proposals were focused on the power amplifier whose consumption is huge compared to the total power of the transmitter.

The first study proposed to smartly merge the peak factor reduction (also known as PAPR reduction) and linearisation steps prior to power amplification. Indeed, the smaller the peak factor of transmitted waveforms, the larger the power amplifier efficiency. Results have shown that this approach is able to head the transmission performance by jointly tuning linearity and peak factor reduction depending on needs and load which directly mitigate the power amplifier consumption.

Last, a study has been proposed to adapt the operating point of the power amplifier in order to reduce the power consumption. Indeed, this component is usually designed for maximum load, performing the highest energy efficiency at maximum RF output power. Nevertheless, when the traffic load decreases, lower RF output power levels are required and the energy efficiency gets worse due to the power amplifier characteristics. Therefore, a reconfigurable

power amplifier was proposed defining different operating points which can be optimized to different power levels, providing energy savings at medium and low traffic load in the base station radio equipment. After evaluating different commercial PAs through simulations, different transistor technologies, GaN and LDMOS, and different output power levels related to macrocell and microcell scenarios have been analyzed. This approach has been applied to hardware level so it could be applied in different base stations. PAE improvement could be estimated as 20% using the proposed dynamically reconfigurable solution versus a conventional PA. As power amplifier could represent 64% DC power consumption breakdown in macrocell scenarios, and 47% in microcell scenarios, the energy efficiency improvement diary in an urban area could be around 13% and 9% respectively.

REFERENCES

[1] INFSO-ICT-247733 EARTH, Deliverable D4.3, "Final Report on Green Radio Technologies", 2012.

[2] INFSO-ICT-24773 EARTH, deliverable D2.3, "Energy efficiency analysis of the reference systems , areas of improvements and target breakdown", January 2011.

[3] Per Skillermark, Pål Frenger, "Enhancing Energy in LTE with Antenna Muting", in Proc. IEEE Vehicular Technology Conference (VTC) 2012 Spring, Yokohama, Japan, May 2012.

[4] Laetitia Falconetti, Pål Frenger, Harald Kallin, Thomas Rimhagen, "Energy Efficiency in Heterogeneous Networks", in Proc. IEEE Online Conference on Green Communications (GreenCom) 2012.

[5] Celtic-Plus SHARING, Deliverable D2.2, "Scenarios, KPIs and Evaluation Methodology for Advanced Cellular Systems", 2014.

[6] 3GPP TSG RAN WG1, R1-114336, "Base Station Power Model", San Francisco, CA, USA, November 2011.

[7] 3GPP TR36.887, "Study on energy saving enhancement for E-UTRAN", rel 12, June 2014.

[8] 3GPP TR36.814, "Further advancements for E-UTRA physical layer aspects", rel9, march 2010.

[9] Kimmo Hiltunen, "Utilizing eNodeB Sleep Mode to Improve the Energy-Efficiency of Dense LTE Networks", in Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC) 2013, London, The United Kingdom, September 2013.

[10] 3GPP TS 36.104, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Base station (BS) radio transmission and reception. Release 12"

[11] Celtic-Plus SHARING, Deliverable D4.1, "D4.1 New opportunities, challenges and innovative concept candidates for SON-Heterogeneous Networks", 2014.

[12] Julian Jakob Bussgang. Cross correlation functions of amplitude-distorted gaussian signals. Research laboratory of electronics, Massachusetts Institute of Technology, Cambridge. Technical Report 216, 1952. "March 26, 1952.".

[13] Paolo Banelli, Geert Leus, and Georgios B Giannakis. Bayesian estimation of clipped gaussian processes with application to OFDM. In Proc. EUSIPCO 2002, volume 1, pages 181–184, 2002.

[14] Désiré Guel and Jacques Palicot. OFDM PAPR reduction based on nonlinear functions without BER degradation and out-of-band emission. In Proceedings of ICSPC 2009, Singapore, May 2009.

[15] O.Abel Gouba and Yves Louet. Digital predistortion expressed as an adding signal technique in OFDM context. In New Circuits and Systems Conference (NEWCAS), 2013 IEEE 11th International, pages 1–4, June 2013.

[16] O. Abel Gouba and Yves Louet. Adding signal for peak-to-average power reduction and predistortion in an orthogonal frequency division multiplexing context. IET Signal Processing, 7(9):879–887, December 2013.

[17] Peter B. Kenington. High-linearity RF amplifier design. Artech House, 2000.

[18] Yves Louet and Jacques Palicot. A classification of methods for efficient power amplification of signals. Annales des Télécommunications, 63(2008-08-07):351–368, August 2008.

[19] Jose Tellado-Mourelo. Peak to average power reduction for multicarrier modulation. PhD thesis, Stanford University, September 2000. Trans.

[20] R.N. Braithwaite. A combined approach to digital predistortion and crest factor

reduction for the linearization of an RF power amplifier. Microwave Theory and Techniques, IEEE Transactions on, 61(1):291–302, 2013.

[21]   Desire Guel and Jacques Palicot. OFDM PAPR reduction based on nonlinear functions without BER degradation and out-of-band emission. In Proceedings of ICSPC 2009, Singapore, Singapore, May 2009.

[22]   Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2013-2018", February 2014.

[23]   R. Q. Hu, Y. Qian, S. Kota, and G. Giambene, "Hetnets-a new paradigm for increasing cellular capacity and coverage [guest editorial]," Wireless Communications, IEEE, vol. 18, no. 3, pp. 8-9, 2011.

[24]   H. Helantera, "Measuring mobile broadband data traffic – End user and access network centric approach", 10th World Telecommunication/ICT Indicators Meeting (WTIM-12), Bangkok, September 25-27, 2012.

[25]   K.I. Pedersen, Y. Wang, S.Strzyz and F. Frederiksen, "Enhanced inter-cell interference coordination in co-channel multi-layer LTE-advanced networks", Wireless Communications, IEEE, 20.3: 120-127, June 2013.

[26]   M. Brau, Y. Corre and Y. Lostanlen, "Assessment of 3D network coverage performance from dense small-cell LTE", IEEE International Conference on Communications (ICC) 2012, Ottawa, June 2012.

[27]   Y. Corre and Y. Lostanlen, "Three-dimensional urban EM wave propagation model for radio network planning and optimization over large areas", IEEE Transactions on Vehicular Technology, Vol. 58, No. 7, pp. 3112-3123, Sept. 2009.

[28]   H. Klessig, A. J. Fehske and G. P. Fettweis, "Energy efficiency gains in interference-limited heterogeneous cellular mobile radio networks with random micro site deployment", in Proc. 34th IEEE Sarnoff Symposium, Princeton, New Jersey, USA, 2011.

[29]   The Climate Group, *SMART 2020: Enabling Low Carbon Economy in Information Age*, June 2008. Available online: http://www.theclimategroup.org/publications

[30]   C. Bianco, F. Cucchietti, and G. Griffa, "Energy consumption trends in the next generation access network; a telco perspective," in Telecommunications Energy Conference, 2007. INTELEC 2007. 29th International, 10 2007, pp. 737 –742.

[31]   L. Souchon, B. Aebischer, J. Roturier, and F. Flipo, "Infrastructure of the information society and its energy demand," in ECEEE 2007 Summer Study. Panel 6. Products and appliances, 2007.

[32]   http://www.energystar.gov/

[33]   http://www.greentouch.org

[34]   Gergely Biczók, Jens Malmodin, Albrecht Fehske, « Economic and Ecological Impact of ICT », Deliverable D2.1 EARTH project. Available at https://bscw.ict-earth.eu/pub/bscw.cgi/d38532/EARTH_WP2_D2.1_v2.pdf

[35]   http://www.fp7-trend.eu/

[36]   J. Restrepo, C. Gruber, and C. Machuca, "Energy profile aware routing," in Proceedings of First International Workshop on Green Communications, 2009.

[37]   Miao, G; Himayat, N.; Li, G. "Energy-efficient ling adaption in frequency-selective channels", IEEE Trans. On Comm. Vol 58, p. 545-554, Feb. 2010.

[38]   F. Haider, C.-X. Wang, X. Hong, H. Haas, D. Yuan and E. Hepsaydir, "Spectral-energy efficiency tradeoff in cognitive radio networks with peak interference power constraints". In Proc. 13Th International Conference on Communication Technology, (Jinan, China) September 2011.

[39]   J. Chabarek, J. Sommers, P. Barford, C. Estan, D. Tsiang, and S. Wright, "Power awareness in network design and routing," in Proceedings of IEEE INFOCOM, 2008.

[40]   J. Restrepo, C. Gruber, and C. Machuca, "Energy profile aware routing," in Proceedings of First International Workshop on Green Communications, 2009.

[41]   L. Chiaraviglio, D. Ciullo, M. Mellia, and E. Leonardi, "How much can the internet be greened?" in Proceedings of Second International Workshop on Green Communications, 2009

[42]   R. Ahlswede, N. Cai, S.-Y. R. Li, R. W. Yeung, « Network Information Flow », *IEEE Transactions on Information Theory*, vol. 46, No. 4, July 2000.

[43]   R. Koetter, M. Médard. "An Algebraic Approach to Network Coding," IEEE/ACM Transactions on Networking, Volume 11, Issue 5, October 2003, pp. 782-796.

[44]   C. Tao, C. Lijun, T. Ho, "Energy Efficient Opportunistic Network Coding for Wireless Networks", *INFOCOM*, 2008.

[45]   C. Fragouli, J. Widmer, J.-Y. Le Boudec, "Efficient Braodcasting Using Network Coding", IEEE/ACM Transactions on Networking, Vol. 16, no. 2, April 2008, pp. 450-463.

[46]   Xiaomeng Shi, Muriel Médard, and Daniel E. Lucani, "When Both Transmitting and Receiving Energies Matter: An Application of Network Coding in Wireless Body Area Networks", available at http://arxiv.org/pdf/1008.0143.pdf

[47]   Peter Briggs, Rajesh Chundury, Jonathan Olsson, "Carrier Internet for Mobile Backhaul", IEEE Communications Magazine, October 2010, Vol 48, N° 10, PP94-100

[48]   DoronEzri ,"Optimizing the Cellular Network for Minimal Emission from Mobile", Green Cellular Next Generation Wireless Green Networks Workshop, 5-6 November 2009, SUPELEC, France

[49]   E. Altman, C. Hasan, J.-M. Gorce, L. Roullet. Green Networking: Downlink Considerations. *NetGCOOP2011 : International conference on NETwork Games, COntrol and OPtimization*, Oct 2011, Paris, France.

[50]   Eitan Altman, Manjesh Kumar Hanawal, Rachid El-Azouzi and ShlomoShamai, "Tradeoffs in Green Cellular Networks". *In Proc.* 3rd GreenMetrics Workshop (San Jose, USA, June 7, 2011).

[51]   Jacques Palicot, "Cognitive Radio: An Enabling Technology for the Green Radio Communications Concept", IWCMC'09, Leipzig, Germany June 2009.

[52]   Honggang ZHANG, "Cognitive Radio for Green Communications and Green Spectrum", COMNETS 08 Co-located with CHINACOM 08, August 25-27, 2008, Hangzhou, China.

[53]   J. Palicot, "Cross Layer Sensors for Green Cognitive Radio" *Annals of Telecom, Towards Green Mobile Networks",* 27 February 2012.

[54]   INFSO-ICT-247733 EARTH, Deliverable D4.3, "Final Report on Green Radio Technologies", 2012.

[55]   Celtic-Plus SHARING, Deliverable D4.2, "Intra-system offloading: innovative concepts and performance evaluation", 2014.

GLOSSARY

ACRONYM	DEFINITION
3GPP	Third Generation Partnership Project
ABS	Almost Blank Sub-frame
ACTS	Advanced Communications Technologies and Services
ACPR	Adjacent Channel Power Ratio
ADSL	Asymmetric Digital Subscriber Line
AMC	Adaptive Modulation and Coding
ANR	Agence Nationale de la Recherche
AP	Access Point
ARPU	Average Revenu Per User
ASIC	Application Specific Integrated Circuit
BAN	Body Area Network
BB	Baseband
BBU	Base Band Unit
BEFEMTO	Broadband evolved Femto
BER	Bit Error Rate
BRAN	Broadband Radio Access Network
BS	Base Station
BTS	Base Transceiver Station
CA	Carrier Aggregation
CAPEX	Capital Expenditure
CCIR	Comité Consultatif International des Radiocommunications
CDF	Cumulative Distribution Function
CDMA	Code Division Multiplexing Access
CDS	Connected Dominant Set

CEPT	Conférence Européenne des Postes et Télécommunications
CO	Confidential
COMP	Coordinated Multi-Point
COST	european Cooperation in Science and Technology
CR	Clipping Ratio
CR	Cognitive Radio
CRE	Cell Range Extension
CREB	Cell Range Expansion Bias
CRC	Cyclic Redundancy Check
CRS	Cell-specific Reference Signal
CS	Coordinated Scheduling
CSG	Closed Subscriber Group
CSI	Channel State Information
CSIT	Channel State Information at Transmitter
CT	Core network and Terminals
CTO	Chief Technical Officer
CTU	Chief Technical Officer
CWC	Centre for Wireless Communications
CoMP	Coordinated Multi-Point
D2D	Device to Device
DARPA	Defense Advanced Research Projects Agency
DL	Downlink
DRX	X-Ray Diffraction (in French)
DSL	Digital Subscriber Loop
DSTL	Defense Science and Technology Laboratory
DTX	Discontinuous Transmission

DVB	Digital Video Broadcasting
EC	European Commission
eICIC	Enhanced Inter-Cell Interference Cancellation
eNB	evolved Node B
EPC	Evolved Packet Core
EPON	Ethernet Passive Optical Network
ETSI	European Telecommunications Standards Institute
EU	European Union
EVM	Error Vector Magnitude
FDD	Frequency Division Duplex
FFR	Fractional Frequency Reuse
FFT	Fast Fourier Transform
FPGA	Field Programmable Gate Array
FRN	Fixed Relay Node
GA	General Assembly
GaN	Gallium Nitrate
GCR	Green Cognitive Radio
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile
GSMA	GSM Alliance
GW	Gateway
HARQ	Hybrid Automatic Repeat reQuest
HDR	Habilitation à Diriger les Recherches
HF	High Frequencies
HO	Hand Over

HSDPA	High Speed Downlink Packet Access
HSPA	High Speed Packet Access
HW	Hardware
HeNB	Home eNB
IA	Interference Alignment
IBO	Input Back Off
IC	Interference Cancellation
ICIC	Inter-Cell Interference Cancellation
ICT	Information and Communication Technology
IFFT	Inverse Fast Fourier Transform
IMT	International Mobile Telecommunications
IP	Internet Protocol
IPR	Intellectual Property Rights
ISD	Inter Site Distance
ITU	International Telecommunication Union
JP	Joint Processing
KPI	Key Performance Indicator
LAN	Local Area Network
LDPC	Low Density Parity Check
LDMOS	Literally Diffused MOS
LMS	Least Mean Square
LTE	Long Term Evolution
LTE-A	Long Term Evolution - Advanced
MAC	Medium Access Control
MBS	Macro Base Station
MC	Multi Carrier

MIMO	Multiple Input Multiple Output (MU-MIMO see MU)
MIS	Maximum Independent Set
MME	Mobility Management Entity
MRN	Mobile Relay Node
MS	Mobile Station
MTC	Machine Type Communications
MU	Multi-User
NAS	Network Access Server
NFC	Near Field Communications
NGMN	Next Generation Mobile Networks
OBO	Output Back Off
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	OFDM Access
OPEX	Operational Expenditure
OSTBC	Orthogonal Space Time Block Code
PA	Power Amplifier
PAE	Power Added Efficiency
PAPR	Peak to Average Power Ratio
PBCH	Physical Broadcast Channel
PC	Personal Computer
PDCP	Packet Data Convergence Protocol
PHY	Physical Layer
PM	Project Manager
POI	Point Of Interest
PSD	Power Spectral Density
PSS	Primary Synchronization Signal

PU	Public
QMR	Quarterly Management Report
QoS	Quality of Service
RAN	Radio Access Network
RAT	Radio Access Technology
RF	Radio Frequency
RLC	Radio Link Control
RN	Relay Node
RNC	Radio Network Controler
RoT	Rise over Thermal noise
RRC	Radio Resource Control
RRM	Radio Resource Management
RTD	Research and Technical Development
SC	Small Cell
SF	Subframe
SINR	Signal to Interference Noise Ratio
SLM	SeLected Mapping
SME	Small and Medium Enterprise
SNR	Signal to Noise Ratio
SON	Self Optimizing/Organizing Network
SSPA	Solid State Power Amplifier
SSS	Secondary Synchronization Signal
SW	Software
TA	Tracking Area
TCO	Total Cost of Ownership
TD	Time Division
TDD	Time Division Duplex

TL	Traffic Load
TM	Task Manager
TR	Technical Requirement
TTI	Transmission Time Interval
TX	Transmit
UE	User Equipment
UK	United Kingdom
UL	Uplink
UMTS	Universal Mobile Telecommunication System
UT	User Terminal
UTRA	Universal Terrestrial Radio Access
UTRAN	Universal Terrestrial Access Network
UWB	Ultra Wide Band
VNI	Visual Networking Index
VPL	Vehicle Penetration Loss
WCDMA	Wideband Code Division Multiplexing Access
Wi-Fi	Wireless Fidelity
WLAN	Wireless Local Area Network
WP	Work Package
WPL	Work Package Leader
WiFi	Wireless Fidelity