

Towards Green and Soft: 5G Design Considerations

Chih-Lin I,

Green Communication Research Center, China Mobile Research Institute

Shuangfeng Han,

Green Communication Research Center, China Mobile Research Institute

Zhikun Xu,

Green Communication Research Center, China Mobile Research Institute

Zhengang Pan

Green Communication Research Center, China Mobile Research Institute

Abstract—As the commercial deployment of 4G systems is picking up, technologists worldwide are beginning to search for next-generation wireless solutions to meet the anticipated demands in the 2020 era given the explosive growth of mobile Internet and the Internet of Things (IoT). This article presents our perspectives of the 5G technologies with two major themes: *Green* and *Soft*. By rethinking the Shannon theorem and traditional cell-centric design, network capacity can be significantly increased while network power consumption is steady or even decreased. The feasibility of the combination of *Green* and *Soft* is investigated through five interconnected areas of research: energy efficiency and spectral efficiency co-design, no more cells, rethinking signalling/control, invisible base stations, and full duplex radio.

Introduction

With the maturing of the fourth generation (4G) standardisation and the ongoing worldwide deployment of 4G cellular networks, research activities on 5G communication technologies have emerged in both the academic and industrial communities. Various organisations from different countries and regions have taken initiatives and launched programs aimed at potential key technologies of 5G: 5GNOW and METIS launched under the European Union's Seventh Framework Programme for Research (FP7) study new waveforms and the fundamentals of 5G to meet the requirements proposed for 2020; the 5G Innovation Centre was established in the United Kingdom to develop a world-class test bed of 5G technologies; the Third Generation Partnership Project (3GPP) has drawn up its draft evolution roadmap out to 2020; and China has kicked off its IMT-2020 Promotion Group to start the study of user demands, spectrum characteristics, and technology trends (see <http://www.miit.gov.cn>).

Since the research on 5G networks is still in an embryonic stage, there are no clear answers for questions such as what 5G is and what the enabling technologies are. However, the broad consensus on the 5G requirements includes higher spectral-efficiency (SE) and energy-efficiency (EE), lower end-to-end latency,

and more connection nodes. From the perspective of China Mobile Research Institute, the world of 5G should reflect two major themes: *Green* and *Soft*.

As global carbon emissions increase and sea levels rise, global weather and air pollution in many large cities across the world are becoming more severe. Consequently, energy saving has been recognised as an urgent issue worldwide. In 2012, the annual average power consumption by information and communications technologies (ICT) was over 200 GW, of which telecoms infrastructure and devices accounted for 25 percent (The Climate Group 2008). In the 5G era, the mobile traffic in wireless communications networks is no longer solely generated by traditional mobile terminals. Increasingly, driven by the widespread applications of IoT, mobile networks are being used to gather data from or to transfer data between machine-type devices. These devices include, for example, intelligent wearable devices, various environmental sensors for temperature, humidity, metering devices, and vehicular sensors for navigation, safety, and traffic management etc. Machine type communications (MTC) are typically characterised with small packet size, low tolerance of transmission delay, low mobility and low frequency of use, etc. It is forecast that the number of connected MTC devices will reach tens of billions. The massive connectivity involved in MTC may not be handled efficiently by the current wireless communication networks, like 3G and 4G. Meanwhile, networks trend to become denser and denser to provide high data rate. Dramatic improvements in EE will be needed and new tools for jointly optimising network SE and EE will be essential.

Several research groups and consortia have been investigating EE of cellular networks, including Mobile VCE, EARTH, and Green-Touch. Mobile VCE has focused on the BS hardware, architecture, and operation, realising energy saving gains of 75–92 percent in simulations (The Climate Group 2008). EARTH has devised an array of new technologies including low-loss antennas, micro discontinuous transmission (DTX), antenna muting, and adaptive sectorisation according to traffic fluctuations, resulting in energy savings of 60–70 percent with less than 5 percent throughput degradation (Skinnermark & Frenger 2012). GreenTouch (<http://www.greentouch.org>) has set up a much more ambitious goal of improving EE 1000 times by 2020. Several operators have been actively developing and deploying green technologies, including green BSs powered solely by renewable energies, and green access infrastructure such as cloud/collaborative/clean radio access network (C-RAN) (China Mobile Research Institute 2011).

Soft is the other critical characteristic of 5G networks. Carrier-grade networks are complex and composed of special-purpose nodes and hardware. New standards and features often require a variety of equipment to be developed and integrated, thus leading to very long launch cycles. In order to accommodate explosive mobile Internet traffic growth and a large number of new applications/services demanding much shorter times to market, much faster turnaround of new network capabilities is required. Dynamic RAN reconfiguration can handle both temporal and spatial domain variation of mobile traffic without over-provisioning homogeneously. *Soft* technologies are the key to resolve these issues.

By separating software and hardware, control plane and data plane, building software over general-purpose processors (GPPs) via programming interfaces and virtualisation technology, it is possible to achieve lower cost and higher efficiency using software defined networks (SDNs) and network functions virtualization (NFV) (Chiosi *et al* 2017). The OpenRoad project at Stanford University introduced Open-flow, FlowVisor, and SNMPVisor to wireless networks to enhance the control plane. Base station virtualisation from NEC concentrated on slicing radio resources at the medium access control (MAC) layer. CloudEPC from Ericsson modified the Long Term Evolution (LTE) control plane to control open-flow switches. CellSDN from Alcatel-Lucent considered a logically centralised control plane, and scalable distributed enforcement of quality of service (QoS) and firewall policies in the data plane. C-RAN implements a soft and virtualised BS with multiple baseband units (BBUs) integrated as virtual machines on the same server, supporting multiple radio access technologies (RATs). A soft end-to-end solution from the core network to the RAN can enable the 5G goals of spectral and energy efficiency.

In the following sections, this article will elaborate on a *Green* and *Soft* 5G vision. In addition to the traditional emphasis on maximising SE, EE must be positioned side by side for joint optimisation. We present an EE/SE co-design framework. The concept of 'no more cells' is highlighted later with user-centric design and C-RAN as key elements of a soft cell infrastructure. The rationale for a fundamental rethinking of signalling and control design in 5G, especially for the MTC communication is provided. This article further discusses the idea of invisible BSs incorporating large scale antenna system (LSAS) technology. Two major issues are investigated, an N by M hybrid beamforming structure, and an irregular antenna array. Finally, the fundamental interference management issues in networks based on full-duplex technologies and potential solutions are identified followed by a summary.

Rethink Shannon: EE and SE Co-Design

Given limited spectrum and ever-increasing capacity demand, SE has been pursued for decades as the top design priority of all major wireless standards, ranging from cellular networks to local and personal area networks. The cellular data rate has been improved from kilobits per second in 2G to gigabits per second in 4G. SE-oriented designs, however, have overlooked the issues of infrastructure power consumption. Currently, RANs consume 70 percent of the total power. In contrast to the exponential growth of traffic volume on mobile Internet, both the associated revenue growth and the network EE improvement lag by orders of magnitude. A sustainable future wireless network must therefore be not only spectrum-efficient but also energy-efficient. Therefore EE and SE joint optimisation is a critical part of 5G research. Looking at traditional cellular systems, there are many opportunities to become greener, from the equipment level, such as more efficient power amplifiers using envelop tracking, to the network level, such as dynamic operation in line with traffic variations both in time and space. For the fundamental principles of EE and SE co-design, one must first revisit the classic Shannon theory and reformulate it in terms of EE and SE. In classic Shannon theory, the channel capacity is a function of the log of the transmit power (P_t) noise power

spectral density (N_0), and system bandwidth (W). The total system power consumption is a sum of P_t and the circuit power P_c , that is,

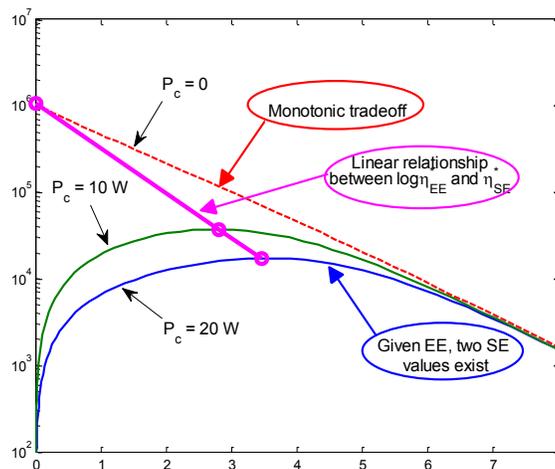
$$P_{tot} = P_t / \rho + P_c$$

where ρ is power amplifier (PA) efficiency defined as the ratio of the input of the PA to the output of the PA. From the definition of EE (Li *et al.* 2011), EE is equal to the channel capacity normalised by the system power consumption. SE is the channel capacity normalised by system bandwidth. The relationship of EE and SE can be shown as a function of PA efficiency and P_c in Fig. 1a. From Fig. 1a, it can be observed that when P_c is zero, there is a monotonic trade-off between η_{EE} and η_{SE} as predicted by the classic Shannon theory. For non-zero P_c , however, η_{EE} increases in the low SE region and decreases in the high SE region with η_{SE} (for a given η_{EE} , there are two values of η_{SE}). As P_c increases, the EE-SE curve appears flatter. Furthermore, when taking the derivative of η_{EE} over η_{SE} , the maximum EE (η_{EE}^*) and its corresponding SE (η_{SE}^*) then satisfy the following:

$$\log_2 \eta_{EE}^* = \frac{\log_2 \rho}{N_0 \ln 2} - \eta_{SE}^*$$

This means there is a linear relationship between $\log_2 \eta_{EE}^*$ and η_{SE}^* , and the EE-SE relationship at the EE optimal points is independent of P_c . This observation implies that as P_c decreases, an exponential EE gain may be obtained at the cost of linear SE loss. Fig 1b compares the EE-SE performance of current Global System for Mobile Communications (GSM) and LTE BSs. LTE performs better than GSM in terms of both SE and EE; both, however, are working in a low SE region, indicating room for improvement.

a) SE and EE relationship for different circuit powers



b) SE and EE relationship for current cellular networks

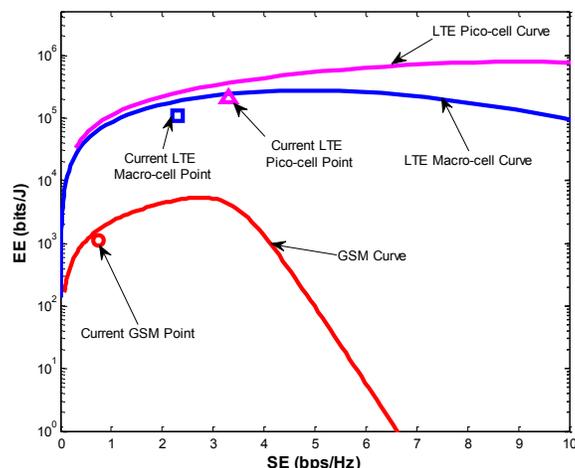


Fig. 1. SE and EE relationship for different circuit powers and current networks

Antenna muting is proposed in EARTH to improve EE, while LSAS stipulates EE improvement by increasing the number of antennas. These seemingly contradicting conclusions are actually consistent with the analysis presented above where the difference is that the former operates in a low SE region, whereas the latter operates in a high SE region. While some progress has been made in EE and SE co-design investigation, there is still a long way to go to develop a unified framework and comprehensive understanding in this area. Especially, new techniques, like spatial modulation (Renzo *et al* 2014), can utilise the antenna resource more efficiently. Ideally, the EE-SE curve in future systems should achieve the following criteria:

- The EE value should be improved for each SE operation point.
- The EE-SE win-win region should be enlarged and the EE-SE trade-off region should be reduced.
- The slope of the EE-SE curve in the tradeoff region should be reduced.

MIMO systems have been studied under these criteria. As shown in Fig.2, besides the linear relationship between $\log_2 \eta_{EE}^*$ and η_{SE}^* , the slopes of the EE-SE curve are also affected by different parameters. We can see that when SE approaches zero, the slope depends only on the bandwidth and circuit power, and is independent of the antenna configuration and the knowledge of CSI at the transmitter. From the solid curves, we can see that reducing the circuit power can improve the EE in the low-SE region significantly. When SE approaches infinity, the slope is only related to the number of data streams, i.e. multiplexing gain, and independent of circuit power. From the dashed curves, it can be observed that when increasing the number of antennas, the EE improvement mainly happens in the high SE region. This observation implies that reducing circuit power has advantage in the low SE region while developing transmission techniques can improve the EE in the high SE region.

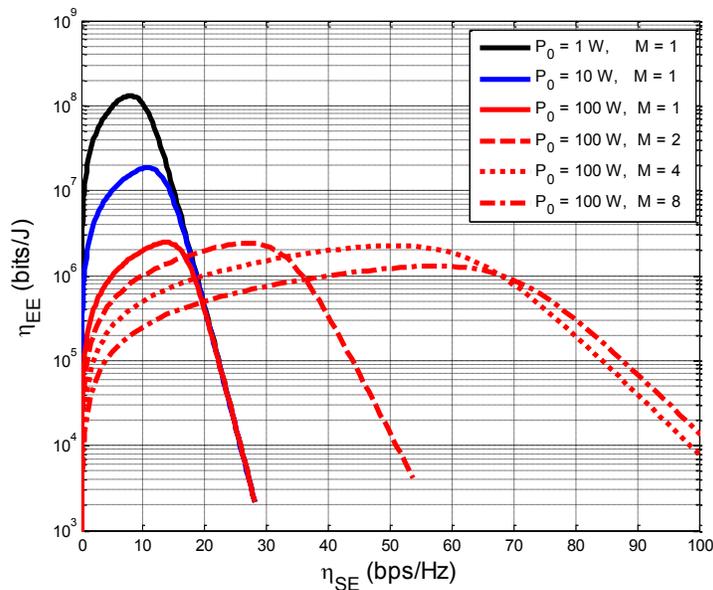


Fig. 2. The impact of the circuit power and the number of antennas on the EE-SE curves.

Rethink Ring and Young: No More Cells

The concept of cellular systems was proposed in 1947 by two researchers from Bell Labs, Douglas H. Ring and W. Rae Young. Since the first generation of cellular standards, this cell-centric design has been maintained through every new generation of standards including 4G. The nature of a homogeneous cell-centric design is that cell planning and optimisation, mobility handling, resource management, signalling and control, coverage, and signal processing are all assumed to be done either for or by each BS uniformly. In a practical deployment, it is clear this system does not match with traffic variations and diverse environments. Relay, coordinated multipoint (CoMP), distributed antenna systems, and heterogeneous networks (HeteNet) have been implemented as short-term solutions to amend these issues. Recently, Beyond Cellular Green Generation, liquid cells, soft cells, and phantom cells have surfaced as potential radio access architectures. These paradigms all lead to the principle of no more cells. 5G design should start with such a paradigm shift, departing from cell-based coverage, resource management, and signal processing, and leaning toward user-centric coverage facilitated by a C-RAN architecture.

User-centric Design

The concept of no more cells is user-centric with amorphous cells, decoupled signalling and data, and decoupled downlink (DL) and uplink (UL). For example, a macro BS would become a signalling BS while small cells would be data-only BSs. In a HetNet scenario, the small cell is within the coverage of a macro cell. Even if the small cell has no traffic, it cannot be turned off in the traditional cell paradigm. But with a control and data decoupling scheme, the macro cell is responsible for control and the small cell only for data. Thus, when there is no data traffic in the small cell, it can be completely turned off to save energy.

New users can access the macro cell, and then the macro cell can coordinate with the small cell for possible data transmission. Based on the channel conditions, service types, and BS traffic loads, the DL and UL can be decoupled to facilitate better resource allocation between cells. This can be illustrated in the following example: Consider two cells where cell 1 is heavily loaded in the downlink and cell 2 loaded in the uplink. In the traditional cell concept, if one user equipment (UE) device is located at the cell boundary with symmetric DL and UL data requirements, and the serving cell is cell 1, its DL requirement may not be satisfied. If the UE device's serving cell is cell 2, its UL requirement may not be satisfied. If there is a user-centric network design, the UE device's DL can be from cell 2 and UL to cell 1, meeting the UE device's data requirement for symmetric DL and UL.

C-RAN

Building on the architecture of distributed BSs where radio units are placed outdoors closer to the antenna and baseband units (BBUs) are placed indoors at cell sites, C-RAN goes one step further by bringing BBUs from multiple BSs to a central pool location. The GPP servers perform baseband processing using virtual machines running on real-time Linux. The centralisation of the baseband processing leads to more energy-efficient cooling, making the C-RAN network architecture an essential part of the design of energy-efficient networks. Energy savings of 70 percent in the OPEX of the BS infrastructure have been realised in 2G and 3G trials inside China. By virtualising the baseband processing, new features can be added to the network within months, as opposed to years in the traditional infrastructure. The centralised baseband processing allows for soft technologies such as CoMP processing, multi-RAT virtualisation, as well as soft and dynamic cell reconfiguration [6]. C-RAN is a revolutionary new type of radio access architecture and another essential element of 5G.

Rethink Signalling and Control: Trillions of Nodes for 5G!?

Existing mobile networks are designed more specifically for conventional and streaming applications such as voice and video. As mobile data traffic grows exponentially, more diversified traffic profiles have emerged. They have brought new challenges to mobile networks, especially small-sized persistent bursty traffic types, such as instant messaging (IM) traffic, that contain frequent texts, photos, and periodic pings. These mobile applications would cause frequent transitions between the Connected and Idle states. As a consequence, these transitions not only increase device battery drainage, but also cause excessive signalling overhead in mobile networks.

Analogous to the small-sized traffic types, the massive connectivity involved in MTC may not be handled efficiently by the current wireless communication networks, like 3G and 4G. The most challenging issue is not that the aggregated MTC data traffic demand is beyond the system capacity but that a large number of devices may access the network simultaneously or an even larger number of small packets might be generated sporadically and sent to the network independently. More importantly, it is estimated that only 7%

of the total wireless traffic will be from IoT nodes in 2020. Therefore, it is well-motivated to design a new signalling mechanism that is optimised to handle such a traffic profile with tens of billions nodes or even trillions of nodes in the IoT networks.

Aggregation of Packet Data from Multiple MTC Devices

There are a number of efforts in this direction. For example, various signalling optimisation schemes have been discussed by Taleb and Kunz (2012) to reduce the signalling load, which is generated during certain procedures and scenarios, like random access etc. Also, considering the fact that many MTC applications are delay-tolerant, low-access priority indication (LAPI) is introduced to alleviate the problem of an overloaded network with MTC devices (3GPP 2011). Jian Zhang et al (2012) introduce a new network convergence approach, where the converged cellular network and wireless sensor network architecture is evolving from layered to flat to decrease the hierarchical signalling exchange between the two networks.

One interesting MTC communication model is proposed in Fig. 3, where the MTC devices are controlled by an aggregator, which functions as the wireless gateway to the cellular network. The services aggregated can be both homogenous and heterogeneous, and sent to the aggregator in a scheduled manner or sporadically. The aggregator will then relay the aggregated data packets to the cellular network in an aggregated manner. Correspondingly, based on the aggregated packet types and data relay modes, the aggregators will request for aggregator specific RRC mechanism. The involved signalling overhead is significantly reduced and the RRC signalling efficiency is dramatically improved. Note that the service aggregation of MTC devices is analogous to the service aggregation of small-sized packets from various applications running on multiple mobile devices.

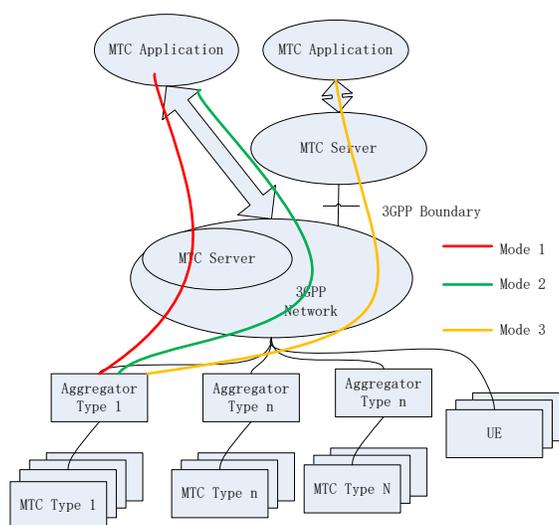


Fig. 3. Aggregation of Packet Data from Multiple MTC Devices

Multiple MTC devices are able to connect to each specific aggregator via local transmission technologies, which could be cable, fibre, WiFi, Bluetooth, ZigBee and even 3GPP cellular access. It is assumed that the connection between MTC devices and aggregators doesn't significantly influence the operator's mobile network and thus, the impact of MTC device's access to aggregators is not considered in this paper. Two relay modes of the aggregators are presented in this paper:

Two proposed relay modes for aggregators

Mode 1: No RRC State Transition (Always Connected)

Suppose in the network (e.g. within one base station), there are N aggregators with same MTC traffic and relay behavior (mode 1) on orthogonal resources, e.g. on different frequency bands (with same bandwidth), and are simultaneously transmitting to the network in a frequency division multiple access (FDMA) manner. Before relaying the aggregated k packets in some time window T , the aggregator is already in connected mode. After packet transmission, the connected mode will be maintained, i.e. there is no need for RRC connection setup in the following windows.

Mode 2: RRC Connection Mode Switched to Idle Mode within Window T

In mode 2, we assume there are N simultaneous aggregators equally sharing the time resource in each window T . The base station allocates all frequency resources to each aggregator, i.e. in a time division multiple access (TDMA) mode. Before transmitting the aggregated packets to the network, each aggregator will enter into connected mode in the time window T . The k packets will be transmitted to the cellular network in window T_{tx} (which is smaller than T) within each window T . Then, after T_{tx} , the aggregator enters into RRC idle mode directly.

For the above two relay modes, the same traffic is relayed in time window T . However, the signalling overhead of these two modes can be different. One analysis can be found in [Chih-Lin *et al* \(2014\)](#).

Invisible Base Stations with Irregular Antenna Array

Massive MIMO or LSAS has been a research focus since Marzetta's seminal paper ([Marzetta 2012](#)). Among the open issues still pending in LSAS, two issues are investigated in this section. Firstly, when a large number of antennas are implemented to achieve better BF gains, implementing the same number of transceivers may not be feasible due to excessive demand on real time signal processing for high BF gains, high power consumption and cost (especially the high cost and power consumption of mixed-signal devices in mm-Wave systems). The hybrid BF structure with much smaller number of digital transceivers than total antenna number will therefore be more practical and cost-effective to deploy. Secondly, although LSAS elegantly addresses the capacity and power consumption challenges, the physical size of an LSAS BS is of concern. The much larger physical footprints of LSAS BSs will not only bring significant tower construction

challenges but also lead to greater confrontation. LSAS BSs with irregular antenna array can help to better adapt to the physical environments, thus easing the antenna installation as shown in Fig. 4.

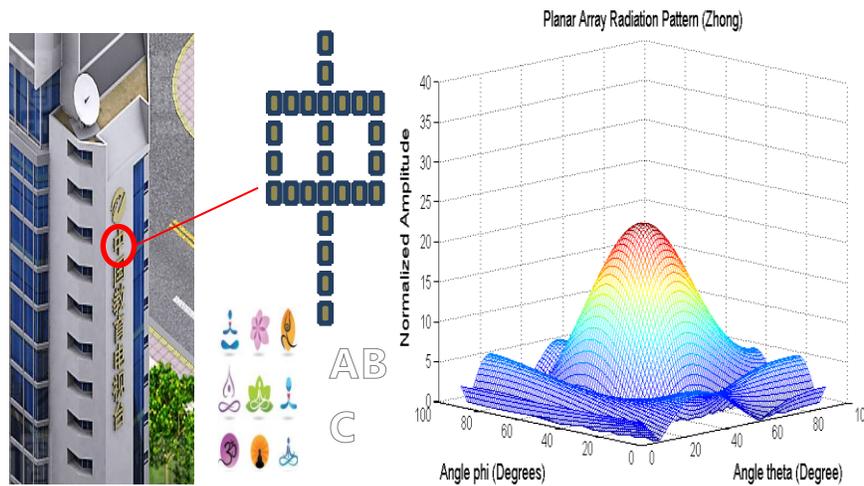


Fig.4. An illustration of invisible base stations

EE-SE Analysis of N by M Active Antenna Structure

Although the generic LSAS requires a complete transceiver chain for each antenna element, for practical considerations, a much smaller number of transceivers than that of antenna elements may be adopted. Unlike current BS RF structures, where each transceiver is connected to a column of antenna elements generating a fixed coverage beam, the LSAS system under investigation is an LSAS system of size $L = N \times M$, where N is the number of transceivers and M is the number of active antennas per transceiver.

An accurate power model is needed to calculate the EE. However, this is not straightforward, since base stations have different types (macro, pico, femto) and are generally produced by different vendors with various implementation technologies. In this section, the following simple power model is used, i.e.,

$$P_{total} = NP + P_{static} = NP + NP_0 + P_{common} + NMP_{rf_circuit},$$

where P_{total} is the total power, NP is the RF power of N transceivers, P_{static} is the static power of the base station, including component NP_0 which scales with the number of transceivers, P_{common} which is common for any transceiver number, and $NMP_{rf_circuit}$ which scales with total antenna number NM . The EE-SE relationship can then be written as:

$$\eta_{EE} = C / P_{total} = \frac{\eta_{SE}}{\left(2^{\frac{r_{SE}}{N}} - 1\right) \frac{N_0}{\eta_{PA}} \frac{N}{M} + \frac{NP_0 + P_{common} + NMP_{rf_circuit}}{W}} \quad (1)$$

EE-SE Relationship at the Green Points

As shown in [Li et al \(2011\)](#), the EE-SE relationship based on Shannon's theory is monotonic, where a higher SE will always lead to a lower EE. When the circuit power is considered, however, there exists a green point on the EE-SE curve where the maximum energy efficiency η_{EE}^* is achieved. Two cases are discussed here for the N by M hybrid BF structure, the case when $NM=L$ (i.e. the total antenna number is fixed to be L , but N and M are variable) and the case when N and M are independent. The analytical goal of the first case is to find the optimal number of transceivers and correspondingly the optimal number of antennas per transceiver given the total number of antennas, while the investigation on the second case is to explore how the independent N and M should be jointly optimised.

It can be derived ([Shuangfeng et al 2014](#)) based on Eq.1 that there exists only one green point on EE-SE curve for each case, i.e., there exists only one η_{SE}^* which maximises the EE performance. The relationship between η_{EE}^* and η_{SE}^* is further given as

$$\lg(\eta_{EE}^*) = -\eta_{SE}^* \lg 2 / N + \lg(M\eta_{PA} / N_0 \ln 2)$$

i.e., $\lg(\eta_{EE}^*)$ scales with η_{SE}^* linearly with a slope of $-\lg 2 / N$. Similar to the EE-SE relationship with classic Shannon theory, a higher η_{SE}^* will always lead to a lower η_{EE}^* . Interestingly, the relationship between η_{EE}^* and η_{SE}^* is independent of P_o , P_{common} , $P_{rf_circuit}$ and W , though, as can be seen from Eq.1, η_{SE}^* and η_{EE}^* are determined based on all the other parameters.

How does N affect EE-SE?

When the required SE is pre-determined, it is desirable that the transceiver number N is optimised, yielding the highest EE performance with the minimum transceiver number. Based on Eq.1, it's found that in the cases $NM=L$ and independent N and M , for any given SE, there exists only one optimal N to yield the best EE. A detailed proof can be found in ([Shuangfeng et al 2014](#)). The practical meaning of the existence of the optimal N is that with a given SE, a system designer doesn't need to implement too many transceivers to achieve the best EE performance.

Assume $P_{rf_circuit} = 1W$, $P_{common} = 50W$, $P_o = 1W$, $W = 2 \times 10^7 \text{Hz}$, $N_0 = 10^{-17} \text{dBm/Hz}$, and a channel gain of -100dB . Considering the $NM=500$ case, the impact of N (from 1 to 10) on EE performance is shown in subplot 1 of Fig. 5, where five SE values are simulated. Note that since N and M are integers, there are only 5 valid (N, M) combinations for $NM=500$, that is, (1,500), (2,250), (4,125), (5,100) and (10,50). It can be observed that on each curve there exists one optimal N that yields the highest EE. For example, when SE is 20bps/Hz, the optimal N is 4. When SE is 8bps/Hz, the optimal N is just 1. The case when N is larger than 10 is not shown in the figure, since it may be difficult to schedule users with negligible inter-user interference when M is very small.

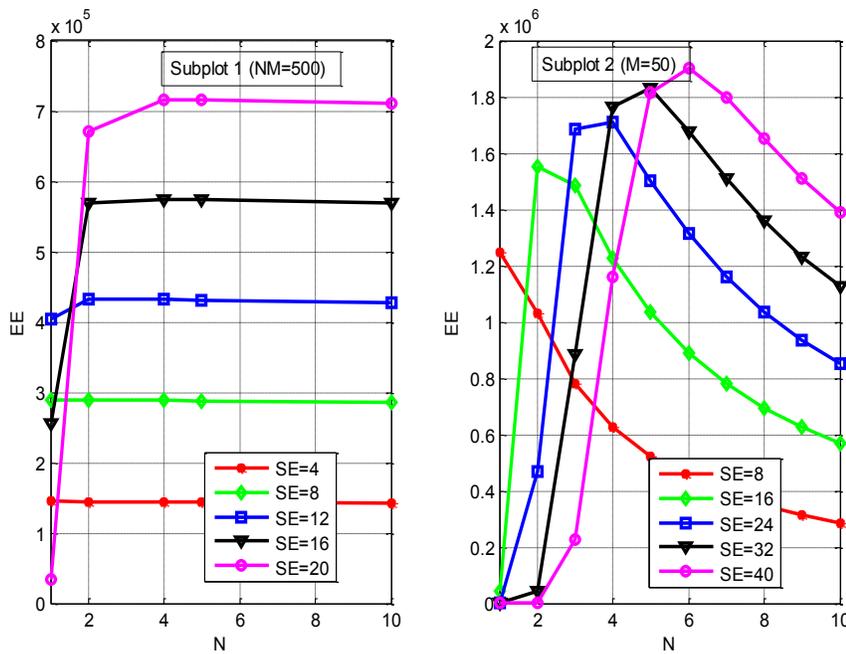


Fig.5. N vs. EE with different SE values

When N and M are independent, the impact of N on EE performance is shown in subplot 2 in Fig.5, where $M=50$ and other parameters are the same with that in subplot 1. Similar to the fixed NM case, on each curve there exists one optimal N . For example, when SE is 40bps/Hz, the optimal N is 6. Different from the fixed NM case, the EE performance is very sensitive to N , because the total antenna number scales with N .

In a practical system operation, the SE requirement may vary according to the traffic load and service types. For example, as shown in subplot 2 in Fig. 5, with the maximum 40bps/Hz SE requirement, the optimum N should be designed to be 6. But when the SE requirement is reduced to 8bps/Hz, the optimal N should be 1. Therefore, it is important that for the possible SE range, the system can be designed with the largest optimal N , and selects the best N according to the SE requirement via transceiver on/off. This can help to further enhance the EE performance according to the system traffic load.

Irregular Antenna Arrays

By integrating the antenna elements into the environment, the BSs can be made virtually invisible. Instead of constructing fake trees, which are often eyesores, multiple active elements can be built in the form of tiles. By separating the single LSAS panel into multiple tiles, the LSAS can be flexibly deployed in an irregular fashion as part of the building facade or signage, and thus blend into the environment. Irregular antenna deployment in a practical environment requires a different system design and adaptive signal processing algorithms. As predicted by sparse antenna array theory, the sidelobes of the sparse array have increased, and due to the smaller number of antennas, the main lobe peak has decreased. Advanced

algorithms regarding subarrays, orthogonal placements, or parasitics can help optimise the beamforming performance of irregular arrays.

In addition to beamforming optimisation, there are several other challenges for irregular antenna arrays. Synchronisation, broadcast, and cell common reference signals in cellular systems are generally transmitted in an omnidirectional manner for better coverage, whereas LSAS panels can only create radiation patterns in front of the panel. Cell coverage will be more challenging since for a given antenna placement, there are many possible coverage scenarios. Another issue is channel modelling for irregular antenna deployment. 3D channel modelling is being investigated in 3GPP and various study groups like WINNER, where generally a regular antenna configuration is assumed. On top of the 2D channel model in WINNER or 3GPP, the elevation angle is added for each ray, where the angle of arrival/angle of departure (AoA/AoD) and the large scale fading with regard to different antennas are assumed to be the same due to the regular spacing in the traditional 2D array. With irregular antenna arrays, however, the spacing and relative position of each antenna may invalidate the above assumption where AoA /AoD and large scale fading may be different for each ray with regard to different LSAS antennas; therefore, modification to the current channel modelling is needed.

Full duplex Radio

Current cellular systems are either frequency division duplex (FDD) or TDD. To double SE as well as improve EE, a full duplex operation should be considered for 5G. A full duplex BS transmits to and receives from different terminals simultaneously using the same frequency resource at the same time. Self-interference cancellation is the key to the success of a full duplex system since high DL interference will make the receivers unable to detect the UL signal. Significant research progress has been made recently in self-interference cancellation technologies, including antenna placement, orthogonal polarisations, analog cancellation, and digital cancellation ([Aryafar 2012](#)). Most of the research, however, has been on either point-to-point relay or a single-cell BS scenario. In this section, FD network deployment scenarios and the interference conditions will be investigated.

FD Network Deployment Scenarios and the Interference Analysis

An FD system can be applied to many practical deployment scenarios. The simplest scenario is the point to point communication (P2P), where the transmission data rate can be doubled. In this case, the self-interference (from Tx to Rx of the same device) is the dominant form of interference. The other scenarios are in cellular networks (homogeneous or heterogeneous) where the base stations and terminals can work in either full duplex or half duplex (HD) modes. With current 3G and 4G networks, new FD BSs may not easily form a homogeneous network with continuous coverage. Rather, FD may serve as a first step for hotspot and indoor capacity enhancement in HetNet with discontinuous coverage, and in the long run may form a new homogeneous network.

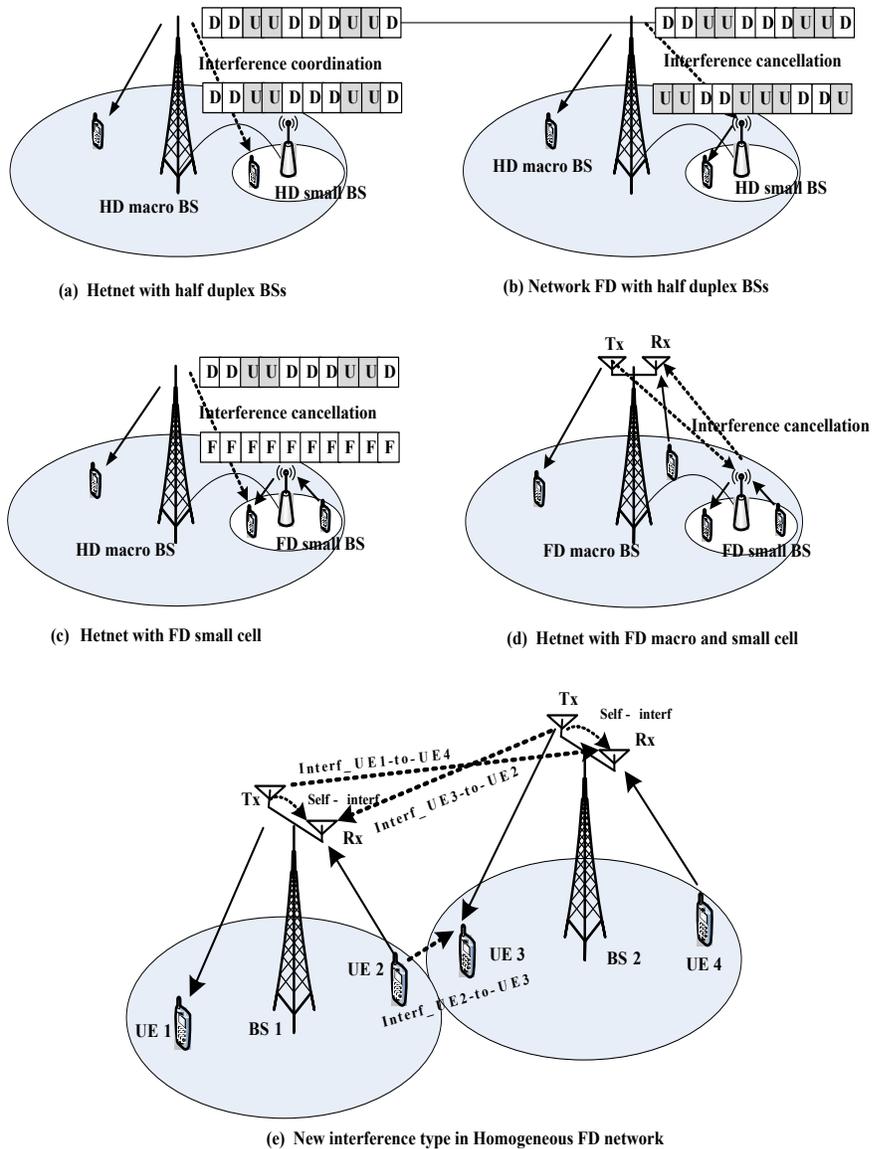


Fig.6. FD Deployment scenario and interference analysis

In the current TDD or FDD systems, the DL to DL interference received at UE and UL to UL interference received at a BS have been extensively studied in literature and standardisation bodies (e.g. CoMP in 3GPP LTE-A and IEEE 802.16m). In a FD system, the interference situation is more severe. For example, the new interference in an FD homogeneous network is shown in Fig. 6(e). (Note the following analysis can be applied to both HetNet and homogeneous networks). BS1 and BS2 are transmitting to UE1 and UE3 respectively in the DL while UE2 and UE4 are transmitting to BS1 and BS2 respectively with same frequency and time resources. In addition to the self-interference from Tx to Rx at each BS, there are intra-cell UL to DL interference $interf_{UE2_to_UE1}$ (from UE2 to UE1) and $interf_{UE4_to_UE3}$ (from UE4 to UE3), inter-cell UL to DL interference $interf_{UE2_to_UE3}$ (from UE2 to UE3), and DL to UL interference $interf_{UE1_to_UE4}$ (from BS1 Tx to BS2 Rx) and $interf_{UE3_to_UE2}$ (from BS2 Tx to BS1 Rx). These interferences have significant impact on whether an FD system works and must be mitigated

properly. Note that DL to DL and UL to UL interference mitigation can readily utilise the CoMP schemes specified in LTE-A, like coordinated scheduling coordinated beamforming and joint processing. More efforts are needed to mitigate the interference for the other 3 types of interference (self interference, UL to DL interference and DL to UL interference). Self interference cancellation has been well studied in the literature ([Ayrafar et al 2012](#)), ([Jain et al 2011](#)), ([Hua et al 2012](#)), ([Knox 2012](#)). Antenna cancellation, analog cancellation, and digital cancellation can be jointly performed to reduce the self interference to a tolerated level. UL(DL) to DL(UL) interference cancellation is still an open issue to solve.

Conclusions

This article has presented five promising areas of research targeting a *Green* and *Soft* 5G system. The fundamental differences between classic Shannon theory and practical systems are first identified and then harmonised into a framework for *EE-SE co-design*. The characteristics of the “no more cells” concept are described from the perspective of infrastructure and architecture variations with particular emphasis on C-RAN as a typical realisation in order to enable various soft technologies. *Rethinking signalling/control* for the MTC communication is then explored, and initial redesign mechanisms are discussed. Virtually *invisible base stations* with irregular LSAS array are envisioned to provide much larger capacity at lower power in high-density areas when integrated into building signage. Optimal configuration of transceivers and active antennas is investigated in terms of EE-SE performance. Finally, deployment scenarios and new interferences in full duplex network are identified, and several candidate solutions are discussed. These five areas provide potential for fundamental breakthroughs, and together with achievements in other research areas, they will lead to a revolutionary new generation of standards suitable for 2020 5G deployment.

References

- Aryafar, E; et al. 2012. “Midu: Enabling MIMO Full Duplex,” *Proc. ACM Mobicom '12*, 2012.
- Chih-Lin I; Shuangfeng Han; Yami Chen; Gang Li. 2014. “Trillions of nodes for 5G!?” submitted to ICC 2014.
- China Mobile Research Institute. 2011. “C-RAN: The Road Towards Green RAN,” Oct. 2011, available at: <http://labs.chinamobile.com/cran>.
- Chiosi, M; Clarke, D; Willis, P. 2012. “Network Functions Virtualization,” Oct. 2012.
- Hua, Y; Liang, P; Ma, Y; Cirik, A. C; Gao, Q. 2012. “A Method for Broadband Full-Duplex MIMO Radio”, *IEEE Signal Process. Lett.*, vol. 19, no. 12, Dec. 2012, pp. 793-796.
- Jain, M; Choi, J; Bharadia, D; Seth, S; Srinivasan, K; Levis, P; Katti, S; Sinha, P. 2011. “Practical, Real-time, Full Duplex Wireless”, *MobiCom'11*, 2011.

- Jian Zhang; Lianhai Shan; Honglin Hu; Yangyang. 2012. Mobile cellular networks and wireless sensor networks: toward convergence, *IEEE Communication Magazine*, vol. 50, issue. 3, pp. 164-169, Mar. 2012
- Knox, M. E; 2012. "Single antenna full duplex communications using a common carrier", *Wireless and Microwave Technology Conference (WAMICON)*, 2012 .
- Li, G. Y. *et al.*, 2011. "Energy-Efficient Wireless Communications: Tutorial, Survey, and Open Issues," *IEEE Wireless Commun.*, vol. 18, no. 6, Dec. 2011, pp. 28–35.
- Marzetta, T. 2012. "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, Nov. 2010, pp. 3590–3600.
- Renzo, M. D; Haas, H; Ghrayeb, A; Sugiura, S; Hanzo, L. 2014. "Spatial modulation for generalized MIMO: challenges, opportunities and implementation," *Proc. IEEE*, vol. 102, no. 1, Jan. 2014, pp. 56–103.
- Shuangfeng Han; Chih-Lin I; Corbett Rowell; Zhikun Xu; Sen Wang; Zhengang Pan. 2014. "Large Scale Antenna System with Hybrid Digital and Analog Beamforming Structure," *International Communication Conference (ICC) workshop 2014*.
- Skillermark, P; Frenger, P. 2012. "Enhancing Energy Efficiency in LTE with Antenna Muting," *IEEE VTC Spring '12*, 2012, May 2012, pp. 1–9.
- Tarik Taleb; Kunz, Andreas. 2012. "Machine Type Communications in 3GPP Networks: Potential, Challenges, and Solutions", *IEEE Communications Magazine*, Mar. 2012, pp. 178-184.
- The Climate Group. 2008. "Smart 2020: Enabling the Low Carbon Economy in the Information Age," available at <http://www.smart2020.org>.
- Third Generation Partnership Project. 2011. 3GPP TS 37.868 v11.0.0. Study on RAN Improvements for Machine-type Communications, Release 11, Sep. 2011

Glossary

BF	Beamforming
BS	Base Station
CoMP	Co-ordinated Multipoint
CSI	Channel State Information
DL	Downlink
EE	Energy Efficiency
FD	Full Duplex
FDD	Frequency Division Duplex
GPP	General Purpose Processors
GSM	Global System for Mobile Communications
HD	Half Duplex
HetNet	Heterogeneous Networks
ICIC	Inter-cell Interference Cancellation
IM	Instant Messaging
LSAS	Large Scale Antenna System
LTE	Long Term Evolution
MAC	Medium Access Control
MIMO	Multiple-Input Multiple-Output
MTC	Machine Type Communications
NFV	Network Functions virtualisation
OPEX	Operating Expenditure
PA	Power Amplifier
RAN	Radio Access Network
RRC	Radio Resource Control
SE	Spectral Efficiency
TDD	Time Division Duplex
UE	User Equipment
UL	Uplink