2

Green Communication Concepts, Energy Metrics and Throughput Efficiency for Wireless Systems

Timothy O'Farrell[1] and Simon Fletcher[2]
[1]*University of Sheffield, Department of Electronic and Electrical Engineering, Sheffield, UK*
[2]*NEC Telecom MODUS Ltd, Surrey, UK*

2.1 Introduction

In the last decade, the communications industry has overcome some extraordinary challenges transforming the way people communicate and access information. Interestingly, this revolution in communication, which touches every aspect of society from the economy, through health care to leisure, has been achieved by an evolutionary approach to enhancing the communications network infrastructure. Over the years, greater access to information has been motivated by the introduction of faster broadband connections (both wired and wireless), versatile and highly portable data centric devices (e.g., smartphones and tablets), and new cloud-based services, which are globally appealing (e.g., social media and YouTube). Importantly, devices such as smartphones have introduced a platform concept that allows third parties to develop and distribute applications (Apps), which has sparked the rapid take-off of mobile broadband. As well as supporting core telecommunication services, network operators, including mobile network operators (MNOs), must provide reliable and inexpensive access to the Internet while the demand for data continues to grow rapidly. Against this backdrop, the industry is already turning its attention to evolving the next generation of broadband infrastructure. For example, the 5G mobile radio access network (RAN) will be characterized by even higher data speeds, prolific amounts of machine-type communication and extremely low latency. Industry is viewing cell densification, heterogeneous networks (HetNets), and massive MIMO technologies as potential approaches to meet this demanding specification. However, this road map comes

Green Communications: Principles, Concepts and Practice, First Edition.
Edited by Konstantinos Samdanis, Peter Rost, Andreas Maeder, Michela Meo and Christos Verikoukis.

with a considerable prices tag in terms of energy and environmental cost as well as capital expenditure costs.

Accommodating the growth of data services by enhancing transmission speeds and increasing the amounts of communication infrastructure presents several key energy consumption challenges which can no longer be discarded because of their impact on network running costs and the environment. Operators are keen to reduce their long-term energy costs [1] given the recent rises in energy prices [2]. From 2009 to 2011, Vodafone reported that the impact of entering new markets together with upgrading existing network infrastructure resulted in a 35% increase in total energy consumption primarily because the number of base station sites owned was doubled during this period [3]. Equally, the challenges imposed by climate change have taken centre stage because scientific studies suggest that global warming directly correlates with Green House Gas (GHG) emissions, in particular CO_2 emissions. As energy consumption and GHG emissions are closely related, a reduction in the former influences strongly the latter. Accordingly, the European Commission has predicated three significant targets for its member states to be fulfilled by 2020; these are 20% reduction in GHG emissions, 20% reduction in total energy consumption, and 20% of energy generation from renewables [4]. This has resulted in companies publically committing to reducing their carbon footprint. For example, by 2020, Vodafone has committed to reduce carbon emissions by 50% from levels in 2008 [5]. Other companies have made similar commitments. While ICT as a whole is responsible for approximately 2% to 3% of global CO_2 emissions [6], it is increasingly evident that ICT is being heavily adopted as a tool to reduce CO_2 emissions in other industries, for example, the electricity supply (green grid) [7], transport, and building management systems, to name a few.

Recently, significant resources have been committed to understanding and remedying the high energy consumption of large-scale communication systems. Since 2008, a number of dedicated international research projects have addressed this issue predominantly focused on cellular mobile systems. For example, the Green Radio (GR) Core 5 Research Program in Mobile VCE [8, 9] has looked at a broad range of energy reduction techniques across the protocol stack and their integration. The GR program focused on intermediate to long-term solutions to achieve energy reductions in mobile networks. Other significant programs that targeted more near-term energy reduction solutions were the Celtic Initiative project OPERANet-I (Optimizing Power Efficiency in Mobile Radio Networks) [10, 11] and the European Framework Program 7 project EARTH (Energy Aware Radio and Network Technologies) [12]. The material presented in this chapter is mainly focused on an analytical framework developed in the GR project that addresses the evolution of an MNO's RAN based on realistic scenarios and traffic data.

The chapter provides an accessible analytical framework that can be used to compare the energy consumption and throughput characteristics of different RAN configurations based on the Long Term Evolution (LTE) family of standards (i.e. LTE and LTE-Advanced). The technical approach is useful for exploring how an LTE RAN may be upgraded in an energy efficient manner to meet the expected growth in mobile data traffic. The approach also allows particular energy saving techniques to be evaluated, such as deployment techniques (e.g., cell size and RAN topology) and power state techniques (e.g., sleepmodes). Drawing on the open literature, the analysis uses accurate energy consumption models of base station and access point equipment in order to account for the overhead, or load independent, energy consumption as well as the load dependent consumption. Therefore, the results reflect the overall operational energy consumption characteristic of a RAN and not just the radio frequency (RF) characteristic. Also,

the Energy Consumption Gain (ERG) and Throughput Gain (TPG) figures of merit, developed by the authors in the GR project, are used to compare RAN configurations. In particular, a new figure of merit called the Energy Throughput Gain (ETG) is identified. The ETG reliably accounts for the difference in energy consumption and throughput between two distinct RAN configurations when transporting application data bits.

The chapter is organized as follows. Section 2.2 provides an overview of the evolution of broadband access networks focusing on radio access. Traffic growth, network upgrading strategies and the role of heterogeneous networks (HetNets) are discussed. Section 2.3 describes the prevailing energy consumption models for base stations and access points and the future trends in the energy consumption of these key subsystems whereas Section 2.4 details the metrics adopted for comparing energy consumption and throughput in distinct RAN configurations. Section 2.5 evaluates the energy consumption and throughput performance of a greenfield deployment scenario based on a regular, homogeneous RAN that uses different cell sizes and base station types. The study is augmented by exploring energy savings obtained from sleep-modes and HetNet deployments. Section 2.6 discusses the key findings drawn from the chapter.

2.2 Broadband Access Evolution

A key factor that has ignited the concern of the energy consumption in broadband access networks, in particular mobile access, is the consistent and rapid growth in broadband traffic. The highly cited Cisco Visual Networking Index (VNI) has reported that in 2012, global average internet traffic increased by 34% while global mobile data traffic grew by 70% [13]. Noticeably, the volume of data traffic consumed per smartphone is increasing rapidly. The Cisco VNI predicts that while portable devices such as smartphones and tablets accounted for 3% of global IP traffic in 2012, this will rise to 22% of IP traffic in 2017. As indicated in Table 2.1, the underlying cause of the exponential growth in mobile data traffic is the combined increase in the number of mobile broadband subscribers, the devices per subscriber, and the traffic consumed per device.

Traffic loads within an RAN can exhibit significant temporal and spatial variations. Though the traffic at any given base station may vary hugely over short time intervals (e.g., seconds or minutes), when averaged over longer time periods (e.g., hours, days or weeks), useful insights emerge that correlate strongly with user behavior. For example, users commuting between the residential and commercial districts of a city results in different busy hour traffic for each district. In commercial areas, the busy hour occurs in the morning while in residential areas, a busy hour is observed between 9 and 10 pm when most traffic is generated indoors by stationary users [14]. This phenomenon is illustrated in Figure 2.1, where normalized traffic intensity versus hour of the day is depicted for a European residential district [15, 16]. Typically, the density of base stations within a city is higher in commercial districts than residential areas. As such, the traffic is unevenly distributed with a significant imbalance between busy and non-busy hours, along with very large variations across the different parts of a city. Despite this temporal and geographic variation of traffic, base stations typically consume a fixed amount of overhead energy commensurate with operating the RAN during the busy hour period. It is widely recognized that this approach to base station deployment and operation is not sustainable from an energy consumption point of view in the face of predicted traffic growth rates.

MNOs have used a number of pathways to evolve their mobile access networks to meet growing traffic demands. Historically, network evolution and equipment upgrades have been

Table 2.1 CAGR for devices and traffic per device

Predicted device and traffic growth	2011–2016 Mbit/month CAGR (%)	2011–2016 Device No. CAGR (%)
M2M module	30	42
Smartphone	77	24
Tablet	52	50
Laptop	27	17
Portable gaming device	27	56

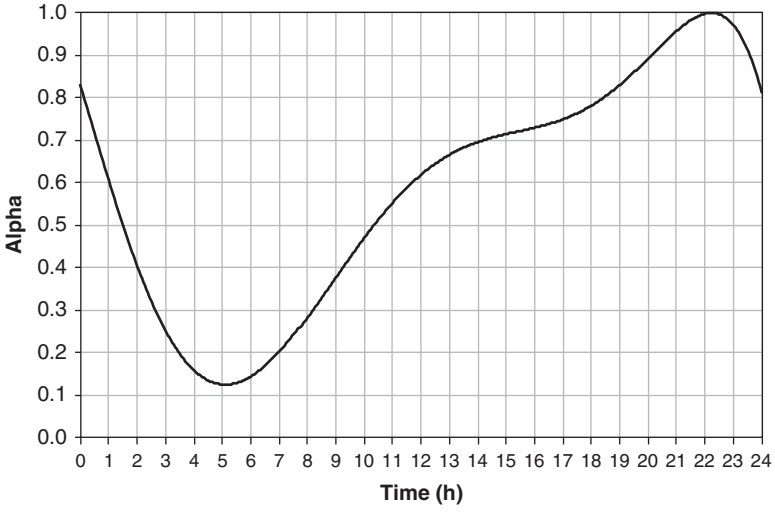

Figure 2.1 Load activity factor α for residential area

driven by capacity and coverage needs without considering network energy consumption constraints. The most significant network evolution pathways have been the introduction of new radio access technologies (RATs) through four successive technology generations. Since the transition in the early 1990s from the 1G analogue mobile systems to 2G digital systems, such as GSM, there has been a constant drive to connect mobile users to the Internet with faster data rates. For this purpose, existing GSM cell sites were upgraded with GPRS and EDGE, and in the early 2000s, a new 3G standard was introduced. Though deployed mostly in densely populated regions, the 3G network in Europe, called UMTS, was deployed with additional cell sites as well as new equipment using newly acquired licensed bands. However, the emergence of handheld devices with enhanced features to access digital content accelerated the upgrade of UMTS to HSPA and HSPA+ with significantly increased downlink peak data rates up to 21 Mbit/s. In 2009, the LTE standard was ratified in 3GPP Release-8 (Rel-8). Commonly referred to as 4G, LTE uses OFDM technology to achieve peak data rates up to 300 Mbit/s if 4×4 MIMO technology is used. Even higher peak data rates will be achieved with LTE-Advance, which is currently being standardized. Recent efforts to reduce the energy consumption of mobile cellular systems, for example, by enabling sleep modes, are planned for future revisions of the LTE specification.

The availability of multiple RATs each with a variety of configurations provides an MNO with the opportunity to deploy a RAN, which is energy-efficient as well as meets capacity and coverage requirements. While the existing 2G and 3G cell sites provide a blueprint on which to build a 4G RAN rapidly, the higher operating frequency of LTE at 2600 MHz together with its higher data rates will necessitate more cell sites in order to achieve seamless coverage, especially in urban areas. Though the lower frequency bands of LTE at 1800 and 800 MHz offer improved coverage, the preferred option to fill the coverage holes between macrocell sites in urban areas is to deploy additional small cells (e.g., micro-, pico-, and metro femtocells), which boost the capacity within a small area. As an example of a HetNet, this layered approach of using macrocells over small cells can also enhance capacity by deploying small cells in traffic hotspots. The traffic served by the small cells releases resources for the macrocell sites. Further, as small cells operate with less transmit RF power, they typically generate less interference while individual units consume less operational energy. For example, small cells using picocell base stations typically operate with RF transmit powers up to 5 W and a 100 m diameter cell. In a HetNet, small-cell deployment outdoors may be realized using an alternative RAT such as WiFi. By operating on a diferent frequency band, WiFi allows macrocells to offload traffic to WiFi access points without increasing the interference level within the macrocell network.

Attempting to serve users located indoors using the outdoor network is problematic owing to the additional high path loss experienced by radio waves passing through walls. The penetration loss can attenuate the radio signal by 10 to 30 dB, causing significant outage. One option to mitigate the effect is to increase the transmit power of outdoor base stations, but this approach is prohibitive in terms of power consumption for the transmission gains achieved. An alternative approach is to deploy femtocells in the form of very low RF transmit power (e.g., 1 mW) base stations operating over very short ranges (e.g., 10 m). By avoiding outdoor-to-indoor path loss, femtocells achieve good indoor coverage to a small group of users. In an enterprise setting, an MNO can plan the deployment of femtocells to meet capacity and coverage needs while in a domestic setting, the operator has no control over the placement of a base station, which can lead to problematic interference conditions. A similar concern exists in domestic WiFi networks where too many collocated WiFi access points induce capacity and latency issues.

Like WiFi access points, each femtocell requires a backhaul connection such as an xDSL connection in homes or an Ethernet connection in offices. In general, the energy consumption of a backhaul connection does not scale linearly with cell size, which results in a high residual backhaul overhead per cell in small-cell deployments. The issue affects outdoor small cells too and is a potential bottleneck as the dense deployment of indoor and outdoor small cells will significantly increase the number of backhaul connections and, therefore, the associated energy consumption.

This highly complex deployment scenario creates a significant challenge for an MNO to plan a RAN, which is efficient in terms of energy, capacity, and coverage. The evolution and upgrade options available to an MNO are subject to tight CAPEX, IMPEX, and OPEX cost constraints. When coupled with the rising cost of energy, there is a pressing need to leverage network planning approaches that minimize energy consumption and maximize capacity and coverage in a trade-off between these three critical design features. It is not enough to rely on Moore's law to reduce energy consumption with equipment upgrades. A proactive network planning methodology that designs for energy efficiency as well as capacity and coverage is needed when faced with exponentially growing traffic volumes. The methodology starts with

a consideration of the energy consumption in network equipment, in particular the base station which is the entity that consumes the most energy in a RAN.

2.3 Cell Site Power Consumption Modeling

The mobile network can be divided into three logical categories: the core, the RAN, and the user devices. Though the number of network elements increases exponentially when enumerating from the core to the edge of the network, the highest proportion of network energy is consumed by the base stations and typically lies between 60% and 80% of the RAN. This fact is significant in RAN evolution because capacity is generally enhanced by adding more base stations. Between 2011 and 2012, Vodafone UK recorded 14,321 operating base stations [17], and this number is set to increase with the introduction of 4G.

Base station sites can vary considerably depending on the shape, size, and composition of the serving cells. However, each base station site has typically seven common key subsystem elements defined as: antennas and feed cables; power amplifier (PA); small-signal RF transceiver, processing unit, backhaul; cooling system; and the mains AC-DC power supply unit that connects to the power grid. This structure can be extended to different categories of base stations including macro, micro, and pico base stations. Auxiliary equipment located at the base station site can include battery backup and lighting.

Feeder cables provide a transmission link between the radio units, typically located at ground level with the base station equipment, and the antennas located at the top of the transmission mask. The cables introduce additional losses, which increase with cable length and RF carrier frequency. Typically, a loss of 3 dB is attributed to feeder cables, which means only half the power from a radio unit reaches the antenna. This relatively high loss can be alleviated by using remote radio head (RRH) technology, which involves locating a sector radio unit in close proximity with the antenna on the mask head. The RRH connects to the system unit, still located at ground level, by fiber but a power supply cable is required to drive the RRH. While a number of RRH configurations are possible, the feeder losses are typically reduced to about 1 dB.

The PA and RF transceiver together constitute the RF module, which can account for approximately 60% of the base station site power consumption. The efficiency of a PA is highly dependent on the peak-to-average power ratio (PAPR) of the signal format being transmitted. In OFDM-based systems like LTE, this can translate into low PA efficiencies of 15% to 35% over a 20 MHz bandwidth. Typically, a dedicated RF module is deployed for each sector and each antenna of a base station site making the RF power overhead a significant portion of the overall site power consumption. By operating a PA close to its saturation region, higher PA efficiencies can be achieved but at the expense of increasing distortion through nonlinear effects. Linearization methods such as Cartesian feedback, feedforward, and digital predistortion [18, 19] alleviate the distortion, allowing PAs to realize maximum efficiencies of 40% to 50%.

The processing unit consists of baseband (BB) signal processing, radio resource management, and site control functionalities. The module also includes digital conversion to and from the RF transceiver. A dedicated system module is usually deployed for each base station sector only. The power consumption of the BB signal processing has continuously benefitted from Moore's law and enhanced processing architecture over time leading to reduced power consumptions for this unit when equipment is upgraded or a new RAT emerges.

From a network planning perspective, the backhaul is an important subsystem element that is frequently omitted from overall RAN energy assessment studies. The backhaul provides a high-capacity connection between the base station site and the core mobile network. This connection may be provided using either fiber optic, microwave, or copper links, depending on the site location. While most sites use a microwave backhaul (48% in the United Kingdom), the proportion of fiber-based backhaul connections is growing. The advent of LTE and small-cell technology to enhance capacity is leading to significantly more base stations being deployed and the impact of this on overall RAN energy consumption needs to be quantified.

Site cooling is an energy-intense facility required at a base station site in order to maintain an appropriate operating temperature inside the base station cabinet. The power consumption of the air-conditioning unit depends on the internal and ambient temperature of the base station cabinet. Typically, an internal and ambient temperature of 25 °C is assumed, which results in a constant power consumption for the air conditioning. Typically, this power consumption can account for 10% to 30% of the overall power consumption of a macrocell base station site. Modern site equipment, in particular RRH technology, is migrating to passive cooling approaches with clear benefits to the overall site energy consumption.

A number of base station site power consumption models are available in the open literature to predict the total power consumption of different types of base station with reasonable accuracy. Significant generic models have been developed in Ref. [20–22] based on measurements from specific vendor equipment. In this chapter, we consider the model developed by the GR project, which is similar to the model in Ref. [20] and is based on manufacturers' data. Figure 2.2 depicts a schematic block diagram of the salient power consuming units of a macrocell base station site using the GR model.

The units divide into three categories: those units used per antenna per sector of the base station site, which includes the PA and the transceiver units; those units used per sector, which

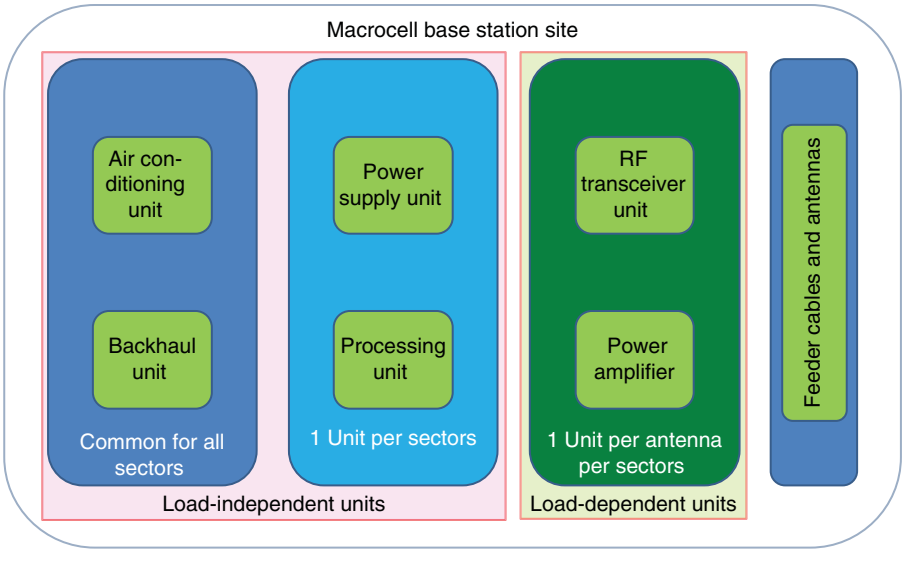

Figure 2.2 Power consuming units of a macrocell base station site

includes the processing and power supply units; and those units used once per base station site, which includes the air-conditioning and backhaul units. A further categorization of the units is based on whether or not there is a dependency on the traffic load. The load-dependent units include the PA and RF transceiver. The power consumption of these components is scaled by a normalized load activity factor α, which ranges from zero to unity. The load-independent units include the power supply, air-conditioning, backhaul, and processing. Though there is some load dependency associated with the backhaul [23] and processing units, its effect is relatively small such that the power consumption of the load-independent units is treated as constant with time. Then, a mathematical expression for the power consumption of a base station site is given by Eq. (2.1).

$$P_{\text{bts}} = P_{\text{ac}} + P_{\text{bh}} + n_{\text{s}} \cdot \{P_{\text{ps}} + P_{\text{pu}}\} + \alpha \cdot n_{\text{s}} \cdot n_{\text{a}} \cdot \left\{ P_{\text{trx}} + \frac{P_{\text{tx}}}{\eta_{\text{pa}} \cdot \eta_{\text{cl}}} \right\} \tag{2.1}$$

The term P_{bts} denotes the overall base station site power consumption. The terms P_{ac}, P_{bh}, P_{ps}, and P_{pu} denote the power consumption of the air-conditioning, backhaul, power supply, and processing units, respectively, which are load-independent, where n_{s} is the number of sectors per base station site. The terms P_{trx} and P_{tx} denote the transceiver and peak PA output power consumption, respectively, which are load-dependent and hence scaled by α. The term n_{a} denotes the number of RF transmission chains, which corresponds to the number of antennas used. This term only scales those units associated with the RF transmission chain. The power consumption of the PA is given by $P_{\text{pa}} = P_{\text{tx}}/\eta_{\text{pa}}$, where η_{pa} denotes the PA efficiency (i.e., output divided by input power at the quiescent operating point). In practice, the PA efficiency has some dependency on α due to the nonlinear characteristic of the PA. The term η_{cl} accounts for the antenna feed cable efficiency. With an appropriate choice of parameters, Eq. (2.1) can be used to model macro-, micro-, pico-, and femtocell base station power consumption with reasonable accuracy as well as account for MIMO antenna configurations. Table 2.2 shows a list of parameter values appropriate for different base station types and their power consumption at full load (i.e., α). These parameters have been extracted from Ref. [20, 21] and combined with data collected in the GR project [24]. An RRH configuration is assumed for the microcell. The estimates in Table 2.2 demonstrate that the significant power savings can be achieved when using a microcell when compared to a macrocell. The combination of reducing the transmit power and the air-conditioning power while improving the feeder cable efficiency reduces the overall base station power consumption by about nine times. Both pico- and femtocells show further significant reductions in overall power consumption. However, these power reductions are achieved at the expense of cell site coverage as discussed later in the chapter. In addition, the proportion of power consumption attributed to backhaul increases from about 0.5% in a macrocell to 48% in a femtocell, when the backhaul is based on a fiber optic link.

2.4 Power and Energy Metrics

Metrics provide a means of quantifying the consumption of power or energy in a mobile cellular network. Typically, power and energy metrics are used to compare and contrast different power/energy saving solutions. However, a metric's effectiveness depends significantly on the network conditions when a measurement is made. For example, when comparing the energy consumption of two different RAN configurations or architectures, there is an implicit

Table 2.2 LTE BS site power consumption estimates based on 2010 vendor data

Equipment	Macrocell	Microcell	Picocell	Femtocell
Number of sectors (n_s)	1	1	1	1
Number of antennas per sector (n_a)	1	1	1	1
Transmit power (W)	40	6.3	0.13	0.1
Transceiver unit (W)	13	6.5	1.0	0.6
Processing unit (W)	30	27	3	2.5
Power supply unit (W)	50	50	5	2
Backhaul (fiber) (W)	10	10	10	10
Air-conditioning unit (W)	225	60	0	0
PA efficiency (η_{pa})	30%	23%	7%	4%
Antenna feed cable loss (dB)	−3	−1	0	0
BS site power (W)	594	188	21	18

requirement that both networks are compared when transporting the same absolute number of application data bits. This premise suggests that energy per bit, with units of J/bit, provides a useful energy metric. The method developed in the GR project to compare the energy consumption of two distinct RANs builds on the notion of a J/bit metric or Energy Consumption Rating (ECR).

The energy metric framework is developed by considering a RAN, covering an area A_{ran}, which must transport in a predetermined period of time a total of M application bits, which are evenly distributed over A_{ran}. The RAN consists of n homogeneous base station sites (BTS) each of approximate circular radius $D/2$, where D is the intersite distance. From Eq. (2.1), the energy consumption of a BTS, E_{BTS}, is given by Eq. (2.2).

$$E_{BTS} = P_{rh}T_{rh} + P_{oh}T_{oh} \tag{2.2}$$

The terms in P_{rh} and P_{oh} denote the load-dependent and load-independent peak or maximum power consumptions, respectively, where rh refers to a radio-head-only consumption and oh refers to an overhead-only consumption. From Eq. (2.1) these terms are given by $P_{rh} = n_s \cdot n_a \cdot \left\{ P_{trx} + \frac{P_{tx}}{\eta_{pa} \cdot \eta_{cl}} \right\}$ and $P_{oh} = P_{ac} + P_{bh} + n_s \cdot \{P_{ps} + P_{pu}\}$. The terms T_{rh} and T_{oh} denote time durations over which load-dependent and independent power is consumed, respectively. T_{rh} is defined in terms of a site-average capacity $T_{rh} \overset{\text{def}}{=} \frac{M/n}{C_{BTS}}$, while T_{oh} is defined in terms of a site-average throughput, $T_{oh} \overset{\text{def}}{=} \frac{M/n}{S_{BTS}}$, both measured in bit/s, and $T_{rh} \leq T_{oh}$. However, $S_{BTS} = \alpha C_{BTS}$ such that the site-average load activity factor $\alpha = T_{rh}/T_{oh} \leq 1$ and $S_{BTS} \leq C_{BTS}$. Then, Eq. (2.2) may be written in terms of C_{BTS}, α, M and n as shown in Eq. (2.3).

$$E_{BTS} = P_{rh} \cdot \frac{M}{n \cdot C_{BTS}} + P_{oh} \cdot \frac{M}{\alpha \cdot n \cdot C_{BTS}} \tag{2.3}$$

The number of base station sites n can be approximated by $n = A_{RAN}/A_{BTS}$, where the site coverage area $A_{BTS} \cong \pi\left(\frac{D}{2}\right)^2$. Then, an expression for the total RAN energy consumption in

terms of the site parameters is given by Eq. (2.4).

$$E_{\mathrm{R}} = n \cdot E_{\mathrm{BTS}} = \left(\frac{P_{\mathrm{rh}}}{C_{\mathrm{BTS}}} + \frac{P_{\mathrm{oh}}}{\alpha \cdot C_{\mathrm{BTS}}} \right) \cdot M \tag{2.4}$$

Dividing E_{R} by M in Eq. (2.4) gives an expression for the RAN energy consumption rating $\mathrm{ECR_R}$, as shown in Eq. (2.5).

$$\mathrm{ECR_R} = \frac{E_{\mathrm{R}}}{M} = \frac{\alpha \cdot P_{\mathrm{rh}} + P_{\mathrm{oh}}}{\alpha \cdot C_{\mathrm{BTS}}} = \frac{P_{\mathrm{BTS}}}{S_{\mathrm{BTS}}} = \mathrm{ECR_{BTS}} \tag{2.5}$$

Equation (2.5) demonstrates the use of the ECR metric when determining the absolute RAN energy consumption because $E_{\mathrm{R}} = M \cdot \mathrm{ECR_R}$. Importantly, both the RAN and BTS ECR are equal for a homogeneous deployment of sites and are given by the ratio of site-average power consumption (P_{BTS}) to site-average throughput (S_{BTS}). Equation (2.5) is formulated on the parameters of a base station site. Without loss of generality, a formulation based on the parameters of a sector (or cell) is straightforward because the total number of cells is equal to $n \times n_{\mathrm{s}}$.

Equation (2.5) can be used as a comparison framework between two or more distinct RAN architectures. Let $E_{\mathrm{R},1}$ and $E_{\mathrm{R},2}$ denote the energy consumption of two distinct RANs ($\mathrm{RAN_1}$ and $\mathrm{RAN_2}$), where $E_{\mathrm{R},1} \geq E_{\mathrm{R},2}$. A figure of merit termed the RAN Energy Consumption Gain is defined as $\mathrm{ECG_R} = \frac{E_{\mathrm{R},1}}{E_{\mathrm{R},2}}$ on the basis that $\mathrm{RAN_2}$ consumes less energy than $\mathrm{RAN_1}$ owing to more energy-efficient base station sites. An expression for the $\mathrm{ECG_R}$ in terms of the base station site parameters is given by Eq. (2.6), where $\mathrm{RAN_1}$ and $\mathrm{RAN_2}$ transport M_1 and M_2 application bits, respectively.

$$\mathrm{ECG_R} = \frac{M_1 \cdot (\alpha_1 \cdot P_{\mathrm{rh},1} + P_{\mathrm{oh},1})/S_1}{M_2 \cdot (\alpha_2 \cdot P_{\mathrm{rh},2} + P_{\mathrm{oh},2})/S_2} = \frac{M_1}{M_2} \cdot \frac{P_1/S_1}{P_2/S_2} = \frac{M_1}{M_2} \cdot \frac{\mathrm{ECR_1}}{\mathrm{ECR_2}} \tag{2.6}$$

In Eq. (2.6), the ratio M_2/M_1 constitutes a second RAN figure of merit termed the RAN Throughput Gain (TPG) denoted by $\mathrm{TPG_R}$. For clarity, in Eq. (2.6), the subscript BTS denoting base station site has been omitted.

Equation (2.6) forms a basis for comparing the energy efficiency of distinct RANs across a broad range of operating configurations. This is achieved by applying the constraints $T_{\mathrm{oh},2} = T_{\mathrm{oh},1}$ and $T_{\mathrm{rh},2} \leq T_{\mathrm{rh},1}$. From Eq. (2.3), the first constraint gives $\frac{M_2}{\alpha_2 n_2 C_2} = \frac{M_1}{\alpha_1 n_1 C_1}$ while the second constraint gives $C_2 \geq C_1$, where n_i for $i = 1, 2 \ldots$ denotes the number of sites per RAN. As $n_1/n_2 \equiv A_2/A_1$, where A_1 and A_2 are the base station site areas and $A_1 \geq A_2$, then the constraint $\frac{M_2}{\alpha_2 n_2 C_2} = \frac{M_1}{\alpha_1 n_1 C_1}$ gives an expression for the RAN TPG as $\mathrm{TPG_R} = \frac{M_2}{M_1} = \frac{S_2/A_2}{S_1/A_1}$. That is, the RAN Throughput Gain is equal to the ratio of the base station site area throughput densities (bit/s/km^2) between $\mathrm{RAN_2}$ and $\mathrm{RAN_1}$. Substituting $\frac{M_2}{M_1} = \frac{S_2/A_2}{S_1/A_1}$ into Eq. (2.6) gives Eq. (2.7).

$$\mathrm{ECG_R} = \frac{1}{\mathrm{TPG_R}} \cdot \frac{\mathrm{ECR_1}}{\mathrm{ECR_2}} = \frac{P_1/A_1}{P_2/A_2} \tag{2.7}$$

Equation (2.7) illustrates a key relationship between the RAN ECG and the base station site area power densities (W/km^2). However, Eq. (2.7) also demonstrates a fundamental relationship between the product $\mathrm{ECG_R} \cdot \mathrm{TPG_R}$, shown in Eq. (2.8), and the ratio of the base station site Energy Consumption Ratings (J/bit).

$$\mathrm{ECG_R} \cdot \mathrm{TPG_R} = \frac{\mathrm{ECR_1}}{\mathrm{ECR_2}} \tag{2.8}$$

The product $\mathrm{ECG_R} \cdot \mathrm{TPG_R}$ forms an overall figure of merit when comparing two RANs, which takes into account the RAN energy efficiency as measured by the ECG and the RAN throughput efficiency as measured by the TPG. Importantly, this composite figure of merit is determined by the site ECR metrics. Equation (2.8) clearly demonstrates the inverse relationship between energy efficiency and throughput efficiency where by increasing one reduces the other while their product is bound by the ratio $\mathrm{ECR}_1/\mathrm{ECR}_2$. Then, the most efficient RAN in terms of least energy consumed and greatest throughput achieved is the one that maximizes the product $\mathrm{ECG_R} \cdot \mathrm{TPG_R}$. In summary, the three defining figures of merit for comparing RANs in terms of the site metrics P/A (W/km^2), S/A (bit/s/km^2), and P/S (J/bit) are shown in Eq. (2.9).

$$\mathrm{ECG_R} = \frac{P_1/A_1}{P_2/A_2}, \quad \mathrm{TPG_R} = \frac{S_2/A_2}{S_1/A_1}, \quad \mathrm{ECG_R} \cdot \mathrm{TPG_R} = \frac{P_1/S_1}{P_2/S_2} \tag{2.9}$$

The energy and throughput efficiencies of various LTE RAN configurations are explored in Section 2.5. However, the formulas in Eq. (2.9) provide a number of key insights into the bounds on RAN ECG and TPG. For example, consider a homogeneous RAN that is signal-to-interference (SIR) limited with load activity factor $\alpha_1 = 1$, intersite distance D_1, and base station site power consumption P_1. If the intersite distance is reduced from D_1 to D_2 using the same base station (i.e., $P_2 = P_1$) with $\alpha_2 = 1$, then $S_2 \cong S_1$ giving $\mathrm{ECG_R} \cdot \mathrm{TPG_R} = 1$. This scenario corresponds to cell densification using the same BTS infrastructure. Equation (2.9) proves that the RAN ECG decreases as D_2 decreases because $\mathrm{ECG_R} = 1/\mathrm{TPG_R} = A_2/A_1 = D_2^2/D_1^2$ while the RAN TPG increases as D_1^2/D_2^2. However, if a lower power base station is used instead (i.e., $P_2 < P_1$), then the expressions in Eq. (2.9) can be used to determine a least intersite distance $D_2 = D_1\sqrt{P_2/P_1}$ with $\mathrm{ECG_R} = 1$ and $\mathrm{TPG_R} = P_1/P_2$. By increasing D_2, it is possible to increase $\mathrm{ECG_R}$ while reducing $\mathrm{TPG_R}$ subject to the constraint $\mathrm{ECG_R} \cdot \mathrm{TPG_R} = P_1/P_2$. This bound demonstrates that joint energy savings and throughput gains are attainable at reduced intersite distances provided the new base station has lower power. In the next section, a variety of energy saving strategies for an LTE-based RAN are explored building on the metric analysis of this section.

2.5 Energy and Throughput Efficiency in LTE Radio Access Networks

In this section, the ECG and TPG figures of merit developed in Section 2.4 are used to investigate an LTE RAN accounting for intercell interference as well as noise. The analysis is based on a Shannon capacity model of the LTE physical layer. The interference is characterized by an average value related to the load activity factor α. The analysis is carried out for a homogeneous deployment of single sector or cell base station sites using antennas characterized by unity gain in the azimuth plane (i.e., omnidirectional) and gain G in the elevation plane. Each base station transmits with the same RF power P_{tx}. This approach provides vivid insights into the planning of RANs based on energy as well as throughput requirements [25].

A key performance metric in RANs is the average signal-to-interference-and-noise ratio (SINR) defined as $\mathrm{SINR} = \frac{S}{I+N}$, where S is the average received signal power, I is the average received interference, and N is the noise power. The total interference from K neighboring sites is given by Eq. (2.10).

$$I = \sum_{k=1}^{K} \alpha_k I_{\max,k} \tag{2.10}$$

The value of the *SINR* depends on the position of the mobile within a cell. Often, the cell-edge SINR is used as a determining quality of service (QOS) metric. Another important QoS metric is the cell-average SINR obtained by averaging the SINR over the cell area. The analysis assumes that the load activity factors α_k in every cell (i.e., serving and interfering) are equal such that $\alpha_k = \alpha\ \forall k$. Then, $I = \alpha I_{\max}$ where $I_{\max} = \sum_{k=1}^{K} I_k$ is the maximum interference power at the mobile. The SINR can be written as a function of the minimum signal-to-interference ratio (SIR), the signal-to-noise ratio (SNR), and the cell load activity factor α as shown in Eq. (2.11) where $\text{SIR}_{\min} = \frac{S}{I_{\max}}$ and $\text{SNR} = \frac{S}{N}$.

$$\text{SINR} = \frac{1}{\frac{\alpha}{\text{SIR}_{\min}} + \frac{1}{\text{SNR}}} \tag{2.11}$$

The SIR and SNR are averaged values taken over the small-scale fading of S and I. The magnitude of $\text{SIR}_{\min}$ at the mobile is determined by the network geometry as well as the antenna configuration. It is common practice to model the interference as a Gaussian noise. Therefore, the SNR is given by $\frac{S}{N} = \frac{\text{EIRP}}{\text{LN}}$, where the effective isotropic radiated power $\text{EIRP} = GP_{tx}$ for transmit antenna gain G. The term in L is a linear path loss given by $L = MKd^{\gamma}$, where M denotes a shadow fading margin, γ is the path loss exponent, K is the median path loss at a distance of one kilometer, and d is the distance in kilometers between the mobile and base station. The dB path loss is modeled by $L_{dB} = 128 + 37.4log_{10}(d_{km})$, which is based on the 3GPP system simulation specification [25]. A dB shadow-fading margin of 8.7 dB is used throughout the analysis, which corresponds to a 95% single-cell area probability for a shadow-fading standard deviation of 8 dB. The Shannon formula is used to provide a mapping between capacity C and SINR as shown in Eq. (2.12) where the terms in EIRP and L have been substituted.

$$C = \text{B}\log_2\left(1 + \frac{1}{\frac{\alpha}{\text{SIR}_{\min}} + \frac{MKd^{\gamma}N}{GP_{tx}}}\right) \tag{2.12}$$

Equation (2.12) gives the capacity or peak rate at the mobile's position for a system bandwidth B. The site- or cell-average capacity is obtained by taking the average of C over the cell area (i.e., $C_{\text{BTS}} = \frac{4}{\pi D^2}\int_{\text{cell area}} C\, dA$). This calculation is usually completed numerically or by system level modeling of the RAN. Then, an expression for the cell-average spectral efficiency is given by $\eta_{\text{BTS}} = \frac{C_{\text{BTS}}}{B}$ and the cell-average throughput is given by $S_{\text{BTS}} = \alpha C_{\text{BTS}}$.

The average values for the cell capacity and throughput are combined with the RAN metrics and figures of merit developed in Section 2.4 to evaluate a number of options for saving energy in a RAN. In the following sequel, energy savings obtained by reducing the cell size, deploying small-cell technologies, and using sleep modes are evaluated. Also, the energy efficiency of various HetNet topologies is investigated. The parameters for these investigations are shown in Table 2.3 with power consumption values being extracted from the data presented in Table 2.2

Table 2.3 BTS technology parameters

Cell	Macro	Micro	Pico	Femto	Units
P_{tx}	46	38	21	20	dBm
P_{oh}	315	147	18	14.5	W
P_{rh}	279	41	2.86	3.1	W
G	18	6	2	0	dBi

and applied to Eq. (2.1) for a base station site with one antenna unit ($n_{\text{a}} = 1$) and, therefore, one sector or cell ($n_{\text{s}} = 1$).

2.5.1 Reducing Cell Size

In this subsection, the impact on energy consumption and throughput is evaluated when the size of a cell is reduced without changing the base station site hardware or system bandwidth. This scenario is equivalent to reducing the intersite distance or increasing the number of base station sites per km^2. The impact of reducing D on ECG_{R}, TPG_{R} and their product ETG_{R} is considered. The RAN with the greater cell size is identified as RAN_1 and, without loss of generality, the load activity factor in RAN_1 is set equal to unity (i.e., $\alpha_1 = 1$). From Eq. (2.9), expressions for the ECG_{R} and TPG_{R} in terms of the intersite distances and cell-average spectral efficiencies are given by Eq. (2.13) where the cell area ratio $\frac{A_2}{A_1} = \left(\frac{D_2}{D_1}\right)^2$.

$$\text{ECG}_{\text{R}} = \left(\frac{D_2}{D_1}\right)^2 \frac{P_{\text{rh},1} + P_{\text{oh},1}}{\alpha_2 P_{\text{rh},2} + P_{\text{oh},2}} \qquad \text{TPG}_{\text{R}} = \left(\frac{D_1}{D_2}\right)^2 \frac{\alpha_2 \eta_2}{\eta_1} \tag{2.13}$$

Figure 2.3(a), (b), and (c) show plots of ECG_{R}, TPG_{R}, and ETG_{R} as a function of the ratio $\frac{D_2}{D_1}$ with α_2 as a parameter, respectively. The results were obtained for the macrocell parameters shown in Tables 2.2 and 2.3 with a value of $D_1 = 1$ km, which gives a cell-average spectral efficiency $\eta_1 = 3.255$ bit/s/Hz. The results demonstrate a number of important conclusions based on the common practice of operators to increase system capacity by increasing the number of base station sites per km^2.

From Figure 2.3(a), the ECG_{R} is always < 1 for $\alpha_2 < 1$ and as D_2 decreases with respect to D_1, the ERG_{R} decreases rapidly with $\text{ECG}_{\text{R}} \ll 1$ for $\frac{D_2}{D_1} < 0.5$. When $\alpha_2 = 1$, the ECG_{R} increases marginally as D_2 decreases because η_2 increases marginally as D_2 decreases. Figure 2.3(a) also illustrates the impact of cell load on energy efficiency. The ECG_{R} increases with decreasing load. That is, a lightly loaded RAN consumes less energy than a heavily loaded RAN. However, the difference in ECG_{R} is relatively small as the load α only affects the radio-head power consumption (P_{rh}), which is typically small compared to the overhead power consumption (P_{oh}). When $T_{\text{rh}} < T_{\text{oh}}$, an opportunity exists to exploit sleep mode principles such as discontinuous transmission (DTX).

In contrast, Figure 2.3(b) shows that the TPG_{R} exhibits the opposite characteristics to the ECG_{R}. That is, the TPG_{R} increases with decreasing D_2 and is lower when the network is more

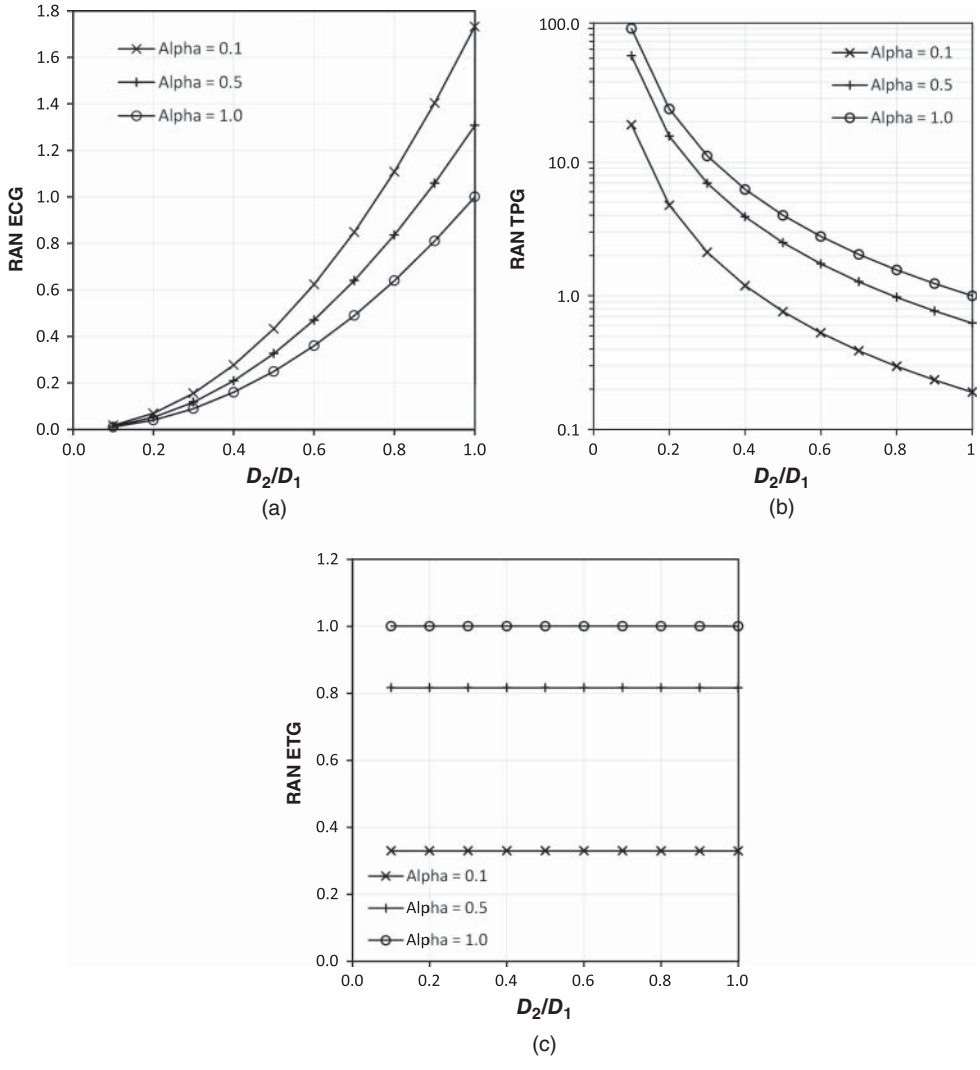

Figure 2.3 Energy and throughput figures of merit for a macrocell RAN. (a) ECG_R, (b) TPG_R, (c) ETG_R

lightly loaded. For values of $D2/D1 < 0.5$, the TPG increases rapidly. A point of interest is when $ECG_R = TPG_R$. This equality indicates that the performance of RAN_2 has achieved the largest possible joint energy consumption and throughput gains with respect to RAN_1. In particular, when $ECG_R = TPG_R = 1$, the joint gain $ETG_R = 1$ corresponds to the case that RAN_2 achieves the same energy and throughput performance as RAN_1 at intersite distance D_2. From Figure 2.2(a) and (b), the only point when $ECG_R = TPG_R = 1$ is when $D_2 = D_1$. This result demonstrates that increasing the RAN throughput by reducing the intersite distance only will always result in an increased RAN energy consumption. Figure 2.3(c) plots $ETG_R = ECG_R \cdot TPG_R$ versus D_2 with α_2 as a parameter. The graphs show that ETG_R remains constant

for $0.1 \leq \frac{D_2}{D_1} \leq 1$ but increases with α_2 to a maximum value of unity. The constant characteristic of $\mathrm{ETG_R}$ can be attributed to the RAN performances being SIR limited.

2.5.2 Reducing Cell Size and BTS Power Consumption

The results in (A) showed that in order to increase throughput by reducing intersite distance, the power consumption of the BTS should be reduced. In this section, the impact of replacing macrocells with microcells using the BTS data in Tables 2.2 and 2.3 is examined. Under full load, the macrocell consumes 594 W while the microcell consumes 184 W. Using the ECG, TPG, and ETG metrics the least intersite distance and the least cell bandwidth are defined and evaluated when $\mathrm{ECG_R} = \mathrm{TPG_R} = 1$ (i.e., $\mathrm{ETG_R} = 1$). That is, $\mathrm{ECG_R} = \mathrm{TPG_R} = 1$ corresponds to the condition that $\mathrm{RAN_2}$ using microcells achieves the same throughput and energy consumption as $\mathrm{RAN_1}$ using macrocells.

Figure 2.4(a), (b), and (c) show plots of $\mathrm{ECG_R}$, $\mathrm{TPG_R}$, and $\mathrm{ETG_R}$ as a function of the ratio $\frac{D_2}{D_1}$ with α_2 as a parameter, respectively. As in (A), the results were obtained for a value of $D_1 = 1$ km corresponding to a cell-average spectral efficiency $\eta_1 = 3.255$ bit/s/Hz in $\mathrm{RAN_1}$. The results show similar trends to those obtained in Figure 2.3. However, as $P_2 \ll P_1$, there is an appreciable range of α_2 values when $\mathrm{ETG_R} > 1$ for $0.1 \leq \frac{D_2}{D_1} \leq 1$. That is, the reduced power consumption of a microcell can be exploited to enhance the throughput of $\mathrm{RAN_2}$ without consuming more energy compared to $\mathrm{RAN_1}$ which uses macrocells.

In order to quantify these gains, the least intersite distance when $\mathrm{ECG_R} = \mathrm{TPG_R} = 1$ can be determined from Eq. (2.13) as shown in Eq. (2.14) when the system bandwidths are the same (i.e., $B_2 = B_1$).

$$\frac{D_2}{D_1} = \sqrt{\frac{\alpha_2 P_{\mathrm{rh},2} + P_{\mathrm{oh},2}}{P_{\mathrm{rh},1} + P_{\mathrm{oh},1}}} = \sqrt{\frac{\alpha_2 \eta_2}{\eta_1}} \tag{2.14}$$

From analysis, an exact solution to Eq. (2.14) is obtained when $\alpha_2 = 0.15$ and $\eta_2 = 5.595$ bit/s/Hz giving $\frac{D_2}{D_1} = 0.508$ or $D_2 = 508$ m. Any value of $D_2 < 508$ m will result in $\mathrm{RAN_2}$ consuming more energy than $\mathrm{RAN_1}$ though achieving a higher RAN throughput. However, the energy consumption gains in $\mathrm{RAN_2}$ are achieved by lowering α_2, which reduces the intercell interference, in order to increase the cell-average spectral efficiency η_2. This approach has the undesired effect of operating each cell of $\mathrm{RAN_2}$ with a very low throughput efficiency ($\frac{\alpha\eta}{B}$), which significantly under utilizes the available radio resource. For example, the cell-average throughput efficiencies in $\mathrm{RAN_1}$ and $\mathrm{RAN_2}$ are 3.255 bit/s/Hz and 0.839 bit/s/Hz, respectively, which correspond to absolute cell-average throughputs of 65 Mbit/s and 16.7 Mbit/s, respectively. This ratio is required so that $\mathrm{RAN_2}$, which has almost four times as many cells as $\mathrm{RAN_1}$, achieves the same RAN throughput and energy consumption when the system bandwidths are the same (i.e., $B_2 = B_1 = 20$ MHz).

In order to avoid operating $\mathrm{RAN_2}$ with a very low throughput efficiency, its system bandwidth B_2 can be reduced instead. The amount of this reduction can be calculated from the metric expressions in Eq. (2.10) for the case when $\alpha_2 = \alpha_1 = 1$ and $\eta_2 \cong \eta_1$. These constraints correspond to both RANs being fully loaded and SIR limited such that the cells in both RANs have approximately the same throughput and spectral efficiencies. Equation (2.15) shows expressions for $\mathrm{ECG_R}$ and $\mathrm{TGP_R}$ in terms of the BTS powers, bandwidths, and intersite distance ratio

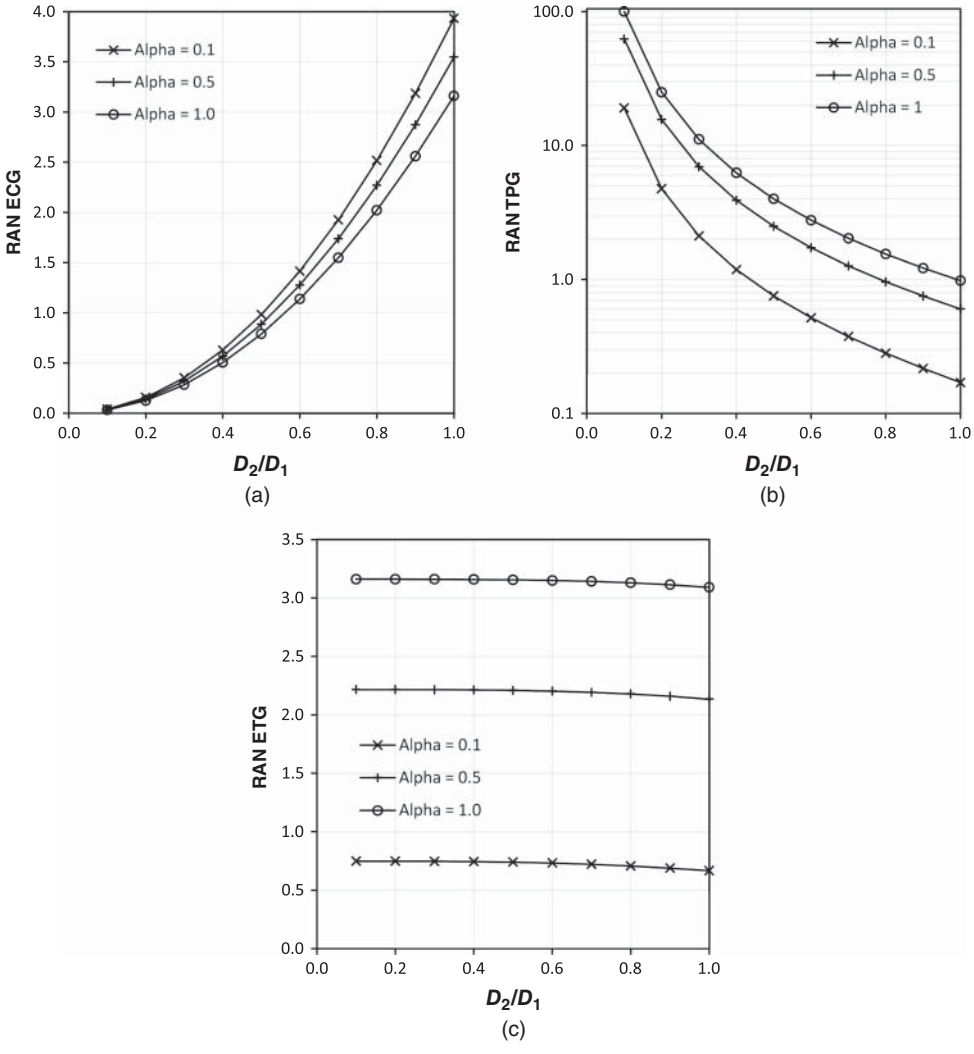

Figure 2.4 Energy and throughput figures of merit for a small-cell RAN. (a) $\mathrm{ECG_R}$, (b) $\mathrm{TPG_R}$, (c) $\mathrm{ETG_R}$

where the cell-average throughputs are given by $S_i = \alpha_i B_i \eta_i, i = 1, 2$.

$$\mathrm{ECG_R} = \left(\frac{D_2}{D_1}\right)^2 \cdot \frac{P_1}{P_2} \qquad \mathrm{TPG_R} = \left(\frac{D_1}{D_2}\right)^2 \cdot \frac{B_2}{B_1} \tag{2.15}$$

For $\mathrm{ECG_R} = \mathrm{TPG_R} = 1$, the first expression in Eq. (2.15) gives $\frac{D_2}{D_1} = \sqrt{\frac{P_2}{P_1}} = 0.563$ or $D_2 =$ 563 m while the second expression gives $\frac{B_2}{B_1} = \left(\frac{D_2}{D_1}\right)^2 = \frac{P_2}{P_1} = 0.317$ or $B_2 = 6.34$ MHz for $B_1 = 20$ MHz. The implication of this result is far reaching–if there is no requirement to increase RAN throughput, then the system bandwidth requirement can be reduced by using

smaller, low powered cells. Alternatively, without affecting the RAN throughput of each, several operators sharing a single wideband infrastructure would be more energy-efficient. Though the power consumption attributed to backhaul has been included in the analysis, the provision of backhaul connectivity in small cells would need to be resolved.

As $\mathrm{ETG_R} = \mathrm{ECG_R} \cdot \mathrm{TPG_R}$, then combining the expressions in Eq. (2.15) gives a fundamental equations for the RAN energy throughput gain in terms of BTS powers and system bandwidths as shown in Eq. (2.16).

$$\mathrm{ETG_R} = \frac{B_2}{B_1} \cdot \frac{P_1}{P_2} = \frac{P_1/B_1}{P_2/B_2} \tag{2.16}$$

The expression for $\mathrm{ETG_R}$ in Eq. (2.16) is independent of the intersite distance ratio $\frac{D_2}{D_1}$ and depends only on the ratio of the power spectral densities between fully loaded RANs that have SIR limited performance. This characteristic is observed in Figures 2.4(c) and 2.3(c) for $\alpha_2 = 1$, where $\mathrm{ETG_R}$ versus $\frac{D_2}{D_1}$ is almost constant. As discussed earlier, a point of particular interest is when $\mathrm{ECG_R} = \mathrm{TPG_R}$, which corresponds to when the performance of RAN_2 achieves the largest possible joint energy consumption and throughput gains with respect to RAN_1. Equating the expressions in Eq. (2.15) gives the expression in Eq. (2.17) for the optimum intersite distance ratio that gives the highest joint energy consumption and throughput gain of RAN_2 over RAN_1. That is, Eq. (2.17) defines the intersite distance ratio for least energy consumption and highest throughput gain.

$$\frac{D_2}{D_1} = \left(\frac{B_2}{B_1} \cdot \frac{P_2}{P_1}\right)^{\frac{1}{4}} \tag{2.17}$$

The 1/4 exponent in Eq. (2.17) demonstrates the stringent restrictions on $\frac{D_2}{D_1}$ when simultaneously constrained by energy and throughput requirements. The macro/microcell scenario considered in Figure 2.3 where $P_1 = 594$ W, $P_2 = 188$ W, and $B_2 = B_1 = 20$ MHz gives $\frac{D_2}{D_1} = 0.75$ or $D_2 = 750$ m for $D_1 = 1000$ m. This result agrees with the plots in Figures 2.4(a) and (b) when $\alpha_2 = 1$. The corresponding values of $\mathrm{ETG_R}$ is given by $\frac{P_1}{P_2} = 3.16$ while $\mathrm{ECG_R} = \mathrm{TPG_R} = \sqrt{\mathrm{ETG_R}} = 1.78$. Equations (2.16) and (2.17) illustrate that the key drivers in RAN energy and throughput performance are system bandwidth and BTS power consumption. The former should be maximized and the latter minimized.

2.5.3 BTS Sleep Mode

The impact of sleep mode can be assessed in the context of changing the intersite distance between the base station sites of a regular, homogeneous macrocell deployment. As considered previously, in the analysis, each BTS uses a single antenna with an omnidirectional azimuth beam pattern. Then, a RAN of n_2 operating cells with an intersite distance D_2 can be transformed into a sleep mode RAN of n_1 operating cells with an intersite distance $D_1 = \sqrt{3}D_2$ by powering down $n_2 - n_1$ cells, where $\frac{n_2}{n_1} = \frac{A_1}{A_2} = \left(\frac{D_1}{D_2}\right)^2 = 3$. This configuration equates to one where no cells in the RAN for intersite spacing D_2 have adjacent cells that are operational when in sleep mode. Providing both intersite distances D_2 and D_1 achieve SIR limited

performance, then there is no need to increase the transmit power on the downlink, though consideration should be given to the uplink sensitivity to ensure that mobile uplink transmissions are satisfactorily received.

For this configuration, the expressions for $\mathrm{ECG_R}$ and $\mathrm{TPG_R}$ in Eq. (2.10) can be modified for the sleep mode case as given in Eqs. (2.18) and (2.19). The term in β is a sleep mode factor on the BTS overhead power consumption with a value between 0 and 1. However, the overhead power consumption due to backhaul is not scaled by β as it is unlikely that the backhaul would be placed into sleep mode for signaling reasons. In the analysis next, a value of $\beta = 0.5$ is used throughout.

$$\begin{aligned}\mathrm{ECG_R} &= \frac{(P_1 + 2\beta P_{\mathrm{oh},1})/A_1}{P_2/A_2} \\ &= \frac{1}{3} \cdot \left(\frac{\alpha_1 P_{\mathrm{rh},1} + (1+2\beta) P_{\mathrm{oh},1} + 2(1-\beta) P_{bh,1}}{\alpha_2 P_{\mathrm{rh},2} + P_{\mathrm{oh},2}} \right) \end{aligned} \tag{2.18}$$

$$\mathrm{TPG_R} = \frac{S_2/A_2}{S_1/A_1} = 3 \cdot \left(\frac{\alpha_2 B_2 \eta_2}{\alpha_1 B_1 \eta_1} \right) \tag{2.19}$$

In order to calculate the $\mathrm{ECG_R}$ in Eq. (2.18) for a macrocell sleep mode scenario, the following equalities are used: $B_2 = B_1, P_{\mathrm{rh},2} = P_{\mathrm{rh},1}$, and $P_{\mathrm{oh},2} = P_{\mathrm{oh},1}$ with values being taken from Tables 2.2 and 2.3. In this macrocell example, a value of $D_2 = 500$ m is used giving an equivalent intersite distance in sleep mode of $D_1 = 866$ m. The condition for establishing sleep mode is realized when $\mathrm{TPG_R} = 1$ for $\alpha_1 = 1$, which corresponds to the sleep mode RAN being fully loaded. Through analysis, the value of η_1 was calculated at 3.26 bit/s/Hz. The equivalent load and spectral efficiency in the RAN without sleep mode activated are given by α_2 and η_2, respectively. Solving for these values when $\mathrm{TPG_R} = 1$ using Eq. (2.19) gives $\alpha_2 = 0.21$ and $\eta_2 = 5.19$ bit/s/Hz. That is, when the load in the RAN falls below ~ 0.21 the RAN will enter sleep mode. At this point, the $\mathrm{ECG_R}$ is determined using Eq. (2.18) with $\alpha_1 = 1$ and $\alpha_2 = 0.21$. Then, for $P_{\mathrm{rh}} = 279$ W and $P_{\mathrm{oh}} = 315$ W, an $\mathrm{ECG_R} = 0.82$ is obtained. A value of $\mathrm{ECG_R} = 0.82$ means that the macrocell RAN in sleep mode consumes 18% less energy compared with normal full load operation. If the exercise is repeated using microcells instead of macrocells for $D_2 = 500$ m, an $\mathrm{ECG_R} = 0.74$ is achieved corresponding to 26% less energy being consumed by the microcell RAN in sleep mode. However, repeating the exercise using picocells for $D_2 = 150$ m gives an $\mathrm{ECG_R} = 0.88$, which corresponds to only 12% less energy being consumed by the picocell RAN in sleep mode. The higher potential energy savings in a microcell RAN is attributed to the larger percentage overhead power in microcells (78% compared to 53% in macrocells), which can be scaled down by β during sleep mode. However, in a picocell RAN, the backhaul power consumption, which accounts for almost 50% of the BTS power, is not scaled by β and becomes a limiting factor. If the backhaul power in picocells is scalable by β, then the energy savings rise to 30%.

2.5.4 Heterogeneous Networks

In this sequel, the energy efficiency of HetNets is explored by considering various small-cell overlays of a macrocell RAN. ECG_H, TPG_H, and ETG_H denote energy, throughput, and joint energy/throughput figures of merit for the HetNet RAN. The calculation of these figures of

merit is performed in a straightforward manner using the fundamental metrics of cell-average area power density and area throughput density as these quantities can be added to represent the superposition of different HetNet layers. The analysis considers the evolution of a homogeneous macrocell RAN with an intersite distance D_1 and an operating bandwidth B_1. The HetNet is realized by adding a homogeneous overlay of small cells with an intersite distance D_2 and operating bandwidth $B_2 = B_1$ MHz. The small-cell overlay is implemented on an adjacent frequency band, thereby eliminating interference between layers. The intersite distance D_2 is selected to give efficient HetNet energy and throughput characteristics. For comparison, the energy/throughput characteristics of homogenous small-cell networks are determined when operating over the same total spectrum bandwidth $(B_1 + B_2)$. Expressions for the ECG_H and TPG_H derived from the area power density and area throughput density are shown in Eqs. (2.20) and (2.21), respectively, where subscript 1 refers to the macrocell RAN underlay and subscript 2 refers to the small-cell RAN overlay.

$$\mathrm{ECG}_H = \frac{P_1/A_1}{P_1/A_1 + P_2/A_2} = \frac{\mathrm{ECG}_2}{1 + \mathrm{ECG}_2} \tag{2.20}$$

$$\mathrm{TPG}_H = \frac{S_1/A_1 + S_2/A_2}{S_1/A_1} = 1 + \mathrm{TPG}_2 \tag{2.21}$$

The terms in ECG_2 and TPG_2 can be evaluated using the formulas in Eq. (2.15) for SIR limited performance when the load factor for both layers is set to unity. The ETG_H is obtained by taking the product of ECG_H and ETG_H. As discussed previously, the joint energy and throughput gain is maximized when $\mathrm{ECG}_H = \mathrm{TPG}_H$. Equating ECG_H and TPG_H, as defined in Eqs. (2.20) and (2.21), gives the expression in Eq. (2.22) in terms of the BTS powers, bandwidths, and intersite distances.

$$\frac{B_2}{B_1} \cdot \left(\frac{D_1}{D_2}\right)^2 + \frac{B_2}{B_1} \cdot \frac{P_1}{B_2} + 1 = 0 \tag{2.22}$$

Substituting $x = D_2/D_1, a = P_1/P_2$ and $b = B_2/B_1$ into Eq. (2.22) gives the general algebraic expression $(1 + ab)x^2 + b = 0$. Evaluation of this function shows that there is no solution for $x = (D_2/D_1)^2$ in the range $0 \leq \left(\frac{D_2}{D_1}\right)^2 \leq 1$. This result illustrates a key characteristic of two tier networks—there is no optimal trade-off between energy savings and throughput gain. The addition of an overlay network always increases the overall energy consumption within the RAN. Equation (2.20) shows that ECG_H is always less than unity. The best way to make a two-tier HetNet energy efficient is to maximize ECG_2, that is, maximize the energy consumption gain in the overlay with respect to the underlay. Therefore, the primary reason for adding a small-cell overlay is to enhance the RAN capacity and this is demonstrated by Eq. (2.21) whereby the overlay throughput gain TPG_2 linearly adds to the underlay throughput TPG_1 to give TPG_H.

The performances of three HetNet topologies are compared and contrasted with various single tier small-cell topologies in Table 2.4. A homogeneous macrocell RAN (i.e., RAN_1) with $D_1 = 500$ m and $B_1 = 10$ MHz is used as a reference baseline for comparing the various topologies explored.

In Table 2.2, the HetNet configurations HN1, HN2, and HN3 correspond to micro/macrocell, pico/macrocell, and pico/microcell RANs, respectively. In the case of HN1&2, D_2 has been selected with respect to D_1 to achieve $\mathrm{ECG}_2 = \mathrm{TPG}_2$ with the original macrocell RAN forming the underlay. In the case of HN3, the underlay has been replaced by a microcell RAN with

Table 2.4 HetNet versus small cell—energy and throughput gains

RAN	Macro	Micro	Micro	Pico	Pico	HN1	HN2	HN3	Units
B_2	20	10	20	10	20	10	10	10	MHz
D_1	500	500	500	500	500	500	500	375	m
D_2	500	375	446	216	257	375	216	216	m
ECG	1	1.78	2.51	5.31	7.52	0.64	0.84	1.33	
TPG	2	1.78	2.51	5.34	7.51	2.78	6.33	7.12	
ETG	2	3.17	6.32	28.35	56.47	1.78	5.33	9.45	

$D_1 = 375$ m, which gives ECG = TPG in the microcell RAN with respect to the original macrocell RAN. Then, a picocell overlay is added with $D_2 = 216$ m, which establishes ECG = TPG in the picocell overlay with respect to the microcell underlay. The results in Table 2.2 show that in HN1&2, the $\text{ECG}_H < 1$ with the pico/macro deployment being more energy-efficient than the micro/macro deployment. Both HetNets achieve $\text{TPG}_H > 1$ with the pico/macro HetNet showing significantly greater throughput gains than the micro/macro HetNet. A similar trend is observed for the ETG_H figure of merit values. In contrast, HN3 achieves an $\text{ECG}_H > 1$, which is attributed to the energy savings achieved by replacing the macrocell underlay with a microcell underlay. The picocell overlay in HN3 helps to boost the throughput significantly in HN3 giving the highest TPG_H and ETG_H of the three HetNet configurations considered.

However, Table 2.2 also shows results for homogeneous small-cell deployments encompassing microcell and picocell RANs only. What is striking about these results is that small-cell RANs offer significantly better energy and throughput gains than HetNets. Even for the same bandwidth of 10 MHz, a picocell RAN can achieve a joint energy/throughput gain of 28.35 compared with the best HetNet, which achieves 9.45. The enhancements are even greater when the small-cell bandwidth is increased to 20 MHz giving an ETG of 56.47 in the picocell RAN. While further increases in the TPG_H and ETG_H for a HetNet can be achieved by allowing both layers to share the total system bandwidth of 20 MHz, the increase in intercell interference limits the gains achievable. An upper bound for a 20 MHz HetNet with under- and overlays sharing the same frequency band can be calculated from the data in Table 2.2. For example, a 20 MHz microcell underlay (Table 2.2, colum 4) with ECG = TPG ~ 2.5 and a 20 MHz picocell overlay (Table 2.2, colum 6) with ECG = TPG ~ 7.51 can be combined to give a 40 MHz HetNet with an $\text{ECG} = ((1/2.5) + (1/7.5))^{-1} = 1.875$, a $\text{TPG} = 2.5 + 7.5 = 10$, and an ETG = 18.75. As these figures of merit represent the HetNet performance when the under- and overlay are on separate 20 MHz bands, the performance when sharing the same 20 MHz band is expected to be significantly less. Therefore, for the prevailing power consumption models used in this analysis, small cells would appear to offer significantly greater potential than HetNets for saving energy while increasing throughput even when backhaul effects are taken into account.

2.6 Conclusions

This chapter has taken a comprehensive look at the energy consumption problem in broadband wireless networks, in particular cellular mobile radio networks. With research on 5G

considering techniques to increase throughput such as cell densification, HetNets, and massive MIMO, there remains a significant dearth of the underlying theories and principles needed to deliver the required increase in data throughput without compromising the overall RAN energy efficiency. The issue has been addressed in this chapter by developing metrics that enable different RAN architectures to be fairly compared in terms of relative energy consumption and throughput figures of merit. In particular, a new composite figure of merit has been developed called the RAN Energy Throughput Gain (ETG_R), which enables salient RAN parameters to be set in order to achieve the best possible combination of energy reduction and throughput gain. The ETG_R is formed by taking the product of the ECG_R and the TPG_R and demonstrates the inverse relationship between throughput gain and energy consumption gain. Importantly, the basis of these figures of merit relates to comparing the fundamental metrics of area power density in W/km^2, area throughput density in $bit/s/km^2$, and, in the case of the ETG_R, the energy consumption rating in J/bit. It is the optimization of these figures of merit that drives the design of an energy-efficient RAN. At the centre of such designs is the requirement for power consumption models of the RAN hardware, in particular the BTS. Such models should accurately account for the load-independent as well as the load-dependent causes of power consumption. Furthermore, the backhaul consumption should be taken into account, particularly in small cells such as pico- and femtocells where it can account for approximately 50% of the BTS power consumption. For these cell types, the backhaul can be a limiting factor to achieving an energy-efficient deployment as backhaul power consumption cannot be readily controlled.

The second part of this chapter examines a number of RAN architectures based on the LTE standard. Four key scenarios were considered: reducing intersite distance, deploying small-cell technologies; introducing sleep modes; and deploying HetNets. In all four cases, the RAN ECG, TPG, and ETG figures of merit were used to characterize the RAN energy and throughput efficiencies. The RAN configurations were evaluated for constant transmit power and SIR limited performance, which is characteristic of urban deployments where macrocell intersite distances of ≤ 1 km are common. When intersite distance was reduced for the same BTS hardware, the figures of merit confirmed that the energy consumption increased monotonically. The throughput also increased monotonically demonstrating the benefit of using smaller cells to increase the RAN capacity but the throughput gains were only achieved at the expense of increased energy consumption.

In order to mitigate this behavior, the macrocells were replaced by lower power microcells. The figures of merit were used to identify two salient deployment options. The first option established the least intersite distance between microcells that resulted in the microcell RAN having the same energy consumption and throughput as the macrocell RAN (i.e., $ECG_R = TPG_R = ETG_R = 1$). For the same system bandwidth, this operating point is achieved by lowering the cell load α to a feasible level that limits the intercell interference. The main enhancements experienced by the microcell RAN in this case were higher peak and cell-edge throughputs (or capacities). By increasing the intersite distance at the feasible load, the ECG_R can be increased above unity but the TPG_R will decrease below unity such that their product ETG_R remains almost equal to one. Alternatively, if the least intersite distance is held constant while the load factor is increased, then the ECG_R remains almost constant at unity while the TPG_R increases giving the desired increase in throughput with reduced cell size for no further increase in RAN energy consumption.

However, operating the microcell RAN at the feasible load significantly under-utilizes the available radio resources. If no further increase in area throughput density is required, then

an alternative strategy is to reduce the system bandwidth B instead of the cell load factor. For this case, each cell of both the microcell and macrocell RANs can be operated at full load with $\alpha = 1$, and, therefore, each cell achieves approximately the same cell-average spectral efficiency. Then, both microcell and macrocell RANs achieve the same power spectral density giving the fundamental relationship that the bandwidth ratio B_2/B_1 is equal to the BTS power ratio P_2/P_1 when $\mathrm{ECG_R} = \mathrm{TPG_R} = 1$. The LTE standard allows the air-interface bandwidth to be selected from a range of bandwidth values equal to 1.4, 3, 5, 10, 20 MHz. As each bandwidth mode operates with full transmit power, further reductions in power consumption could be considered by selecting the minimum transmit power needed to achieve SIR limited performance while meeting a cell-edge peak throughput rate.

The second option established the least intersite distance for $\mathrm{ECG_R} = \mathrm{TPG_R}$ when the load factor $\alpha = 1$. That is, all RANs are compared at maximum load. This condition established an operating point, which jointly maximized the $\mathrm{ECG_R}$ and the $\mathrm{TPG_R}$ given by Eq. (2.17). In particular, when the systems had the same bandwidths and their performances were SIR limited, the ratio of intersite distances and the ratio of BTS powers followed a fourth power law. This relationship confirmed that reductions in BTS power consumption can be traded for joint improvements in energy and throughput efficiency.

The behavior of a RAN in sleep mode was investigated. Though placing cells into sleep mode offers energy savings by scaling down the overhead power in nonoperational cells excluding backhaul, the savings are realized for low load factors. Over the diurnal day, the opportunity for entering sleep mode may be limited leading to small overall energy savings. The savings from sleeping microcells is potentially greater than that obtained by sleeping macro- or picocells. With respect to macrocells, the microcells exhibit a higher proportion of overhead power that can be scaled down during sleep mode. However, in picocells, the backhaul power consumption, which is not scaled down, becomes a limiting factor.

The energy and throughput efficiency of HetNets were investigated in the context of evolving a macrocell RAN into a large-cell underlay and small- cell overlay using separate frequency bands to eliminate interference between the layers. In addition, the energy and throughput performance of a number of homogeneous small-cell deployments were compared with the HetNet options. The main conclusions drawn were that small-cell deployments achieved considerably greater energy savings and throughput gains than HetNets. In HetNets, the addition of an overlay contributes an extra energy consuming component into the overall RAN while inhibiting the optimum trade-off between energy savings and throughput gains. The main advantage of a small-cell overlay is the enhancement of throughput. However, the gains remain significantly less than those achievable by using a homogeneous deployment of small cells. Even when the two tiers share the same frequency band, allowing interlayer interference, the upper bound on potential energy savings and throughput gains were considerably less than those achieved by small cells.

Overall, the chapter has established a robust evaluation framework for assessing the energy and throughput performance of an LTE-based RAN. The results demonstrate the fundamental roles played by the cell-site power consumption model, including the backhaul, the cell-average area power and throughput densities (efficiencies), and the energy consumption rating in joules/bit. Based on the models and analyses provided, the evolution of a macrocell cell RAN into a high throughput, energy-efficient RAN might be satisfactorily achieved by first evolving into a HetNet of small cells over macrocells followed by removing the macrocell underlayer leaving a small-cell RAN. An important ingredient in future small-cell

deployments will be the minimization of the energy consumption in the backhaul. The backhaul power consumption can account for 50% of the overall BTS power consumption in a picocell or femtocell and can become a limiting factor especially when considering sleep mode options. An interesting solution to this problem might be to transform the macrocell underlay into a collection of massive MIMO-based macrocells providing backhaul connectivity to numerous off-grid, small cells. When cloud RAN techniques are taken into the technology mix, the options for energy and throughput efficient RANs become prolific. Hopefully, this chapter can contribute to this rapidly expanding field by establishing a consistent methodology for accurately assessing energy and throughput trade-offs in RAN planning.

References

[1] J. He, P. Loskot, T. O'Farrell, et al., "Energy efficient architectures and techniques for green radio access networks," in Communications and Networking in China (Chinacom), Beijing 2010.

[2] J.H. Tsai, H.W. Lo, and W.C. Chou, "Evaluation of mobile services for the future of 3G operators," IJMC, vol. 7, no. 4, pp. 470–493, 2009.

[3] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, "The global footprint of mobile communications: the ecological and economic perspective," IEEE Commun. Mag., vol. 49, no. 8, pp. 55–62, 2011.

[4] EC, "Impacts of information and communication technologies on energy efficiency," *Bio Intelligence Service*, 2008.

[5] Vodafone, "Sustainablility report-for the year ending 31 March 2011," *Vodafone Group plc*, 2011.

[6] W. Riaz, J. Gutierrez, and J. Pedersen, "Strategies for the next generation green ICT infrastructure," in *ISABEL International Symposium on Applied Sciences in Biomedical and Communication Technologies*, Bratislava, 2009.

[7] B. Heile, "Smart grids for green communications," IEEE Wireless Commun., vol. 17, pp. 4–6, 2010.

[8] C. Han, et al., "Green radio: radio techniques to enable energy-efficient wireless networks," IEEE Commun. Mag., vol. 49, no. 6, pp. 46–54, 2011.

[9] B. Badic, T. O'Farrrell, P. Loskot, and J. He, "Energy efficient radio access architectures for green radio: Large versus small cell size deployment," in *IEEE Vehicular Technology Conference Fall (VTC 2009-Fall)*, Anchorage, Alaska, USA, pp. 1–5, 2009.

[10] H.-O. Scheck, "ICT & wireless networks and their impact on global warming," in *Procedings of European Wireless Conference*, Lucca, Italy, pp. 911–915, 2010.

[11] M. Lasanen, et al., "Environmental friendly mobile radio networks: approaches of the European OPERA-Net 2 project," in *20th International Conference on Telecommunications (ICT 2013)*, Casablanca, Morocco, May 6–8, 2013.

[12] G. Auer, et al., "How much energy is needed to run a wireless network?," IEEE Wireless Commun., vol. 49, no. 6, pp. 40–49, 2011.

[13] Cisco, Cisco Visual Networking index:global Mobile, Cisco, 2012.

[14] K. Gerwig, "Tablet devices could change user behaviour and network capacity planning," *SearchTelecom.com*, March 2012.

[15] O. Holland, et al., "Opportunistic load and spectrum management for mobile communications energy efficiency," in *IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2011)*, Totonto, 2011.

[16] J. Elling, M. Sorensen, P. Mogensen, and E. Lang, "Mobile broadband network evolution towards 2015-a Copenhagen area case study," Telektronikk, vol. 106, no. 1, pp. 138–148, 2010.

[17] Vodafone, "Performance data summary," 2012. [Online]. Available: http://www.vodafone.co.uk/our -responsibilities/performance-data-summary/. [Accessed 22 March 2014].

[18] Digital Communications Knowledge Transfer Network, "Energy Efficient Wireless Communications (Green Radio Access Networks)," in *Wireless Technology and Spectrum Working Group*, London, 2011.

[19] J. Louhi, "Energy efficiency of modern cellular base stations," in *29th International Telecommunications Energy Conference*, Rome, 2007.

[20] M. Deruyck, W. Joseph, and L. Martens, "Power consumption model for macrocell and microcell base stations," Trans. Emerging Telecommun. Technol., vol. 25, no. 3, pp. 320–333, 2014.

[21] F. Richter, A. Fehske, and G. Fettweis, "Energy efficiency aspects of base station deployment strategies for cellular networks," in *IEEE 70th Vehicular Technology Conference (VTC 2009-Fall)*, Anchorage, Alaska, USA, 2009.

[22] G. Micallef, "*Energy efficient evolution of mobile broadband networks*," PhD Thesis, Aalborg University, 2013.

[23] S. Tombaz, et al., "Impact of backhauling power consumption on the deployment of heterogeneous mobile networks," in *IEEE Global Communications Conference (GLOBECOM 2011)*, Houston, TX, 2011.

[24] W. Guo, C. Turyagyenda, H. Hamdoun, S. Wang, P. Loskot and T. O'Farrell, et al., "Towards a low energy LTE cellular network: Architectures", Proceedings of 19th European Signal Processing Conference, Barcelona, Spain, 29 August - 2 September 2011, Invited Paper.

[25] J. Salo, M. Nur-Alam, and K. Chang, "Practical introduction to LTE radio planning," 2010. [Online]. Available: http://digitus.itk.ppke.hu/ takacsgy/lte_rf_wp_02Nov2010.pdf. [Accessed 22 March 2014].