5G Green Networking: Enabling Technologies, Potentials, and Challenges

Kemal Davaslioglu and Richard D. Gitlin
Department of Electrical Engineering
University of South Florida
Email: {kemald, richgitlin}@usf.edu

Abstract—We are in the midst of a major change in mobile wireless networks. Driven by the massive number of mobile-connected devices and the constant increase in the data rates, the primary objective of wireless network operators has been to satisfy the throughput of users and maximize the network capacity. However, this has lead to an energy inefficient network design. Our goal in this paper is to discuss the potential solutions and key enabling technologies that will facilitate network operators introduce power savings and improve energy efficiency in next-generation wireless networks. Specifically, we focus on the energy efficiency aspects of massive multiple-input multiple output systems (or also referred to as Large-Scale Antenna Systems), millimeter-wave communications, and dense deployment of small cells. With the goal of an energy-efficient network design, we identify the recent advances, quantify how much gain can be achieved, and present a comprehensive summary of open problems in these areas.

I. INTRODUCTION

Over the past 20 years, demand for ubiquitous connectivity and data rates have been constantly increasing, while the spectrum resources have been scarce [1]. Unlike wireline communications where installing new optical cables can satisfy the demand, spectrum scarcity creates significant problems. Enabling technologies for 5G Networks, such as large-scale antenna systems (LSAS), harnessing new spectrum bands through using millimeter wave communications or carrier aggregation, and dense deployment of small cells will provide viable solutions to achieve pervasive connectivity and high data rates. While overcoming these challenges, it is becoming more evident that the carbon footprint of information and communications technologies should not be overlooked, see the examples in [1]. In this paper, we study how these enabling technologies can achieve a green 5G Network. We will briefly describe the technologies that facilitate green networking: Massive MIMO, mmWave communications, and small cells.

Massive multiple-input multiple-output (MIMO) or Large-Scale Antenna Systems (LSAS), is an form of multi-user technology which employs an array of unconventionally large number of antennas (on the order of hundreds or more) that are low-powered and physically small in size. These antennas are individually controlled to transmit very narrow beams to concentrate power efficiently to the users. Unlike point-to-point MIMO, massive MIMO systems enjoy high multiplexing gain even in the case of line-of-sight (LOS) conditions (assuming that angular separation between the terminals is larger than the array resolution). As the number of antenna elements gets larger, better angular resolution can be achieved and interference can be reduced. This facilitates significant power savings.

Millimeter wave (mmWave) refers to the spectrum between 3-300 GHz. Due to the scarcity in cellular frequencies, there has been a growing interest in using the mmWave bands for short-range and fixed wireless communications. For example, the new wireless local area network (WLAN) standard IEEE 802.11ad operates on the 57-64 GHz oxygen absorption band to provide multigigabit data rates using channel bandwidths of 2160 MHz. Local multipoint distribution service (LMDS) uses 27-31 GHz for broadband wireless point-to-multipoint operation for the last mile [2]. In these applications (see also [3] for more applications in mmWave), channel bandwidths are typically two orders of magnitude more than the available bands in the cellular frequencies. Despite the higher losses in the propagation and attenuation characteristics, mmWave communications provide high antenna gains as they employ large antenna arrays with much smaller form factor and can achieve multigigabit data rates due to the vast bandwidth [4].

Small cells are low-cost low-powered base stations that are deployed as an underlay for the macrocell tier to offload low-mobility users, provide seamless coverage, and most importantly improve the user experience. Dense deployment of small cells will play a key role in 5G Networks to complement the macrocell base stations and often are this network architecture is often referred to as a Heterogeneous Network (HetNet). From a green networking perspective, reducing the link distances is the key advantage of small cells to achieve substantial energy savings in the network [5].

In the following, we focus our attention on these enabling technologies. We study their unique advantages, present a quantification of the achievable theoretical gain and a summary of experiment gains. Despite the high potential of these technologies, we point out their current limitations and the challenges ahead in each technology. The rest of this paper is organized as follows. Section II describes the main properties of these enabling technologies. In Section III, we provide an overview of how much gain can be achieved. Section III summarizes challenges ahead and offers guidance on the research directions. In Section V, we conclude with a comprehensive view of how these three enabling technologies can work hand in hand to pave the way to provide a green 5G Network.
II. ENABLING TECHNOLOGIES

A. Massive MIMO

In multi-user MIMO systems, channel resources (such as time and frequency) are spatially shared by users rather than an orthogonal use which is commonly practiced in point-to-point MIMO. When the resources are spatially shared, the resulting interference can be eliminated using precoders such as vector perturbation and lattice-aided methods, and decoders such as successive interference cancellation, which bring diversity per link and provide the degrees-of-freedom (DoF) to separate users in the spatial domain [6], [7]. Massive MIMO takes this one step further and does not necessitate such complex precoding/decoding schemes which are not scalable to process the inputs of large number of antennas. In general, three decoding schemes are used for massive MIMO systems. First, the simplest choice is maximum ratio combining (MRC) that coherently adds the signals from $M$ antennas using the channel estimates. This amplifies the signal strength proportionally and this is called as array gain. The second widely used method is zero-forcing (ZF) in which the goal is to minimize the interference, but this comes at the cost of reducing the array gain. Minimum mean squared error (MMSE) combining performs in between the two methods by amplifying signal strength and suppressing interference [8], [9]. On one hand, in the high SNR regime, ZF outperforms matched filtering and conjugate beamforming, and achieves highest spectral efficiency [8]. On the other hand, the advantage of the latter two decoding schemes is that distributed and decentralized signal processing can be carried out locally at each antenna. This can increase the robustness and resilience of the antenna system to failures in hard terrain conditions. As we will discuss later in Section III, these decoders enable us to turn off several antennas in low traffic conditions when the objective is to save power. With very large antenna arrays, the simplest methods such as conjugate beamforming and matched filter decoding perform nearly optimal [10]. Thus, massive MIMO systems can be used to replace the high-power consuming 40 W power amplifiers at macrocells to increase network energy efficiency. To this end, simulations have been conducted in [10] where the authors compare the performance of a current Long Term Evolution (LTE) macrocell base station and a massive MIMO base station. Their results demonstrate that the network energy efficiency can be increased by up to three orders of magnitude in a dense urban deployment using 128 antennas at a macrocell base station and carefully selecting the transmit power and concurrent users, see Fig. 4 of [10]. Obviously, real applications will deviate from theoretical results within some margin, but it is clear that tremendous power savings and throughput gains are achievable using massive MIMO.

System performance closely depends on how accurate the channel state information (CSI) is for both uplink and downlink channels. In current cellular systems, a base station sends pilot signals and its user terminals receive, quantize, and feed the CSI back to the base stations. For massive MIMO applications where the number of antennas at a base station is envisioned to be two orders of magnitude more than the current cellular systems, CSI acquisition would require much more resources. On the uplink, the problem is easily addressed by user terminals transmitting pilot signals and base station antennas estimating the channel based on these pilots. The required overhead scales linearly with the number of users, $K$, and independent of the number of antennas, $M$. Time division duplex (TDD) systems benefit from the channel reciprocity to estimate the propagation channel. Since channel reciprocity does not hold for frequency division duplex (FDD) systems, CSI is acquired by a combination of downlink training and feedback and we summarize the corresponding challenges later in Section IV-2.

An interesting phenomenon in massive MIMO is the channel hardening effect. Using tools from random matrix theory, it is demonstrated in [11], [12] that the effects of small-scale fading can be averaged out as the number of antennas increase. In this paper, we omit its proof for brevity and focus on its importance. With channel hardening, the distribution of singular values of the channel matrix approaches a deterministic function [12], or simply, the channel randomness vanishes and it starts to look deterministic. This fact greatly simplifies resource allocation, particularly frequency allocation, and power control reduces to admission control on how to determine the number of active terminals [13].

B. Exploiting the Millimeter Wave Spectrum

Due to the scarcity of available spectrum in the microwave bands, research efforts have focused on mmWave communications and carrier aggregation. Specifically, mmWave communication offers a potential 100 GHz new spectrum for mobile communications, which is about 200 times of what current cellular frequencies offer [2]. According to the Friis’ equation, received power decreases with the square of frequency. At higher bands with smaller wavelengths, a smaller received power can be expected. However, as a fair comparison, we can see that for the same antenna size, more antennas can be packed in mmWave and this can achieve substantial antenna gains. In terms of channel exponent, the measurement results in [14] demonstrate that mmWave signals have comparable PL exponent compared to the microwave signals when beamforming is applied at transmit and receive antennas. In [14], it is reported that at distances of up to 200-300 m, PL exponents of 3.2-4.58 and 1.68-2.3 are observed for NLOS and LOS environments, respectively. Due to the constructive addition of reflected paths in LOS environments, PL exponents lower than 2 are frequent [14]. When coupled with massive MIMO, mmWave transmissions are highly directional. With proper encoder/precoder schemes, interference will have a much smaller detrimental effect than it has in the current cellular networks. An interesting remark is that in order to harness the underutilized mmWave using beamforming, basic multiple access procedures of cellular systems such as cell search, synchronization, and random access need to be redesigned [15]. There are several challenges for mmWave communications such as channel characterization of the mmWave, connectivity.
in NLOS conditions, and coverage and mobility problems, which need to be solved for an effective deployment, and we will discuss these in detail in Section IV. When the performance of the current LTE standards and a mmWave system at 28 GHz are compared under fair conditions, reference [15] demonstrates that more than 40 times throughput improvement is achievable, assuming the mmWave system uses a TDD with a 20% overhead and a 50% uplink/downlink duty cycle.

C. Convergence to a Dense Heterogeneous Network

In HetNets, a macrocell base station provides connectivity and service for high-mobility users, while low-powered small cells such as picocells, femtocells, relays, and radio remote heads serve low-mobility users at higher data rates, all of which constitutes an umbrella-like coverage. In this architecture, each base station type has different capabilities, transmit power, range, power consumption, access, backhaul, and operating functionalities. The base stations in a HetNet need to have self-optimizing, self-organizing, and self-healing capabilities. They need to dynamically adapt to the traffic load and user mobility conditions in the network through smart user association and handovers mechanisms, effective intercell interference mitigation systems, and support coordinated multipoint (CoMP) transmissions. Current standards of 4G LTE and LTE-Advanced already include the deployment of small cells in a HetNet architecture. It is envisioned that the infrastructure of next-generation 5G networks will have a much denser small cell deployment [16]. To mitigate the created interference in a dense deployment, 4G standards propose several enhanced intercell interference coordination (eICIC) methods, see [1], [17] for a comprehensive overview. Despite several open issues which are identified in Section IV, the gains that can be achieved by a dense small cell deployment is substantial. For example, the simulation study by Qualcomm [16] estimates that a dense small cell deployment of 144 small cells per macrocell can provide more than two orders of magnitude (up to 145x).

III. ACHIEVABLE GAINS

In this part, we provide an overview of the theoretical and practical gains that can be achieved through using the enabling technologies for 5G Networks.

A. Theoretical Gains

First, from a green communication perspective, the major contribution of massive MIMO systems is to reduce the transmit power levels significantly while keeping the service quality the same. For example, in [11], the authors identify that as the number of antennas at a BS, $M$ grows asymptotically, the uplink transmit power per user can be reduced by $\frac{1}{M}$ when perfect CSI is available and by $\frac{1}{\sqrt{M}}$ when CSI is not perfect. Although initial results in the seminal work in [18] are derived for the asymptotic case, the results would still hold for finite antennas [13]. Also, it is shown that close to optimal results can be achieved with simple precoders and decoders such as eigenbeamforming (BF) and matched filtering (MF) [19]. It is shown in [8] that these significant gains can be achieved with robust and decentralized signal processing methods such as just a conjugate beamformers.

Having very high degrees of freedom (DoF) in massive MIMO relaxes many of the strict constraints on the hardware. One important example is relaxing the constraints on power amplifiers. As discussed in detail in [1], the majority of the energy inefficiency in a cellular network resides at the base stations. Due to the high peak-to-average ratio (PAPR) requirements, high power back-off values are used. Since the transmit powers can be reduced proportional to the number of antennas, the PAPR can be significantly decreased and eventually a constant envelope can be used with very modest penalties [20]. This fact alone would provide substantial power savings and improve the energy efficiency in a cellular network. It can be expected that LTE macrocells will be replaced by massive MIMO macrocells in the near future [10]. This would not only help a network operator provide the rate demands but also be green at the same time.

The vast spectrum in the mmWave offers significant gains. In [15], it is shown that a 20-fold increase in overall cell capacity can be achieved solely due to the bandwidth without any spatial multiplexing or advanced schemes. The link budget analysis in [2] shows that a very high data rate of 2 Gbps is achievable at 1 km distance using mmWave in an urban mobile environment. A good example that demonstrates the joint deployment of a massive MIMO and mmWave technologies is presented in [21]. The authors show that average cell throughput and cell-edge user performance can be improved by a factor of 30 for a system where a macrocell base station uses 12 horn antennas (one per sector), users have 4 element uniform linear array for receiver beamforming, and a system bandwidth of 500 MHz at 28 GHz.

B. Prototypes and Experiment Results

The gains we summarized above mostly demonstrate the theoretical gains which are very significant. Recently, a number of prototypes and testbeds have been built demonstrating the gains in massive MIMO systems, see [14], [22], [23], [24]. These projects have been crucial in stimulating research on massive MIMO and have provided evidence that the theoretical gains are achievable. For example, one of the first prototypes, Argos [22] demonstrated that for a base station with 64 antennas serving 15 clients simultaneously, the total system capacity can be increased to 85 bps/Hz with ZF multi-user beamforming (MUBF) and 35 bps/Hz with conjugate MUBF. In other words, the system capacity can be increased by factors of 3 to 6.7 and at the same time can use 1/64th of the transmission power of a base station with a single antenna [22]. Its successor, Faros [23] improves the transmission range at over 250 meters outdoors by 40 dB at 2.4 GHz and achieves a transmit power less than 100 $\mu$W per antenna while using 108 antennas that serve 5 concurrent users [23]. In another prototype, researchers have shown in [24] that a base station with 128 antennas simultaneously serving 16 co-located users that are spaced far apart in a non-line-of-sight (NLOS) scenario can
achieve 75% to 90% of the asymptotic capacity depending on the angular separation of the users. As a mmWave prototype using a 520 MHz bandwidth at 27.925 GHz, reference [14] presents experiment results that achieve a 528 Mbps data rate with 32 antennas at a few-hundred meters of outdoor coverage while supporting a user mobility of 8 kmp/h in an NLOS environment. These prototypes have demonstrated that the substantial theoretical gains of large scale antenna arrays in massive MIMO and mmWave communications are in fact realizable.

IV. CHALLENGES AHEAD

Although there are substantial achievable gains with these enabling technologies, there are several major challenges and still open problems.

1) Duplexing: The duplexing mode determines the interference conditions in a cellular network. In TDD, base station transmissions are synchronized such that a base station’s downlink and a user’s uplink in different cells do not interfere. The main advantage of TDD is the channel reciprocity which greatly simplifies CSI acquisition. However, most of the legacy systems in 3G and 4G networks in North America and Europe employ FDD so that most network operators have paired spectrum bands. The bottleneck here is the overhead required for CSI acquisition and pilot signals significantly increase for massive MIMO systems (see the pilot contamination issues in [9], [25]). Clearly, efficient and scalable methods are required for FDD systems.

Duplexing users of multiple tiers in a HetNet is also an important problem. A recent paper [25] addresses this problem by studying four different schemes in a multi-tier architecture: (i) FDD, (ii) TDD in which macrocells and small cells operate in non-overlapping bands, (iii) coTDD in which macrocell and small cells operate in the same spectrum, (iv) coRTDD in which macrocell and small cells operate in the same spectrum but the order of the uplink and downlink periods are reversed. Depending on the network deployment, one mode may outperform another, see Fig. 5 in [25]. A hybrid duplexing scheme that adapts to the interference conditions, user mobility, and network load can provide extra energy savings. While developing such methods, care needs to be given to avoid outage scenarios such as the interference between small cell’s downlink and macrocell associated user’s uplink.

2) CSI Acquisition and Limitations on the Simultaneous Users: As we discussed above, the performance of a massive MIMO system is closely related to the CSI availability. Uplink CSI acquisition is prone to pilot contamination [7], [9], [19]. The ultimate limitation of of massive MIMO is the number of users simultaneously served. This is limitation is due to the overhead required for CSI acquisition, see [9]. Thus, efficient and scalable methods with low overhead are required to estimate the channel for high-mobility users. Also, note that channel reciprocity refers to the channel propagation and it does not compensate for random phase and amplitude differences in the transceiver front-end. To obtain full channel reciprocity, methods such as self-calibration and external calibration mechanisms can be employed [25]. A potential solution addressing the CSI acquisition problem is given in [9] where the authors propose to deploy extremely large arrays on the sides of skyscrapers. These can be reserved to provide service for low-mobility users to take away some of the load in smaller arrays in which the excess capacity can be used for high mobility users.

3) Antenna and User Selection to Maximize Energy Efficiency: Turning off some antennas and introducing sleep modes in low load periods are effective ways to provide substantial energy savings in cellular networks, see [1]. The same principles can be applied to massive MIMO systems where some of the RF chains can be turned off with a small performance penalty. An optimized system would dynamically select the antennas and dynamically turn-off portions of the hardware in low traffic. A recent study in [26] investigates how many antennas are needed to achieve the desired performance limit and how many more antennas are needed by the MF to achieve MMSE performance. The problem of finding the optimal number of antennas that maximize energy efficiency is investigated in [8], [27] for linear precoding schemes of conjugate and zero-forcing beamforming. Other studies in [28], [29] jointly optimize uplink and downlink transmission to maximize energy efficiency and determine the optimal number of base station antennas, simultaneous users, and transmit power to uniformly cover a service area. For the uplink, it is proposed in [30] that the network energy efficiency is maximized when some users are switched off. Another good reference in this area is [29] in which studies the dependence of network energy efficiency on the number of base station transmit antennas, number of users, base station power consumption characteristics, and precoder/decoder design. Moreover, when a densely deployed macrocells is considered with an underlay of small cells using microwave bands, network performance can be improved by flexible clustering and efficient user selection [31]. We can see that there is a potential for significant savings in network power consumption by dynamically selecting the active antennas at a base station and the precoder/decoder design, determining the optimal number of concurrent users, and associated transmit power levels in both uplink and downlink.

4) Need for low-cost and low-power hardware: In [1], it is discussed that energy efficient hardware components would substantially improve network energy efficiency. This message is even more critical for massive MIMO and mmWave systems that have a large number of antennas. Although theoretical findings and prototypes demonstrate significant reductions are realizable in the transmitted power, more research needs to be carried out to decrease the amount of power consumed in the hardware components. As the number of antennas scales up, the advances in low power application-specific integrated circuits (ASICs), single chip ADCs/DACs [32], efficient RF power amplification and combining [15], and conformal antenna arrays will play crucial roles in 5G green networks to determine the achievable energy savings.
5) Propagation Characteristics of mmWave: Along with the path loss (PL) characteristics in the mmWave, which are mainly dominated by free-space PL, other channel propagation characteristics also cause detrimental effects depending on the spectrum and link distance. For instance, in the 5764 GHz band, atmospheric attenuation due to oxygen absorption can cause an additional loss of 15 dB/km [2]. At 80 GHz, a 10 meter foliage penetration would create 23.5 dB loss which is 15 dB higher compared to the one at 3 GHz [2]. Furthermore, due to lack of channel models, there has been several recent studies to fully characterize and model the mmWave channel, see [15], [33] and references therein. Although recent studies shed some light on how to model the mmWave channel, there limited to a specific outdoor urban environment. Others terrains and indoor channel models for conference rooms, cubicles, livingroom, etc. also need to be fully characterized.

6) Supporting Multi-user Diversity and Heterogeneous Applications: Unlike current systems, 5G Networks need to provide multiple streams to multiple users simultaneously. This brings challenges for medium access control (MAC) protocols. Next generation MAC-layer protocols need to scale for multicell coordination in mitigating intercell interference and accommodate the large number of antennas and users. While facing these challenges, they also need to support a variety of different applications with various delay and bit rate requirements. Support of adaptive frame sizes can facilitate power savings in the PHY layer.

7) Supporting Massive MIMO and cell densification on the higher Protocol Layers: Massive MIMO and mmWave systems benefit from supporting multiple users simultaneously. This poses new problems on the higher layers of the protocol stack such as providing connectivity, uniform user experience, fast handshaking protocols [9] and reserving spare frequency bands for the uplink pilot signals of new users. Also, the highly-directional transmissions create the hidden node and exposed node problems (see [34]), and robust and efficient neighbor discovery protocols needs to be developed. Two good examples along these lines are the phantom cells for dense small cells [35] and the booster cell in an anchor-booster architecture for mmWave and HetNet integration [36]. Both studies address the mobility and coverage issues, and propose to differentiate the control (C-plane) and user data plane (U-plane). For instance, a mobile user can be connected to a macrocell using microwave bands through the C-plane maintaining its connectivity and connect to a small cell using mmWave through the U-plane for a high-data rate communication. Despite their simplicity, these schemes allow for seamless connectivity, reduce handover failures significantly, and achieve better load balancing. Since inefficiencies such as re-transmissions or repeated connection requests are avoided, these protocol changes improve the network energy efficiency.

8) Non-Cellular Applications: The impact of these enabling technologies are not limited to cellular networks. For example, in [9], it is discussed that billboard-sized arrays in suburban and rural areas can be deployed to provide high-speed fixed wireless access service to multiple houses.

Another study in [37] investigates using distributed antennas for massive MIMO systems. Recently, new use cases for massive MIMO and mmWave focus on directional multihop relaying for wireless backhaul transmissions [2], [4].

9) Security: Current communications systems are prone to signal jamming and are easy to eavesdrop. A unique feature of Massive MIMO is that it offers higher DoF to cancel the jamming signals and to minimize the eavesdropper’s capacity. However, intentional jamming is still an issue. Specifically, the uplink pilots in TDD are vulnerable to jamming due to the low transmit power levels of user terminals. A recent study in [20] addresses this problem by jointly using channel estimation and decoding to avoid jamming effects. However, robust and resilient algorithms are needed to exploit the benefits of massive MIMO against jamming while still being green.

V. CONCLUSIONS

In this paper, enabling technologies of 5G networks have been investigated from the point of view of network energy efficiency and green communications. We summarized their unique advantages, identified the achievable gains from both theoretical and practical perspectives, and pointed out the current challenges ahead in each category. It is clear that the challenges in 5G cannot be solved through a single solution. In what follows, we provide an overview of how combinations of these technologies can be deployed and what their potentials are when used in a complementary fashion.

A. Massive MIMO and mmWave Applications

Current MIMO systems typically operate at microwave frequencies where the transceiver hardware design is very well-known [9], and scaling up the number of antenna elements from current systems to massive MIMO can be certainly achieved. Using the mmWave spectrum, many more antennas can be packed to a given area which makes it possible to achieve a compact-size low-powered massive MIMO implementation with a practical form factor. However, the challenges of both systems such as the lack of full channel characterization in mmWave, shorter coherence time and
higher Doppler spread restricting the user mobility and spatial multiplexing capabilities, maturity of transceiver design for massive MIMO, etc. need to be solved such that these large gains and very high potential power savings can be realized.

B. Massive MIMO, mmWave, and Small Cells Applications

A fully synchronized TDD-based HetNet deployment in which a macrocell base station employs massive MIMO with an underlay of small cell deployments using the mmWave bands can benefit from the complementary advantages of all enablers. Massive MIMO provides interference suppression and spatial multiplexing, mmWave offers underutilized large bandwidths, while small cells reduce link distances and mitigate coverage holes in the deployment area. The channel reciprocity of TDD systems allows devices to reuse the received interference covariance matrix estimate which is used for interference-aware precoding, see e.g., [25]. With this setup, simple precoders using only local information can be used. This facilitates a distributed and scalable implementation which is very much needed. Moreover, using well-known beamforming schemes, interference in the system can be isolated, and thus a high frequency reuse can be carried out. Coupled with the abundant spectrum in the mmWave and reduced link distances, dense small cells can operate on the energy-efficient regime. Through such a complementary and robust design, it is evident that the network energy efficiency can be significantly improved while addressing the challenges of 5G Networks.

REFERENCES