What Will 5G Be?

Jeffrey G. Andrews, *Fellow, IEEE*, Stefano Buzzi, *Senior Member, IEEE*, Wan Choi, *Senior Member, IEEE*, Stephen V. Hanly, *Member, IEEE*, Angel Lozano, *Fellow, IEEE*, Anthony C. K. Soong, *Fellow, IEEE*, and Jianzhong Charlie Zhang, *Senior Member, IEEE*

***Abstract*—What will 5G be? What it will *not* be is an incremental advance on 4G. The previous four generations of cellular technology have each been a major paradigm shift that has broken backward compatibility. Indeed, 5G will need to be a paradigm shift that includes very high carrier frequencies with massive bandwidths, extreme base station and device densities, and unprecedented numbers of antennas. However, unlike the previous four generations, it will also be highly integrative: tying any new 5G air interface and spectrum together with LTE and WiFi to provide universal high-rate coverage and a seamless user experience. To support this, the core network will also have to reach unprecedented levels of flexibility and intelligence, spectrum regulation will need to be rethought and improved, and energy and cost efficiencies will become even more critical considerations. This paper discusses all of these topics, identifying key challenges for future research and preliminary 5G standardization activities, while providing a comprehensive overview of the current literature, and in particular of the papers appearing in this special issue.**

***Index Terms*—Cellular systems, energy efficiency, HetNets, massive MIMO, millimeter wave, small cells.**

I. Introduction

A. The Road to 5G

In just the past year, preliminary interest and discussions about a possible 5G standard have evolved into a full-fledged conversation that has captured the attention and imagination of researchers and engineers around the world. As the long-term evolution (LTE) system embodying 4G has now been deployed and is reaching maturity, where only incremental improvements and small amounts of new spectrum can be expected, it is natural for researchers to ponder "what's next?" [1]. However, this is not a mere intellectual exercise. Thanks largely to the annual visual network index (VNI) reports released by Cisco, we have quantitative evidence that the wireless data explosion is real and will continue. Driven largely by smartphones, tablets, and video streaming, the most recent (Feb. 2014) VNI report [2] and forecast makes plain that an incremental approach will not come close to meeting the demands that networks will face by 2020.

In just a decade, the amount of IP data handled by wireless networks will have increased by well over a factor of 100: from under 3 exabytes in 2010 to over 190 exabytes by 2018, on pace to exceed 500 exabytes by 2020. This deluge of data has been driven chiefly by video thus far, but new unforeseen applications can reasonably be expected to materialize by 2020. In addition to the sheer volume of data, the number of devices and the data rates will continue to grow exponentially. The number of devices could reach the tens or even hundreds of billions by the time 5G comes to fruition, due to many new applications beyond personal communications [3]–[5]. It is our duty as engineers to meet these intense demands via innovative new technologies that are smart and efficient yet grounded in reality. Academia is engaging in collaborative projects such as METIS [6] and 5GNOW [7], while industry is driving preliminary 5G standardization activities (cf. Section IV-B). To further strengthen these activities, the public-private partnership for 5G infrastructure recently constituted in Europe will funnel massive amounts of funds into related research [8].

This article is an attempt to summarize and overview many of these exciting developments, including the papers in this special issue. In addition to the highly visible demand for ever more network capacity, there are a number of other factors that make 5G interesting, including the potentially disruptive move to millimeter wave (mmWave) spectrum, new market-driven ways of allocating and re-allocating bandwidth, a major ongoing virtualization in the core network that might progressively spread to the edges, the possibility of an "Internet of Things" comprised of billions of miscellaneous devices, and the increasing integration of past and current cellular and WiFi standards to provide a ubiquitous high-rate, low-latency experience for network users.

This editorial commences with our view of the "big three" 5G technologies: ultra-densification, mmWave, and massive multiple-input multiple-output (MIMO). Then, we consider important issues concerning the basic transmission waveform, the increasing virtualization of the network infrastructure, and the need for greatly increased energy efficiency. Finally, we provide a comprehensive discussion of the equally important regulatory and standardization issues that will need to be addressed for 5G, with a particular focus on needed innovation in spectrum regulation.

Manuscript received December 1, 2013; revised April 25, 2014; accepted May 25, 2014. Date of publication June 3, 2014; date of current version July 14, 2014.

J. G. Andrews is with the University of Texas at Austin, Austin, TX 78712 USA (e-mail: jandrews@ece.utexas.edu).

S. Buzzi is with the University of Cassino and Southern Latium, 03043 Cassino (FR), Italy, and also with Consorzio Nazionale Interuniversitario per le Telecomunicazioni, 03043 Cassino Frosione, Italy (e-mail: buzzi@unicas.it).

W. Choi is with the Department of Electrical Engineering, Korea Advanced Institute of Science, Daejeon 305701, Korea (e-mail: wchoi@kaist.edu).

S. V. Hanly is with Macquarie University, Sydney, N.S.W. 2109, Australia (e-mail: stephen.hanly@mq.edu.au).

A. Lozano is with the Universitat Pompeu Fabra, 08002 Barcelona, Spain (e-mail: angel.lozano@upf.edu).

A. C. K. Soong is with Huawei Technologies, Plano, TX 75024 USA (e-mail: anthony.soong@huawei.com).

J. C. Zhang is with Samsung Electronics, Richardson, TX 75082 USA (e-mail: jianzhong.z@samsung.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSAC.2014.2328098

B. Engineering Requirements for 5G

In order to more concretely understand the engineering challenges facing 5G, and to plan to meet them, it is necessary to first identify the requirements for a 5G system. The following items are requirements in each key dimension, but it should be stressed that not all of these need to be satisfied simultaneously. Different applications will place different requirements on the performance, and peak requirements that will need to be satisfied in certain configurations are mentioned below. For example, very-high-rate applications such as streaming high-definition video may have relaxed latency and reliability requirements compared to driverless cars or public safety applications, where latency and reliability are paramount but lower data rates can be tolerated.

1) Data Rate: The need to support the mobile data traffic explosion is unquestionably the main driver behind 5G. Data rate can be measured in several different ways, and there will be a 5G goal target for each such metric:

a) **Aggregate data rate** or **area capacity** refers to the total amount of data the network can serve, characterized in bits/s per unit area. The general consensus is that this quantity will need to increase by roughly 1000× from 4G to 5G.
b) **Edge rate** or 5% **rate** is the worst data rate that a user can reasonably expect to receive when in range of the network, and so is an important metric and has a concrete engineering meaning. Goals for the 5G edge rate range from 100 Mbps (easily enough to support high-definition streaming) to as much as 1 Gbps. Meeting 100 Mbps for 95% of users will be extraordinarily challenging, even with major technological advances. This requires about a 100× advance since current 4G systems have a typical 5% rate of about 1 Mbps, although the precise number varies quite widely depending on the load, the cell size, and other factors.
c) **Peak rate** is the best-case data rate that a user can hope to achieve under any conceivable network configuration. The peak rate is a marketing number, devoid of much meaning to engineers and likely to be in the range of tens of Gbps.

Meeting the requirements in (a)-(b), which are about 1000× and 100× current 4G technology, respectively, are the main focus of this paper.

2) Latency: Current 4G roundtrip latencies are on the order of about 15 ms, and are based on the 1 ms subframe time with necessary overheads for resource allocation and access. Although this latency is sufficient for most current services, anticipated 5G applications include two-way gaming, novel cloud-based technologies such as those that may be touchscreen activated (the “tactile Internet” [9]), and virtual and enhanced reality (e.g., Google glass or other wearable computing devices). As a result, 5G will need to be able to support a roundtrip latency of about 1 ms, an order of magnitude faster than 4G. In addition to shrinking down the subframe structure, such severe latency constraints may have important implications on design choices at several layers of the protocol stack and the core network (cf. Section III).

3) Energy and Cost: As we move to 5G, costs and energy consumption will, ideally, decrease, but at least they should not increase on a per-link basis. Since the per-link data rates being offered will be increasing by about 100×, this means that the Joules per bit and cost per bit will need to fall by at least 100×. In this article, we do not address energy and cost in a quantitative fashion, but we are intentionally advocating technological solutions that promise reasonable cost and power scaling. For example, mmWave spectrum should be 10–100× cheaper per Hz than the 3G and 4G spectrum below 3 GHz. Similarly, small cells should be 10–100× cheaper and more power efficient than macrocells. A major cost consideration for 5G, even more so than in 4G due to the new BS densities and increased bandwidth, is the backhaul from the network edges into the core. We address backhaul and other economic considerations in Section IV-C. As for energy efficiency, we address this more substantially in Section III-C.

C. Device Types and Quantities

5G will need to be able to efficiently support a much larger and more diverse set of devices. With the expected rise of machine-to-machine communication, a single macrocell may need to support 10 000 or more low-rate devices along with its traditional high-rate mobile users. This will require wholesale changes to the control plane and network management relative to 4G, whose overhead channels and state machines are not designed for such a diverse and large subscriber base.

II. Key Technologies to Get to 1000× Data Rate

Of the requirements outlined in Section I-B, certainly the one that gets the most attention is the need for radically higher data rates across the board. Our view is that the required 1000× will, for the most part, be achieved through combined gains in three categories:

a) Extreme densification and offloading to improve the area spectral efficiency. Put differently, more active nodes per unit area and Hz.
b) Increased bandwidth, primarily by moving toward and into mmWave spectrum but also by making better use of WiFi’s unlicensed spectrum in the 5-GHz band. Altogether, more Hz.
c) Increased spectral efficiency, primarily through advances in MIMO, to support more bits/s/Hz per node.

The combination of more nodes per unit area and Hz, more Hz, and more bits/s/Hz per node, will compound into *many* more bits/s per unit area. Other ideas not in the above categories, e.g., interference management through BS cooperation [10]–[23] may also contribute improvements, but the lion’s share of the surge in capacity should come from ideas in the above categories. In the remainder of this section, these are distilled in some detail.

A. Extreme Densification and Offloading

A straightforward but extremely effective way to increase the network capacity is to make the cells smaller. This approach has been demonstrated over several cellular generations [24], [25]. The first such generation, in the early 1980s, had cell sizes on the order of hundreds of square kms. Since then, those sizes have been progressively shrinking and by now they are often fractions of a square km in urban areas. In Japan, for instance, the spacing between BSs can be as small as two hundred meters, giving a coverage area well under a tenth of a square km. Networks are now rapidly evolving [26] to include nested small cells such as picocells (range under 100 meters) and femtocells (WiFi-like range) [27], as well as distributed antenna systems [28] that are functionally similar to picocells from a capacity and coverage standpoint but have all their baseband processing at a central site and share cell IDs.

Cell shrinking has numerous benefits, the most important being the reuse of spectrum across a geographic area and the ensuing reduction in the number of users competing for resources at each BS. Contrary to widespread belief, as long as power-law pathloss models hold the signal-to-interference ratio (SIR) is preserved as the network densifies [29].[1] Thus, in principle, cells can shrink almost indefinitely without a sacrifice in SIR, until nearly every BS serves a single user (or is idle). This allows each BS to devote its resources, as well as its backhaul connection, to an ever-smaller number of users.

As the densification becomes extreme, some challenges arise:

- Preserving the expected cell-splitting gains as each BS becomes more lightly loaded, particularly low-power nodes.
- Determining appropriate associations between users and BSs across multiple radio access technologies (RATs), which is crucial for optimizing the edge rate.
- Supporting mobility through such a highly heterogeneous network.
- Affording the rising costs of installation, maintenance and backhaul.

We next briefly discuss these challenges, particularly in view of the other technologies raised in this article.

1) Base Station Densification Gains: We define the BS densification gain $\rho(\lambda_1, \lambda_2) > 0$ as the *effective* increase in data rate relative to the increase in network density, which is a proxy here for cost. Specifically, if we achieve a data rate R_1 (could be any measure thereof, e.g., edge rate or aggregate) when the BS density is λ_1 BSs/km^2 and then we consider a higher BS density λ_2, with corresponding rate R_2, then the densification gain is the slope of the rate increase over that destiny range:

$$\rho(\lambda_1, \lambda_2) = \frac{(R_2 - R_1)/R_1}{(\lambda_2 - \lambda_1)/\lambda_1}. \tag{1}$$

For example, if the network density is doubled, and the edge data rate increases by 50% (for example since some of the added BSs were lightly loaded), then the densification gain is $\rho = 0.5$. In some applications with channel access protocols like CSMA that are inefficient in high density, it is possible that $\rho < 0$, which is colloquially referred to as "the tragedy of the commons", but for cellular network with a centralized MAC we can safely assume $\rho > 0$.

In an interference-limited network with full buffers, the signal-to-interference-plus-noise ratio (SINR) is essentially equal to the SIR and, because the SIR distribution remains approximately constant as the network densifies, the best case scenario is $\rho \approx 1$. In reality, buffers are not always full, and small cells tend to become more lightly loaded than macrocells as the network densifies. Altogether, the SINR usually increases with density: in noise-limited networks because of the increase in received signal power, and in interference-limited networks because the lightly loaded small cells generate less interference (while still providing an option for connectivity) [31]. Nevertheless, at microwave frequencies the gain in SINR is not enough to keep up with the decrease in small-cell utilization and thus $\rho < 1$. In an extreme case, consider λ_1 and R_1 held fixed with $\lambda_2 \to \infty$. In this asymptotic setting, the small cells compete for a finite pool of users, becoming ever more lightly loaded, and thus $\rho \to 0$.

Empirically and theoretically, we observe that ρ improves and can approach 1 with macro-BS muting (termed eICIC in 3GPP) vs. the macrocells transmitting all the time and thus interfering with the small cells all the time.

An intriguing aspect of mmWave frequencies is that densification gains $\rho \gg 1$ may be possible. This is because, as discussed in Section II-B, at these frequencies communication is largely noise-limited and increasing the density not only splits the cell resources and lightens the load, but it may increase the SINR dramatically. As a striking example of this, it was recently shown in [32] that, under a plausible urban grid-based deployment, increasing the BS count in a given area from 36 to 96—which decreased the inter-BS distance from 170 meters down to 85 meters—increased the 5% cell-edge rate from 24.5 Mbps up to 1396 Mbps, giving $\rho = 9.9$. While conceding that this massive densification gain corresponds to a particular setup and model, it is nevertheless remarkable.

In general, quantifying and optimizing the densification gains in a wide variety of deployment scenarios and network models is a key area for continued small-cell research.

2) Multi-RAT Association: Networks will continue to become increasingly heterogeneous as we move toward 5G. A key feature therein will be increased integration between different RATs, with a typical 5G-enabled device having radios capable of supporting not only a potentially new 5G standard (e.g., at mmWave frequencies), but also 3G, numerous releases of 4G LTE including possibly LTE-Unlicensed [33], several types of WiFi, and perhaps direct device-to-device (D2D) communication, all across a great many spectral bands. Hence, determining which standard(s) and spectrum to utilize and which BS(s) or users to associate with will be a truly complex task for the network [34] (see Fig. 1).

Determining the optimal user association is, for general utility functions, a massive combinatorial optimization problem that depends on the SINR from every user to every BS, the instantaneous load at each BS, the choices of other users in the network, and possibly other constraints such as the requirement

[1] The power-law pathloss model ceases to apply in the near field, very close to the transmitter [30].

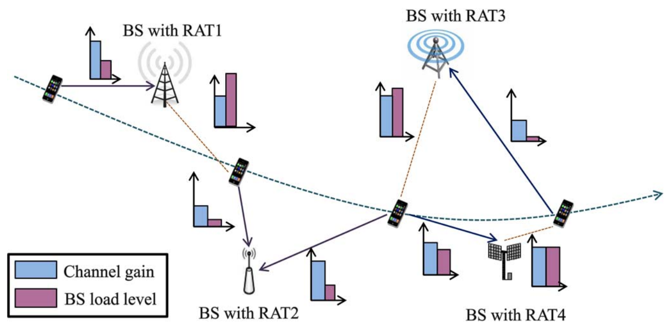

Fig. 1. User association in a multi-RAT network over many frequency bands is complex. In this simplified scenario, a mobile user in turn associates with different BSs based on a tradeoff between the gain to that BS and the traffic load (congestion) that it is experiencing.

to utilize the same BS and standard in both uplink and downlink (to facilitate functioning control channels for resource allocation and feedback) [35], [36]. Therefore, simplified procedures must be adopted [37], an example of which appears in this special issue [38]. The key such simplified procedures are "biasing" and macrocell "blanking" or "muting". Biasing refers to associating with a small cell even if it provides a lower SINR than the macrocell, and is useful for pushing users off of the heavily loaded macrocell and onto the lightly loaded small cell. Everyone wins: the remaining macrocell users get more resources while the biased users have a lower SINR/spectral efficiency but can utilize a large number of resource blocks on the small cell, ultimately attaining a higher data rate. Blanking refers to shutting off the macrocell transmissions for some fraction of the time, preferably while the biased small cell users are being served. This raises all the small-cell SINRs considerably—enough to justify actually shutting down even congested macrocell BSs—while also providing a mechanism for the biased users to hear common control channels that would otherwise be swamped by the macrocells.

Even a simple, seemingly highly suboptimal association approach based on aggressive but static biasing (about 10–20 dB, depending on various factors) toward small cells and blanking about half of the macrocell transmissions has been shown to increase edge rates by as much as 500% [39], [40]. The joint problem of user association and resource allocation in two-tier heterogeneous networks (HetNets), with adaptive tuning of the biasing and blanking in each cell, is considered in [36] and [41]–[46]. An interesting model of hotspot traffic is considered in [42]–[44] where it is shown that, under various network utility metrics, the optimal cell association is determined by rate ratio bias, rather than power (or SINR) bias.

It will be interesting to extend these models to more general scenarios. A dynamic model of cell range expansion is considered in [47], where traffic arrives as a Poisson process in time and the feasible arrival rates, for which a stabilizing scheduling policy exists, are characterized. User association and load balancing in a HetNet, with massive MIMO at the BSs, is considered in [48]. The problem of determining the optimal associations when there are multiple RATS, operating at different frequencies and using different protocols, has not yet received much attention. However, an interesting game theoretic approach is taken in [49] to the RAT-selection problem, where convergence to Nash equilibria and the Pareto-efficiency of these equilibria are studied. A related paper in this special issue [50] explores the interaction between cellular operators and WiFi network owners.

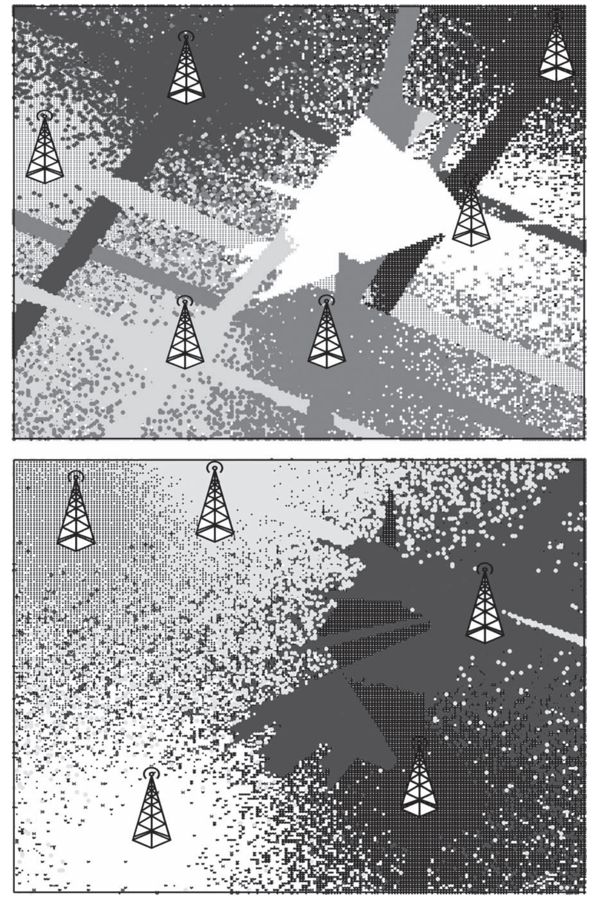

Fig. 2. Calculated mmWave BS associations with real building locations [58]. The shaded regions correspond to association with the BS centered at that shade. Blocking, LOS vs. non-LOS propagation, and beam directionality render our usual notion of cell boundaries obsolete.

Adding mmWave into the picture adds significant additional complexity, since even the notion of a cell boundary is blurry at mmWave frequencies given the strong impact of blockages, which often result in nearby BSs being bypassed in favor of farther ones that are unblocked (cf. Fig. 2). On the positive side, interference is much less important in mmWave (cf. Section II-B), and thus, the need for blanking is reduced.

In summary, there is a great deal of scope for modeling, analyzing and optimizing BS-user associations in 5G.

3) Mobility Support: Clearly, the continued network densification and increased heterogeneity poses challenges for the support of mobility. Although a hefty share of data is served to stationary indoor users, the support of mobility and always-on connectivity is arguably the single most important feature of cellular networks relative to WiFi. Because modeling and analyzing the effect of mobility on network performance is difficult, we expect to see somewhat *ad hoc* solutions such as in LTE Rel-11 [51] where user-specific virtual cells are defined to distinguish the physical cell from a broader area where the user can roam without the need for handoff, communicating with any BS or subset of BSs in that area. Or in mmWave, restricting highly mobile users to macrocells and microwave frequencies, thereby forcing them to tolerate lower rates. Handoffs will be particularly challenging at mmWave frequencies since transmit and receive beams must be aligned to communicate. Indeed, the entire paradigm of a handoff initiated and managed at layer 3 by the core network will likely not exist in 5G; instead, handoffs may be opportunistic, based on mmWave beam alignments, or indistinguishable from PHY/MAC interference management techniques whereby users communicate with multiple coordinated BSs, as exemplified by [52] in this special issue.

4) Cost: Evolving to ever-smaller cells requires ever-smaller, lower-power and cheaper BSs, and there is no fundamental reason a BS needs to be more expensive than a user device or a WiFi node [26]. Nevertheless, obtaining permits, ensuring fast and reliable backhaul connections, and paying large monthly site rental fees for operator-controlled small-cell placements have proven a major hindrance to the growth of picocell, distributed antennas, and other enterprise-quality small cell deployments. Of these, only the backhaul is primarily a technical challenge. Regulatory reforms and infrastructure sharing (cf. Section IV-C) may help address the other challenges.

Turning to end-user-deployed femtocells and WiFi access points, these are certainly much more cost-effective both from a capital and operating expense perspective [24]. However, major concerns exist here too. These include the coordination and management of the network to provide enterprise-grade service, which given the scale of the deployments requires automated self-organization [53]. A further challenge is that these end-user deployments utilize the end-user's backhaul connection and access point, both of which the end-user has a vested interest in not sharing, and in some countries a legal requirement not to. Anecdotally, all readers of this article are familiar with the scenario where a dozen WiFi access points are within range, but all are secured and inaccessible. From an engineering perspective, this closed-access status quo is highly inefficient and the cost for 5G would be greatly reduced in an open-access paradigm for small cells. One preliminary but successful example is Fon, which as of press time boasts over 13 million shared WiFi access points.

5G and all networks beyond it will be extremely dense and heterogeneous, which introduces many new challenges for network modeling, analysis, design and optimization. We further discuss some of the nonobvious intersections of extreme densification with mmWave and massive MIMO, respectively, in the next two sections. Before proceeding, however, we briefly mention that besides cell shrinking a second approach to densification exists in the form of D2D communication. This allows users in close proximity to establish direct communication, replacing two relatively long radio hops via the BS with a single and shorter direct hop. Provided there is sufficient spatial locality in the wireless traffic, this can bring about reduced power consumption and/or higher data rates, and a diminished latency [54]–[56]. Reference [57] in this special issue proposes a novel way of scheduling concurrent D2D transmissions so as to densify while offering interference protection guarantees.

B. Millimeter Wave

Terrestrial wireless systems have largely restricted their operation to the relatively slim range of microwave frequencies that extends from several hundred MHz to a few GHz and corresponds to wavelengths in the range of a few centimeters up to about a meter. By now though, this spectral band—often called "beachfront spectrum"—has become nearly fully occupied, in particular at peak times and in peak markets. Regardless of the efficacy of densification and offloading, much more bandwidth is needed [59], [60].

Although beachfront bandwidth allocations can be made significantly more efficient by modernizing regulatory and allocation procedures, as discussed in Section IV-A, to put large amounts of new bandwidth into play there is only one way to go: up in frequency. Fortunately, vast amounts of relatively idle spectrum do exist in the mmWave range of 30–300 GHz, where wavelengths are 1–10 mm. There are also several GHz of plausible spectrum in the 20–30 GHz range.

The main reason that mmWave spectrum lies idle is that, until recently, it had been deemed unsuitable for mobile communications because of rather hostile propagation qualities, including strong pathloss, atmospheric and rain absorption, low diffraction around obstacles and penetration through objects, and, further, because of strong phase noise and exorbitant equipment costs. The dominant perception had therefore been that such frequencies, and in particular the large unlicensed band around 60 GHz [61], were suitable mainly for very-short-range transmission [62]–[64]. Thus, the focus had been on WiFi (with the WiGiG standard in the 60-GHz band) and on fixed wireless in the 28, 38, 71–76 and 81–86 GHz. However, semiconductors are maturing, their costs and power consumption rapidly falling—largely thanks to the progress of the aforementioned short-range standard—and the other obstacles related to propagation are now considered increasingly surmountable given time and focused effort [65]–[70].

1) Propagation Issues: Concerning mmWave propagation for 5G, the main issues under investigation are:

Pathloss: If the electrical size of the antennas (i.e., their size measured by the wavelength $\lambda = c/f_c$ where f_c is the carrier frequency) is kept constant, as the frequency increases the antennas shrink and their effective aperture scales with $\lambda^2/4\pi$; then, the free-space pathloss between a transmit and a receive antenna grows with f_c^2. Thus, increasing f_c by an order of magnitude, say from 3 to 30 GHz, adds 20 dB of power loss regardless of the transmit-receive distance. However, if the antenna aperture at one end of the link is kept constant as the frequency increases, then the free-space pathloss remains unchanged. Further, if both the transmit and receive antenna apertures are held constant, then the free-space pathloss actually *diminishes* with f_c^2: a power gain that would help counter the higher noise floor associated with broader signal bandwidths.

Although preserving the electrical size of the antennas is desirable for a number of reasons, maintaining at the same time the aperture is possible utilizing arrays, which aggregate the individual antenna apertures: as the antennas shrink with frequency, progressively more of them must be added within the original area. The main challenge becomes cophasing these antennas so that they steer and/or collect energy productively. This challenge becomes more pronounced when the channel changes rapidly, for instance due to mobility (whose effect in terms of Doppler shift increases linearly with frequency) or due to rapid alterations in the physical orientation of the devices.

Blocking: MmWave signals exhibit reduced diffraction and a more specular propagation than their microwave counterparts, and hence they are much more susceptible to blockages. This results in a nearly bimodal channel depending on the presence or absence of Line-of-Sight (LoS). According to recent measurements [68], [70], as the transmit-receive distance grows the pathloss accrues close to the free-space value of 20 dB/decade under LoS propagation, but drops to 40 dB/decade plus an additional blocking loss of 15–40 dB otherwise. Because of the sensitivity to blockages, a given link can rapidly transition from usable to unusable and, unlike small-scale fading, large-scale obstructions cannot be circumvented with standard small-scale diversity countermeasures. New channel models capturing these effects are much needed, and in fact currently being developed [68], [71], [72] and applied to system-level analysis [58], [73]–[75] and simulation studies such as [76] and [77] in this special issue.

Atmospheric and Rain Absorption: The absorption due to air and rain is noticeable, especially the 15 dB/km oxygen absorption within the 60-GHz band (which is in fact why this band is unlicensed), but it is inconsequential for the urban cellular deployments currently envisioned [65], [67] where BS spacings might be on the order of 200 m. In fact, such absorption is beneficial since it further attenuates interference from more distant BSs, effectively increasing the isolation of each cell.

The main conclusion is that the propagation losses for mmWave frequencies are surmountable, but require large antenna arrays to steer the beam energy and collect it coherently. While physically feasible, the notion of narrow-beam communication is new to cellular communications and poses difficulties, which we next discuss.

Large Arrays, Narrow Beams: Building a wireless system out of narrow and focused beams is highly nontrivial and changes many traditional aspects of wireless system design. MmWave beams are highly directional, almost like flashlights, which completely changes the interference behavior as well as the sensitivity to misaligned beams. The interference adopts an on/off behavior where most beams do not interfere, but strong interference does occur intermittently. Overall, interference is de-emphasized and mmWave links may often be noise-limited, which is a major reversal from 4G. Indeed, even the notion of a "cell" is likely to be very different in a mmWave system since, rather than distance, blocking is often the first-order effect on the received signal power. This is illustrated in Fig. 2.

Link Acquisition: A key challenge for narrow beams is the difficulty in establishing associations between users and BSs, both for initial access and for handoff. To find each other, a user and a BS may need to scan lots of angular positions

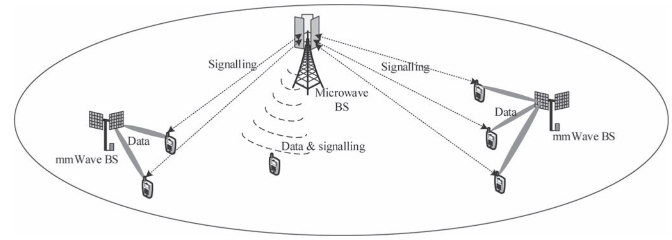

Fig. 3. MmWave-enabled network with phantom cells.

where a narrow beam could possibly be found, or deploy extremely large coding/spreading gains over a wider beam that is successively narrowed in a multistage acquisition procedure. Developing solutions to this problem, particularly in the context of high mobility, is an important research challenge.

Leveraging the Legacy 4G Network: A concurrent utilization of microwave and mmWave frequencies could go a long way toward overcoming some of the above hurdles. An interesting proposal in that respect is the notion of "phantom cells" (relabeled "soft cells" within 3GPP) [78], [79], where mmWave frequencies would be employed for payload data transmission from small-cell BSs while the control plane would operate at microwave frequencies from macro BSs (cf. Fig. 3). This would ensure stable and reliable control connections, based on which blazing fast data transmissions could be arranged over short-range mmWave links [80]. Sporadic interruptions of these mmWave links would then be far less consequential, as control links would remain in place and lost data could be recovered through retransmissions.

Novel Transceiver Architectures Needed: Despite the progress made in WiFi mmWave systems, nontrivial hardware issues remain, and in some cases will directly affect how the communication aspects are designed. Chief among these is the still-exorbitant power consumption of particularly the analog-to-digital (A/D) but also the digital-to-analog (D/A) converters operating on enormous bandwidths. A main consequence is that, although large antenna arrays and high receiver sensitivities are needed to deal with the pathloss, having customary fully digital beamformers for each antenna appears to be unfeasible. More likely are structures based on old-fashioned analog phase shifters or, perhaps, hybrid structures where groups of antennas share a single A/D and D/A [81]–[84]. On the flip side, offering some relief from these difficulties, the channels are sparser and thus the acquisition of channel-state information is facilitated; in particular, channel estimation and beamforming techniques exploiting sparsity in the framework of compressed sensing are being explored [85], [86].

C. Massive MIMO

Stemming from research that blossomed in the late 1990s [87], [88], MIMO communication was introduced into WiFi systems around 2006 and into 3G cellular shortly thereafter. In essence, MIMO embodies the spatial dimension of the communication that arises once a multiplicity of antennas are available at BSs and mobile devices. If the entries of the channel matrix that ensues exhibit—by virtue of spacing, cross-polarization and/or angular disposition—sufficient statistical independence, multiple spatial dimensions become available for signaling and the spectral efficiency multiplies accordingly [89], [90].

In single-user MIMO (SU-MIMO), the dimensions are limited by the number of antennas that can be accommodated on a mobile device. However, by having each BS communicate with several users concurrently, the multiuser version of MIMO (MU-MIMO) can effectively pull together the antennas at those users and overcome this bottleneck. Then, the signaling dimensions are given by the smallest between the aggregate number of antennas at those users and the number of antennas at the BS.

Furthermore, in what is now known as coordinated multipoint (CoMP) transmission/reception, multiple BSs can cooperate and act as a single effective MIMO transceiver thereby turning some of the interference in the system into useful signals; this concept in fact underpins many of the approaches to interference and mobility management mentioned earlier in this section.

Well-established by the time LTE was developed, MIMO was a native ingredient thereof with two-to-four antennas per mobile device and as many as eight per BS sector, and it appeared that, because of form factors and other apparent limitations, such was the extent to which MIMO could be leveraged. Marzetta was instrumental in articulating a vision in which the number of antennas increased by more than an order of magnitude, first in a 2007 presentation [91] with the details formalized in a landmark paper [92]. The proposal was to equip BSs with a number of antennas *much* larger than the number of active users per time–frequency signaling resource, and given that under reasonable time–frequency selectivities accurate channel estimation can be conducted for at most some tens of users per resource, this condition puts the number of antennas per BS into the hundreds. This bold idea, initially termed "large-scale antenna systems" but now more popularly known as "massive MIMO," offers enticing benefits:

- Enormous enhancements in spectral efficiency without the need for increased BS densification, with the possibility—as is always the case—of trading some of those enhancements off for power efficiency improvements [93], [94].
- Smoothed out channel responses because of the vast spatial diversity, which brings about the favorable action of the law of large numbers. In essence, all small-scale randomness abates as the number of channel observations grows.
- Simple transmit/receive structures because of the quasi-orthogonal nature of the channels between each BS and the set of active users sharing the same signaling resource. For a given number of active users, such orthogonality sharpens as the number of BS antennas grows and simple linear transceivers, even plain single-user beamforming, perform close-to-optimally.

The promise of these benefits has elevated massive MIMO to a central position in preliminary 5G discussions [95], with a foreseen role of providing a high-capacity umbrella of ubiquitous coverage in support of underlying tiers of small cells. However, for massive MIMO to become a reality, several challenges must first be overcome and the remainder of this section is devoted to their dissection. For recent contributions on these and other aspects, the reader is referred to a companion special issue on massive MIMO [96]. The present special issue contains further new contributions, mentioned throughout the discussion that follows, plus reference [97] dealing with the massification of MIMO multicasting [98], [99].

1) Pilot Contamination and Overhead Reduction: Pilot transmissions can be made orthogonal among same-cell users, to facilitate cleaner channel estimates [100], [101], but must be reused across cells—for otherwise all available resources would end up consumed by pilots. This inevitably causes interference among pilots in different cells and hence puts a floor on the quality of the channel estimates. This interference, so-called "pilot contamination," does not vanish as the number of BS antennas grows large, and so is the one impairment that remains asymptotically. However, pilot contamination is a relatively secondary factor for all but colossal numbers of antennas [102]. Furthermore, various methods to reduce and even eliminate pilot contamination via low-intensity BS coordination have already been formulated [103], [104]. Still, a careful design of the pilot structures is required to avoid an explosion in overhead. The ideas being considered to reign in pilot overheads include exploiting spatial correlations, so as to share pilot symbols among antennas, and also segregating the pilots into classes (e.g., channel strength gauging for link adaptation v. data detection) such that each class can be transmitted at the necessary rate, and no faster.

2) Architectural Challenges: A more serious challenge to the realization of the massive MIMO vision has to do with its architecture. The vision requires radically different BS structures where, in lieu of a few high-power amplifiers feeding a handful of sector antennas, we would have a myriad of tiny antennas fed by correspondingly low-power amplifiers; most likely each antenna would have to be integrated with its own amplifier. Scalability, antenna correlations and mutual couplings, and cost, are some of the issues that must be sorted out. At the same time, opportunities arise for innovative topologies such as conformal arrays along rooftops or on building facades, and we next dwell on a specific topological aspect in which innovation is taking place.

Within this special issue, [105] explores alternative and highly innovative antenna designs based on the utilization of an electromagnetic lens-focusing antenna.

3) Full-Dimension MIMO and Elevation Beamforming: Existing BSs mostly feature linear horizontal arrays, which in tower structures can only accommodate a limited number of antennas, due to form factors, and which only exploit the azimuth angle dimension. By adopting planar 2D arrays as in Fig. 3 and further exploiting the elevation angle, the so-called full-dimension MIMO (FD-MIMO) can house many more antennas with the same form factor [106]. As a side benefit, tailored vertical beams increase the signal power and reduce interference to users in neighboring cells. Some preliminary cell average and edge data rates obtained from Samsung's network simulator are listed in Table I where, with numbers of antennas still modest for what massive MIMO is envisioned to be, multiple-fold improvements are already observed.

TABLE I
FD-MIMO SYSTEM-LEVEL DOWNLINK SIMULATION RESULTS AT 2.5 GHZ. HALF-WAVELENGTH ANTENNA SPACINGS IN BOTH THE HORIZONTAL AND VERTICAL DIMENSIONS AT THE BSS, 2 ANTENNAS PER USER, 30% OVERHEAD. THE BASELINE IS SU-MIMO WITH 4 ANTENNAS PER BS, AND THE FD-MIMO RESULTS (AVERAGE AND EDGE DATA RATES) ARE FOR MU-MIMO WITH 16 AND 64 ANTENNAS, RESPECTIVELY, CORRESPONDING TO 4×4 AND 8×8 PLANAR ARRAYS PER BS SECTOR

	SU-MIMO	FD-MIMO 16	FD-MIMO 64
Aggregate Data Rate (b/s/Hz/cell)	2.32	3.28	6.37
Edge Data Rate (b/s/Hz)	0.063	0.1	0.4

4) Channel Models: Parallel to the architectural issues run those related to channel models, which to be sound require extensive field measurements. Antenna correlations and couplings for massive arrays with relevant topologies must be determined, and a proper modeling of their impact must be established; in particular, the degree of actual channel orthogonalization in the face of such nonidealities must be verified. And, for FD-MIMO, besides azimuth, the modeling needs to incorporate elevation [106]–[108], which is a dimension on which far less data exists concerning power spectra and angle spreads. A 3D channel modeling study currently under way within 3GPP is expected to shed light on these various issues [109]. References [108] and [110] in this special issue also deal with this subject.

5) Coexistence With Small Cells: As mentioned earlier, massive MIMO BSs would most likely have to coexist with tiers of small cells, which would not be equipped with massive MIMO due to their smaller form factor. Although the simplest alternative is to segregate the corresponding transmissions in frequency, the large number of excess antennas at massive MIMO BSs may offer the opportunity of spatial nulling and interference avoidance with relative simplicity and little penalty. To confirm the feasibility of this idea, put forth in [111] and further developed in [112] within this special issue, comprehensive channel models are again needed.

As networks become dense and more traffic is offloaded to small cells, the number of active users per cell will diminish and the need for massive MIMO may decrease. Aspects such as cost and backhaul will ultimately determine the balance between these complementary ideas.

6) Coexistence With mmWave: As discussed in Section II-B, mmWave communication requires many antennas for beamsteering. The antennas are much smaller at these frequencies and thus very large numbers thereof can conceivably fit into portable devices, and these antennas can indeed provide beamforming power gain but also MIMO opportunities as considered in [113] within this special issue. Any application of massive MIMO at mmWave frequencies would have to find the correct balance between power gain/interference reduction and parallelization.

III. DESIGN ISSUES FOR 5G

In addition to supporting 1000× higher data rates, 5G networks must decrease latencies, lower energy consumption, lower costs, and support many low-rate connections. In this section, we discuss important ongoing research areas that support these requirements. We begin with the most fundamental aspect of the physical layer—the waveform—and then consider the evolution of cloud-based and virtualized network architectures, latency and control signaling, and energy efficiency.

A. The Waveform: Signaling and Multiple Access

The signaling and multiple access formats, i.e., the waveform design, have changed significantly at each cellular generation and to a large extent they have been each generation's defining technical feature. They have also often been the subject of fierce intellectual and industrial disputes, which have played out in the wider media. The 1G approach, based on analog frequency modulation with FDMA, transformed into a digital format for 2G and, although it employed both FDMA and TDMA for multiple access, was generally known as "TDMA" due to the novelty of time multiplexing. Meanwhile, a niche spread spectrum/CDMA standard that was developed by Qualcomm to compete for 2G [114] became the dominant approach to all global 3G standards. Once the limitations of CDMA for high-speed data became inescapable, there was a discreet but unmistakable retreat back toward TDMA, with minimal spectrum spreading retained and with the important addition of channel-aware scheduling [115]. Due to the increasing signal bandwidths needed to support data applications, orthogonal frequency-division multiplexing (OFDM) was unanimously adopted for 4G in conjunction with scheduled FDMA/TDMA as the virtues of orthogonality were viewed with renewed appreciation.

In light of this history, it is natural to ponder the possibility that the transition to 5G could involve yet another major change in the signaling and multiple access formats.

1) OFDM and OFDMA: The Default Approach: OFDM has become the dominant signaling format for high-speed wireless communication, forming the basis of all current WiFi standards and of LTE, and further of wireline technologies such as digital subscriber lines, digital TV, and commercial radio. Its qualities include:

- A natural way to cope with frequency selectivity.
- Computationally efficient implementation via FFT/IFFT blocks and simple frequency-domain equalizers.
- An excellent pairing for MIMO, since OFDM allows for the spatial interference from multiantenna transmission to be dealt with at a subcarrier level, without the added complication of intersymbol interference.

From a multiple access vantage point, OFDM invites dynamic fine-grained resource allocation schemes in the digital domain, and the term OFDMA is employed to denote orthogonal multiple access at a subcarrier level. In combination with TDMA, this parcels the time–frequency grid into small units known as resource blocks that can be easily discriminated through digital filtering [116]. Being able to do frequency and time slot allocation digitally also enables more adaptive and sophisticated interference management techniques such as fractional frequency reuse or spectrum partitions between small cells and macrocells. Finally, given its near-universal adoption, industry has by now a great deal of experience with its implementation, and tricky aspects of OFDM such as frequency offset correction and synchronization have been essentially conquered.

2) Drawbacks of OFDM: Given this impressive list of qualities, and the large amount of inertia in its favor, OFDM is the unquestionable frontrunner for 5G. However, some weak points do exist that could possibly become more pronounced in 5G networks.

First, the peak-to-average-power ratio (PAPR) is higher in OFDM than in other formats since the envelope samples are nearly Gaussian due to the summation of uncorrelated inputs in the IFFT. Although a Gaussian signal distribution is capacity-achieving under an average power constraint [117], in the face of an actual power amplifier a high PAPR sets up an unattractive tradeoff between the linearity of the transmitted signal and the cost of the amplifier. This problem can be largely overcome by precoding the OFDM signals at the cost of a slightly more involved equalization process at the receiver and a slight power penalty; indeed, this is already being done in the LTE uplink [118].

Second, OFDM's spectral efficiency is satisfactory, but could perhaps be further improved upon if the requirements of strict orthogonality were relaxed and if the cyclic prefixes (CPs) that prevent interblock interference were smaller or discarded. Paper [119] in this special issue proposes a novel OFDMA-based modulation scheme named frequency and quadrature amplitude modulation (FQAM), which is shown to improve the downlink cell-edge rate.

Perhaps the main source of concerns, or at least of open questions, is the applicability of OFDM to mmWave spectrum given the enormous bandwidths therein and the difficulty of developing efficient power amplifiers at those frequencies. For example, a paper in this special issue proposes a single-carrier signaling with null cyclic prefix as an alternative to OFDM at mmWave frequencies [76].

3) Potential Alternatives to OFDM: To address OFDM's weaknesses, we now overview some alternative approaches being actively investigated. Most of these, however, can be considered incremental departures from OFDM rather than the step-function changes that took place in previous cellular generations. Further tutorial treatment can be found in a recent tutorial [120].

Time–Frequency Packing: Time–frequency packing [121] and faster-than-Nyquist signaling [122]–[124] have been recently proposed to circumvent the limitations of strict orthogonality and CP. In contrast to OFDM, where the product of the symbol interval and the subcarrier spacing equals 1, in faster-than-Nyquist signaling products smaller than 1 can be accommodated and spectral efficiency improvements on the order of 25% have been claimed.

Nonorthogonal Signals: There is a growing interest in multicarrier formats, such as filterbank multicarrier [125], that are natively nonorthogonal and thus do not require prior synchronization of distributed transmitters. A new format termed universal filtered multicarrier (UFMC) has been proposed whereby, starting with an OFDM signal, filtering is performed on groups of adjacent subcarriers with the aim of reducing sidelobe levels and intercarrier interference resulting from poor time/frequency synchronization [126], [127] (see Fig. 4).

Filterbank Multicarrier: To address the drawbacks of rectangular time windowing in OFDM, namely the need for large guard bands, [128] shows that the use of filterbank multicarrier permits a robust estimation of very large propagation delays and of arbitrarily high carrier frequency offsets, whereas OFDM would have required a very long CP to attain the same performance levels.

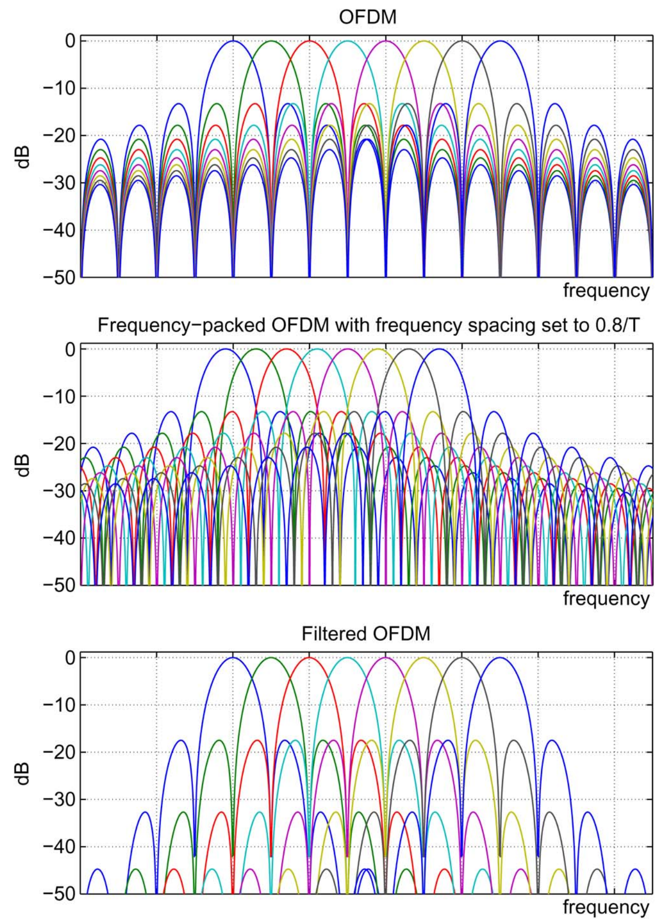

Fig. 4. Frequency-domain magnitude responses of some adjacent waveforms for OFDM, frequency-packed OFDM, and filtered OFDM. The two signaling formats alternative to OFDM trade subcarrier orthogonality for either better spectral efficiency (frequency-packed OFDM) or lower out-of-band emissions (filtered OFDM).

Generalized Frequency Division Multiplexing: GFDM is a multicarrier technique that adopts a shortened CP through the tail biting technique and is particularly well suited for noncontiguous frequency bands [129], [130], which makes it attractive for spectrum sharing where frequency-domain holes may have to be adaptively filled.

Single Carrier: Single-carrier transmission has also been attracting renewed interest, chiefly due to the development of low-complexity nonlinear equalizers implemented in the frequency domain [131].

Tunable OFDM: We conclude with our own opinion that OFDM could be well adapted to different 5G requirements by allowing some of its parameters to be tunable, rather than designed for essentially the worst-case multipath delay spread. In particular, given the increasingly software-defined nature of radios, the FFT block size, the subcarrier spacing and the CP length could change with the channel conditions: in scenarios with small delay spreads—notably dense urban/small cells and mmWave channels—the subcarrier spacing could grow and the FFT size and the CP could be significantly shortened to lower the latency, the PAPR, the CP's power and bandwidth penalty, and the computational complexity; in channels with longer delay spreads, that could revert to narrower subcarriers, longer FFT blocks, and a longer CP. This is conceptually similar to the null-cyclic-prefix single-carrier scheme proposed by NSN in this special issue [76], which is essentially OFDM with DFT-precoding (to reduce PAPR) and a punctured variable-length null prefix that is fixed with oversampling at the receiver.

B. Cloud-Based Networking

Although this special issue is mainly focused on the air interface, for the sake of completeness we briefly touch on the exciting changes taking place at the network level. In that respect, the most relevant event is the movement of data to the cloud so that it can be accessed from anywhere and via a variety of platforms. This fundamentally redefines the endpoints and the time frame for which network services are provisioned. It requires that the network be much more nimble, flexible and scalable. As such, two technology trends will become paramount in the future: network function virtualization (NFV) and software defined networking (SDN). Together, these trends represent the biggest advance in mobile communication networking in the last 20 years, bound to fundamentally change the way network services are provided. Although the move toward virtualization is thus far taking place only within the core network, this trend might eventually expand toward the edges. In fact, the term cloud-RAN is already being utilized, but for now largely to refer to schemes whereby multiple BSs are allowed to cooperate [132]. If and when the BSs themselves become virtualized—down to the MAC and PHY—this term will be thoroughly justified [133].

1) Network Function Virtualization: NFV enables network functions that were traditionally tied to hardware appliances to run on cloud computing infrastructure in a data center. It should be noted that this does not imply that the NFV infrastructure will be equivalent to commercial cloud or enterprise cloud. What is expected is that there will be a high degree of reuse of what commercial cloud offers.

It is natural to expect that some requirements of mobile networks such as the separation of the data plane, control plane and management plane, will not be feasible within the commercial cloud. Nevertheless, the separation of the network functions from the hardware infrastructure will be the cornerstone of future architectures. The key benefit will be the ability to elastically support network functional demands. Furthermore, this new architecture will allow for significant nimbleness through the creation of virtual networks and of new types of network services [134]. A detailed description of the NFV architecture is beyond the scope of this paper, and interested readers can consult [134]–[136] and the references therein.

As virtualization of the communication network gains traction in the industry, an old concept, dating back to the 1990s, will emerge: the provision of user-controlled management in network elements. Advances in computing technology have reached a level where this vision can become a reality, with the ensuring architecture having recently been termed software defined networking (SDN).

2) Software Defined Networking: SDN is an architectural framework for creating intelligent programmable networks. Specifically, it is defined as an architecture where the control and data planes are decoupled, network intelligence and state are logically centralized, and the underlying network infrastructure is abstracted from the application [137].

The key ingredients of SDN are an open interface between the entities in the control and data planes, as well as programmability of the network entities by external applications. The main benefits of this architecture are the logical decoupling of the network intelligence to separate software-based controllers, exposing the network capabilities through an application program interface, and enabling the application to request and manipulate services provided by the network [138].

From a wireless core network point of view, NFV and SDN should be viewed as tools for provisioning the next generation of core networks with many issues still open in terms of scalability, migration from current structures, management and automation, and security.

C) Energy Efficiency: As specified in our stated requirements for 5G, the energy efficiency of the communication chain—typically measured in either Joules/bit or bits/Joule—will need to improve by about the same amount as the data rate just to maintain the power consumption. And by more if such consumption is to be reduced. This implies a several-order-of-magnitude increase in energy efficiency, which is extremely challenging. Unsurprisingly, in recent years there has been a surge of interest in the topic of energy efficient communications, as can be seen from the number of recent special issues, conferences and research projects devoted to "green communications" [139]–[141]. In addition to laudable environmental concerns, it is simply not viable from a logistical, cost or battery-technology point of view to continually increase power consumption.

Due to the rapidly increasing network density (cf. Section II-A), the access network consumes the largest share of the energy [142]. Research has focused on the following areas.

1) Resource Allocation: The literature is rich in contributions dealing with the design of resource allocation strategies aimed at the optimization of the system energy efficiency [143]–[149]; the common message of these papers is that, by accepting a moderate reduction in the data rates that could otherwise be achieved, large energy savings can be attained. Within this special issue, [150] introduces an energy-efficient coordinated beamforming design for HetNets.

2) Network Planning: Energy-efficient network planning strategies include techniques for minimizing the number of BSs for a coverage target [151] and the design of adaptive BS sleep/wake algorithms for energy savings [152]–[155]. The underlying philosophy of these papers is that, since networks have been designed to meet peak-hour traffic, energy can be saved by (partially) switching off BSs when they have no active users or simply very low traffic. Of course, there are different degrees of hibernation available for a BS[2] and attention must be paid in order to avoid unpleasant coverage holes; this is usually accomplished through an increase of the transmitted power from nearby BSs.

3) Renewable Energy: Another intriguing possibility is that of BSs powered by renewable energy sources such as solar power [156]. This is of urgent interest in developing countries lacking a reliable and ubiquitous power grid, but it is also intriguing more broadly as it allows "drop and play" small cell deployment (if wireless backhaul is available) rather than "plug and play". A recent paper showed that in a dense HetNet, plausible per-BS traffic loads can actually be served solely by energy harvesting BSs [157]. A more relaxed scenario is considered in [158], where the resource allocation makes efficient use of both renewable and traditional energy sources.

4) Hardware Solutions: Finally, much of the power consumption issues will be dealt with by hardware engineers, with recent work in low-loss antennas, antenna muting, and adaptive sectorization according to traffic requirements (see, e.g., [159]).

In summary, energy efficiency will be a major research theme for 5G, spanning many of the other topics in this article:

- True cloud-RAN could provide an additional opportunity for energy efficiency since the centralization of the baseband processing might save energy [160], especially if advances on green data centers are leveraged [161].
- The tradeoff between having many small cells or fewer macrocells given their very different power consumptions is also of considerable interest [162].
- A complete characterization of the energy consumed by the circuitry needed for massive MIMO is currently lacking.
- MmWave energy efficiency will be particularly crucial given the unprecedented bandwidths [163].

IV. Spectrum Regulation and Standardization for 5G

Departing from strictly technical issues, we now turn our attention to the crucial intersections that 5G technologies will encounter with public policy, industry standardization, and economic considerations.

A. Spectrum Policy and Allocation

As discussed in Section II-B, the beachfront microwave spectrum is already saturated in peak markets at peak times while large amounts of idle spectrum do exist in the mmWave realm. Due to the different propagation characteristics, and recalling the concept of phantom cells, future systems will need to integrate a broad range of frequencies: low frequencies for wide coverage, mobility support, and control, and high frequencies for small cells. This will require new approaches to spectrum policy and allocation methods. Topics such as massive MIMO and small cells, which address the efficient use of spectrum, must also be considered important issues in spectrum policy. Needless to say, spectrum allocation and policy is an essential topic for 5G, so this section considers the pros and cons of different approaches to spectrum regulation in that context.

1) Exclusive Licenses: The traditional approach to spectrum policy is for the regulator to award an exclusive license to a particular band for a particular purpose, subject to limitations (e.g., power levels or geographic coverage). Exclusive access gives full interference management control to the licensee and provides an incentive for investments in infrastructure, allowing for quality-of-service guarantees. Downsides include high entry barriers because of elevated sunk costs, both in the spectrum itself and in infrastructure, and that such allocations are inherently inefficient since they occur over very long time scales—typically decades—and thus the spectrum is rarely allocated to the party able to make the best economic use of it.

To address these inefficiencies, market-based approaches have been propounded [164]. Attempting to implement this idea, spectrum auctions have been conducted recently to refarm spectrum, a process whereby long-held commercial radio and TV allocations are moved to different (smaller) bands releasing precious spectrum for wireless communications; a prime example of this is the so-called "digital dividend" auctions arising from the digitization of radio and TV. However, there are claims that spectrum markets have thus far not been successful in providing efficient allocations because such markets are not sufficiently fluid due to the high cost of the infrastructure [165]. According to these claims, spectrum and infrastructure cannot be easily decoupled.

2) Unlicensed Spectrum: At the other extreme, regulators can designate a band to be "open access," meaning that there is no spectrum license and thus users can share the band provided their devices are certified (by class licenses). Examples are the industrial, scientific and medical (ISM) bands, which are utilized by many devices including microwave ovens, medical devices, sensor networks, cordless phones and especially by WiFi. With open access, barriers to entry are much lower and there is enhanced competition and innovation, as the incredible success of WiFi and other ISM-band applications makes plain.

The downside of open access is potentially unmanageable interference, no quality-of-service guarantees, and, possibly, the "tragedy of the commons," where no one achieves a desired outcome. Still, it is useful to consider the possibility of open access for bands utilized in small cells as future networks may involve multiple players and lower entry barriers may be needed to secure the emergence of small-cell infrastructures.

Although interference is indeed a significant problem in current open access networks, it is interesting to note that cellular operators nevertheless rely heavily on WiFi offloading: currently about half of all cellular data traffic is proactively offloaded through unlicensed spectrum [2]. WiFi hotspots are nothing but small cells that spatially reuse ISM frequencies. At mmWave frequencies, the main issue is signal strength rather than interference, and it is therefore plausible that mmWave bands be unlicensed, or at a minimum several licensees will share a given band under certain new regulations. This question is of pressing interest for 5G.

3) Spectrum Sharing: Options do exist halfway between exclusive licenses and open access, such as the opportunistic

[2]As an example, a BS serving few users may choose to operate on a reduced set of subcarriers, or it may switch off some of its sectors.

use of TV white space. While the potential of reusing this spectrum is enticing, it is not crystal clear that reliable communication services can be delivered that way. Alternatively, Authorized Shared Access [166] and Licensed Shared Access [167] are regulatory frameworks that allow spectrum sharing by a limited number of parties each having a license under carefully specified conditions. Users agree on how the spectrum is to be shared, seeking interference protection from each other, thereby increasing the predictability and reliability of their services.

4) Market-Based Approaches to Spectrum Allocation: Given the advantages of exclusive licenses for ensuring quality of service, it is likely that most beachfront spectrum will continue to be allocated that way. Nevertheless, better utilization could likely be obtained if spectrum markets could become more fluid [164]. To that end, liberal licenses do not, in principle, preclude trading and reallocation on a fast time scale, rendering spectrum allocations much more dynamic. Close attention must be paid to the definition of spectrum assets, which have a space as well as a time scale, and the smaller the scales, the more fluid the market [168].

In small cells, traffic is much more volatile than in macrocells and operators may find it beneficial to enter into sharing arrangements for both spectrum and infrastructure. Dynamic spectrum markets may emerge, managed by brokers, allowing licenses to spectrum assets to be bought and sold—or leased—on time scales of hours, minutes or even ms [169]. Along these lines, an interesting possibility is for a decoupling of infrastructure, spectrum and services [169]. In particular, there may be a separation between spectrum owners and operators. Various entities may own and/or share a network of BSs, and buy and sell spectrum assets from spectrum owners, via brokers. These network owners may offer capacity to operators, which in turn would serve the end customers with performance guarantees. All of this, however, would require very adaptable and frequency agile radios.

We conclude this discussion by noting that offloading onto unlicensed spectrum such as TV whitespace or mmWave bands could have unexpected results. In particular, adding an unlicensed shared band to an environment where a set of operators have exclusive bands can lead to an overall *decrease* in the total welfare (Braess' paradox) [170]. This is because operators might have an incentive to offload traffic even when this runs counter to the overall social welfare, defined as the total profit of the operators and the utilities of the users, minus the costs. An operator might have an incentive to increase prices so that some traffic is diverted to the unlicensed band, where the cost of interference is shared with other operators, and this price increase more than offsets the operator's benefits. Further, while unlicensed spectrum generally lowers barriers to entry and increases competition, the opposite could occur and in some circumstances a single monopoly operator could emerge [171] within the unlicensed bands.

B. Regulation and Standardization

1) 5G Standardization Status: Several regional forums and projects have been established to shape the 5G vision and to study its key enabling technologies [6], [172]–[174]. For example, the aforementioned EU project METIS has already released documents on scenarios and requirements [175], [176]. Meanwhile, 5G has been increasingly referred to as "IMT-2020" in many industry forums and international telecommunications union (ITU) working groups [177] with the goal, as the name suggests, of beginning commercial deployments around 2020.

To explore 5G user requirements and to elaborate a standards agenda to be driven by them, the ETSI held a future mobile summit [178] in Nov. 2013. The summit concluded, in line with the thesis of this paper, that an evolution of LTE may not be sufficient to meet the anticipated 5G requirements. That conclusion notwithstanding, 5G standardization has not yet formally started within 3GPP, which is currently finalizing LTE Rel-12 (the third release for the LTE-Advanced family of 4G standards). The timing of 5G standardization has not even been agreed upon, although it is not expected to start until later Rel-14 or Rel-15, likely around 2016–2017. However, many ongoing and proposed study items for Rel-12 are already closely related to 5G candidate technologies covered in this paper (e.g., massive MIMO) and thus, in that sense, the seeds of 5G are being planted in 3GPP. Whether an entirely new standards body will emerge for 5G as envisioned in this paper is unclear; the ongoing success of 3GPP relative to its erstwhile competitors (3GPP2 and the WiMAX Forum) certainly gives it an advantage, although a name change to 5GPP would seem to be a minimal step.

2) 5G Spectrum Standardization: Spectrum standardization and harmonization efforts for 5G have begun within the ITU. Studies are under way on the feasibility of bands above 6 GHz [179], including technical aspects such as channel modeling, semiconductor readiness, coverage, mobility support, potential deployment scenarios and coexistence with existing networks.

To be available for 5G, mmWave spectrum has to be repurposed by national regulators for mobile applications and agreement must be reached in ITU world radiocommunication conferences (WRC) on the global bands for mmWave communications. These processes tend to be tedious and lengthy, and there are many hurdles to clear before the spectrum can indeed be available. On the ITU side, WRC-18 is shaping up as the time and venue to agree on mmWave spectrum allocations for 5G.

In addition to the ITU, many national regulators have also started their own studies on mmWave spectrum for mobile communications. In the USA, the technological advisory council of the federal communications committee (FCC) has carried out extensive investigations on mmWave technology in the last few years and it is possible that FCC will issue a notice of inquiry in 2014, which is always the first step in FCC's rulemaking process for allocation of any new frequency bands. As discussed above, it is also unclear how such bands will be allocated or even how they *should* be allocated, and the technical community should actively engage the FCC to make sure they are allocated in a manner conducive to meeting 5G requirements. Historically, other national regulators have tended to follow the FCC's lead on spectrum policy.

C. Economic Considerations

The economic costs involved in moving to 5G are substantial. Even if spectrum costs can be greatly reduced through the approaches discussed above, it is still a major challenge for carriers to densify their networks to the extent needed to meet our stated 5G requirements. Two major challenges are that BS sites are currently expensive to rent, and so is the backhaul needed to connect them to the core network.

1) Infrastructure Sharing: One possible new business model could be based on infrastructure sharing, where the owners of infrastructure and the operators are different. There are several ways in which infrastructure could be shared.

Passive Sharing: The passive elements of a network include the sites (physical space, rooftops, towers, masts and pylons), the backhaul connection, power supplies, and air-conditioning. Operators could cover larger geographical areas at a lower cost and with less power consumption if they shared sites, and this might be of particular importance in dense 5G networks [180]. Regulation could be required to force major operators to share their sites and improve competition.

Active Sharing: Active infrastructure sharing would involve antennas, BSs, radio access networks and even core networks. BS and/or radio access network sharing may be particularly attractive when rolling out small-cell networks [181]. This type of sharing could lead to collusion, with anticompetitive agreements on prices and services [180]. Regulations are required to prevent such collusion, but on the positive side are the economies of scale.

Mobile Virtual Network Operators: A small cell may be operated by a mobile virtual network operator that does not own any spectrum but has entered into an agreement with another operator to gain access to its spectrum within the small cell. The small cell may provide coverage to an enterprise or business such that, when a user leaves the enterprise, it roams onto the other operator's network.

Offloading: Roaming is traditionally used to increase coverage in scenarios when service providers' geographical reaches are limited. However, in 5G, and as discussed above, traffic may be offloaded for a different reason: spatial and temporal demand fluctuations. Such fluctuations will be greater in small-cell networks. Recent papers consider the incentive for investment under various revenue-sharing contracts [182], [183]. It is shown in [182] that sharing increases investment, and the incentive is greater if the owner of the infrastructure gets the larger fraction of the revenue when overflow traffic is carried. A bargaining approach for data offloading from a cellular network onto a collection of WiFi or femtocell networks is considered in [50] in this special issue.

2) Backhaul: A major consideration that has been considered in several places throughout the paper is backhaul, which will be more challenging to provide for hyper-dense ultra-fast networks. However, we find optimism in three directions.

- Fiber deployments worldwide continue to mature and reach farther and farther into urban corridors.
- Wireless backhaul solutions are improving by leaps and bounds, with considerable startup activity driving innovation and competition. Further, mmWave frequencies could be utilized for much of the small-cell backhauling due to their ambivalence to interference. This may in fact be the first serious deployment of non-LoS mmWave with massive beamforming gains given that the backhaul connection is quite static and outdoors-to-outdoors, and thus more amenable to precise beam alignment.
- Backhaul optimization is becoming a pressing concern, given its new status as a performance-limiting factor, and this is addressed in [184], [185] in this special issue. The problem of jointly optimizing resources in the radio network and across the backhaul is considered in [184]. Compression techniques for uplink cloud-RAN are developed in [185]. Another approach is the proactive caching of high bandwidth content like popular video [186].

V. Conclusion

It is an exciting time in the wireless industry and for wireless research at large. Daunting new requirements for 5G are already unleashing a flurry of creative thinking and a sense of urgency in bringing innovative new technologies into reality. Even just two years ago, a mmWave cellular system was considered something of a fantasy; now it is almost considered an inevitability. As this article has highlighted, it is a long road ahead to truly disruptive 5G networks. Many technical challenges remain spanning all layers of the protocol stack and their implementation, as well as many intersections with regulatory, policy, and business considerations. We hope that this article and those in this special issue will help to move us forward along this road.

Acknowledgment

The authors thank Arunabha Ghosh (AT&T Labs), Robert W. Heath, Jr. (UT Austin), and Federico Boccardi (Vodafone) for very helpful feedback and suggestions on the paper.

References

[1] B. Clerckx, A. Lozano, S. Sesia, C. van Rensburg, and C. B. Papadias, "3GPP LTE and LTE-Advanced," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, no. 1, p. 472 124, Sep. 2009.

[2] Cisco, Visual Networking Index, Feb. 2014, white paper at Cisco.com.

[3] M. Corson *et al.*, "Towards proximity-aware internetworking," *IEEE Wireless Commun. Mag.*, vol. 17, no. 6, pp. 26–33, Dec. 2010.

[4] A. Maeder, P. Rost, and D. Staehle, "The challenge of M2M communications for the cellular radio access network," in *Proc. Würzburg Workshop IP, Joint ITG Euro-NF Workshop" Vis. Future Gener. Netw. " EuroView*, Aug. 2011, pp. 1–2.

[5] Analysys Mason Report S. Hilton, Machine-to-Machine Device Connections: Worldwide Forecast 2010–2020 2010, Analysys Mason Report.

[6] FP7 European Project 317669 METIS (Mobile and Wireless Communications Enablers for the Twenty-Twenty Information Society) 2012. [Online]. Available: https://www.metis2020.com/

[7] FP7 European Project 318555 5G NOW (5th Generation Non-Orthogonal Waveforms for Asynchronous Signalling) 2012. [Online]. Available: http://www.5gnow.eu/

[8] 5G-Infrastructure Public-Private Partnership 2013. [Online]. Available: http://5g-ppp.eu/

[9] G. P. Fettweis, "The tactile Internet: Applications and challenges," *IEEE Veh. Technol. Mag.*, vol. 9, no. 1, pp. 64–70, Mar. 2014.

[10] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. Part I. System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[11] N. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[12] G. J. Foschini, K. Karakayali, and R. A. Valenzuela, "Enormous spectral efficiency of isolated multiple antenna links emerges in a coordinated cellular network," *Proc. Inst. Elect. Eng.—Commun.* , vol. 153, no. 4, pp. 548–555, Aug. 2006.

[13] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network MIMO: Overcoming intercell interference in indoor wireless systems," in *Proc. Asilomar Conf. Signals Syst. Comput.*, Nov. 2007, pp. 83–87.

[14] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degrees of freedom for the K user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[15] M. Maddah-Ali, A. Motahari, and A. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.

[16] S. Venkatesan, H. Huang, A. Lozano, and R. Valenzuela, "A WiMAX-based implementation of network MIMO for indoor wireless systems," *EURASIP J. Adv. Signal Process.*, vol. 2009, no. 1, p. 963 547, Sep. 2009.

[17] H. Huang *et al.*, "Increasing downlink cellular throughput with limited network MIMO coordination," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2983–2989, Jun. 2009.

[18] S. W. Peters and R. W. H. , Jr., "Interference alignment via alternating minimization," in *Proc. IEEE ICASSP*, 2009, pp. 2445–2448.

[19] D. Gesbert *et al.*, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[20] C. Suh, M. Ho, and D. N. C. Tse, "Downlink interference alignment," *IEEE Trans. Commun.*, vol. 59, no. 9, pp. 2616–2626, Sep. 2011.

[21] P. Wang, H. Wang, L. Ping, and X. Lin, "On the capacity of MIMO cellular systems with base station cooperation," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3720–3731, Nov. 2011.

[22] O. Simeone *et al.*, "Information theoretic considerations for wireless cellular systems: The impact of cooperation," in *Foundations and Trends in Communications and Information Theory*, vol. 7. Delft, The Netherlands: Now Publishers, 2012.

[23] A. Lozano, R. W. Heath, and J. G. Andrews, "Fundamental limits of cooperation," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5213–5226, Sep. 2013.

[24] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: A survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sep. 2008.

[25] M. Dohler, R. W. Heath, A. Lozano, C. B. Papadias, and R. A. Valenzuela, "Is the PHY layer dead?" *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 159–165, Apr. 2011.

[26] J. G. Andrews, "Seven ways that HetNets are a cellular paradigm shift," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 136–144, Mar. 2013.

[27] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: Past, present, future," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 497–508, Apr. 2012.

[28] R. W. Heath, S. W. Peters, Y. Wang, and J. Zhang, "A current perspective on distributed antenna systems for the downlink of cellular systems," *IEEE Commun. Mag.*, vol. 51, no. 4, pp. 161–167, Apr. 2013.

[29] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of K-tier downlink heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 550–560, Apr. 2012.

[30] D. Ramasamy, R. Ganti, and U. Madhow, "On the capacity of picocellular networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 241–245.

[31] H. S. Dhillon, R. K. Ganti, and J. G. Andrews, "Load-aware modeling and analysis of heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1666–1677, Apr. 2013.

[32] S. Larew, T. Thomas, and A. Ghosh, "Air interface design and ray tracing study for 5G millimeter wave communications," in *Proc. IEEE GLOBECOM Workshop*, Dec. 2013, pp. 117–122.

[33] Qualcomm, Extending LTE Advanced to Unlicensed Spectrum, Feb. 2014. [Online]. Available: http://www.slideshare.net/qualcommwirelessevolution/extending-LTE-Advanced-to-unlicensed-spectrum-31732634

[34] O. Galinina *et al.*, "Capturing spatial randomness of heterogeneous cellular/WLAN deployments with dynamic traffic," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.

[35] Q. Ye *et al.*, "User association for load balancing in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2706–2716, Jun. 2013.

[36] S. Corroy, L. Falconetti, and R. Mathar, "Cell association in small heterogeneous networks: Downlink sum rate and min rate maximization," in *Proc. IEEE Wireless Netw. Commun. Conf.*, Apr. 2012, pp. 888–892.

[37] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. S. Dhillon, "An overview of load balancing in HetNets: Old myths and open problems," *IEEE Commun. Mag.*, vol. 21, no. 2, pp. 18–25, Apr. 2014.

[38] K. Shen and W. Yu, "Distributed pricing-based user association for downlink heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.

[39] S. Singh and J. G. Andrews, "Joint resource partitioning and offloading in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 888–901, Feb. 2014.

[40] Q. Ye, M. Shalash, C. Caramanis, and J. G. Andrews, "On/off macrocells and load balancing in heterogeneous cellular networks," in *Proc. IEEE GLOBECOM*, Atlanta, GA, USA, 2013, pp. 3814–3819.

[41] D. Fooladivanda and C. Rosenberg, "Joint resource allocation and user association for heterogeneous wireless cellular networks," *IEEE Trans. Commun.*, vol. 12, no. 1, pp. 248–257, Jan. 2013.

[42] A. Bedekar and R. Agrawal, "Optimal muting and load balancing for eICIC," in *Proc. Int. Symp. Modeling Optim. Mobile, Ad Hoc WiOpt*, May 2013, pp. 280–287.

[43] S. Borst, S. Hanly, and P. Whiting, "Throughput utility optimization in HetNets," in *Proc. IEEE Veh. Technol. Conf.*, Jun. 2013, pp. 1–5.

[44] S. Borst, S. Hanly, and P. Whiting, "Optimal resource allocation in HetNets," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2013, pp. 5437–5441.

[45] S. Deb, P. Monogioudis, J. Miernik, and J. P. Seymour, "Algorithms for enhanced inter-cell interference coordination (eICIC) in LTE HetNets," *IEEE/ACM Trans. Netw.*, vol. 22, no. 1, pp. 137–150, Feb. 2014.

[46] B. Soret, K. Pedersen, T. E. Kolding, H. Kröner, and I. Maniatis, "Fast muting adaptation for LTE-A HetNets with remote radio heads," in *Proc. IEEE GLOBECOM*, Dec. 2013.

[47] S. Hanly and P. Whiting, "On the capacity of HetNets," in *Proc. ITA*, Feb. 2014, pp. 1–9.

[48] D. Bethanabhotla, O. Bursalioglu, H. C. Papadopoulos, and G. Caire, "User association and load balancing for cellular massive MIMO," in *Proc. ITA*, Feb. 2014, pp. 1–10.

[49] E. Aryafar, M. W. A. Keshavarz-Haddad, and M. Chiang, "RAT selection games in HetNets," in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 998–1006.

[50] L. Gao, G. Iosifidis, J. Huang, L. Tassiulas, and D. Li, "Bargaining-based mobile data offloading," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.

[51] J. Lee *et al.*, "Coordinated multipoint transmission and reception in LTE-Advanced systems," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 44–50, Nov. 2012.

[52] J. Kim, H.-W. Lee, and S. Chong, "Virtual cell beamforming in cooperative networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.

[53] O. Aliu, A. Imran, M. Imran, and B. Evans, "A survey of self organisation in future cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 336–361, Jan. 2013.

[54] B. Kaufman and B. Aazhang, "Cellular networks with an overlaid device to device network," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Oct. 2008, pp. 1537–1541.

[55] C. Yu, K. Doppler, C. B. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for device-to-device communication underlaying cellular networks," *IEEE Wireless Commun.*, vol. 10, no. 8, pp. 2752–2763, Aug. 2011.

[56] "Special issue on Internet of Things," *IEEE Wireless Commun. Mag.*, no. 6, Dec. 2010.

[57] N. NaderiAlizadeh and S. Avestimehr, "ITLinQ: A new approach for spectrum sharing in device-to-device communication systems," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.

[58] S. Singh, M. Kulkarni, and J. G. Andrews, "A tractable model for rate in noise limited mmwave cellular networks," in *Proc. IEEE Asilomar*, Nov. 2014.

[59] Connecting America: The National Broadband Plan (Chapter 5) 2010. [Online]. Available: http://download.broadband.gov/plan/national-broadband-plan.pdf

[60] "Estimated spectrum bandwidth requirements for the future development of IMT-2000 and IMT-Advanced," ITU-R, Geneva, Switzerland, Rep. M.2078, 2006.

[61] R. C. Daniels and R. W. Heath, "60 GHz wireless communications: Emerging requirements and design recommendations," *IEEE Veh. Technol. Mag.*, vol. 2, no. 3, pp. 41–50, Sep. 2007.

[62] S. J. Vaughan-Nicholos, "Gigabit Wi-Fi is on its way," *IEEE Comput.*, vol. 43, no. 11, pp. 11–14, Nov. 2010.

[63] T. Baykas, C.-S. Sum, Z. Lan, and J. Wang, "IEEE 802.15.3c: The first IEEE wireless standard for data rates over 1 Gb/s," *IEEE Commun. Mag.*, vol. 49, no. 7, pp. 114–121, Jul. 2011.

[64] T. S. Rappaport, J. N. Murdock, and F. Gutierrez, "State of the art in 60-GHz integrated circuits and systems for wireless communications," *Proc. IEEE*, vol. 99, no. 8, pp. 1390–1436, Aug. 2011.
[65] M. Marcus and B. Pattan, "Millimeter wave propagation; Spectrum management implications," *IEEE Microw. Mag.*, vol. 6, no. 2, pp. 54–62, Jun. 2005.
[66] A. Alejos, M. G. Sanchez, and I. Cuinas, "Measurement and analysis of propagation mechanisms at 40 GHz: Viability of site shielding forced by obstacles," *IEEE Trans. Veh. Technol.*, vol. 57, no. 6, pp. 3369–3380, Nov. 2008.
[67] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.
[68] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, no. 1, pp. 335–349, Aug. 2013.
[69] W. Roh *et al.*, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, Feb. 2014.
[70] S. Rangan, T. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.
[71] T. S. Rappaport, "Broadband millimeter-wave propagation measurements and models using adaptive-beam antennas for outdoor urban cellular communications," *IEEE Trans. Antennas Propag.*, vol. 61, no. 4, pp. 1850–1859, Apr. 2013.
[72] M. Kulkarni, S. Singh, and J. G. Andrews, "Coverage and rate trends in dense urban mmwave cellular networks," in *Proc. IEEE Globecom*, Dec. 2014.
[73] T. Bai, R. Vaze, and R. W. Heath, "Analysis of blockage effects on urban cellular networks," *IEEE Trans. Wireless Commun.*, 2013, arXiv preprint arXiv:1309.4141, to be published.
[74] T. Bai and R. W. Heath, Jr, "Coverage analysis for millimeter wave cellular networks with blockage effects," in *Proc. IEEE Global Signal Inf. Process. Conf.*, Dec. 2013, pp. 727–730.
[75] T. Bai, R. Vaze, and R. W. Heath, Jr., "Using random shape theory to model blockage in random cellular networks," in *Proc. Int. Conf. SPCOM*, Jul. 2012, pp. 1–5.
[76] A. Ghosh *et al.*, "Millimeter wave enhanced local area systems: A high data rate approach for future wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.
[77] M. R. Akdeniz *et al.*, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.
[78] H. Ishii, Y. Kishiyama, and H. Takahashi, "Novel architecture for LTE-B: C-plane/U-plane split and phantom cell concept," in *Proc. IEEE GLOBECOM Workshop*, Dec. 2012, pp. 624–630.
[79] Q. Li, H. Niu, G. Wu, and R. Q. Hu, "Anchor-booster based heterogeneous networks with mmWave capable booster cells," in *Proc. IEEE GLOBECOM*, Dec. 2013, pp. 93–98.
[80] T. Bai and R. W. Heath, "Analysis of millimeter wave cellular networks with overlaid microwave base stations," in *Proc. ASILOMAR Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2014.
[81] X. Zhang, A. Molish, and S. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4091–4103, Nov. 2005.
[82] J. Wang *et al.*, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1390–1399, Oct. 2009.
[83] V. Venkateswaran and A.-J. van der Veen, "Analog beamforming in MIMO communications with phase shift networks and online channel estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4131–4143, Aug. 2010.
[84] S. Hur *et al.*, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.
[85] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
[86] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, arXiv preprint arXiv:1401.7426, to be published.
[87] G. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.
[88] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.
[89] F. R. Farrokhi, A. Lozano, G. J. Foschini, and R. A. Valenzuela, "Spectral efficiency of FDMA/TDMA wireless systems with transmit and receive antenna arrays," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 591–599, Jan. 2002.
[90] A. Lozano and A. M. Tulino, "Capacity of multiple-transmit multiple-receive antenna architectures," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3117–3128, Dec. 2002.
[91] T. L. Marzetta, "The case for MANY (greater than 16) antennas as the base station," in *Proc. ITA*, San Diego, CA, USA, Jan. 2007.
[92] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Sep. 2010.
[93] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
[94] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
[95] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
[96] M. Matthaiou, G. K. Karagiannidis, E. G. Larsson, T. L. Marzetta, and R. Schober, "Large-scale multiple antenna wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 113–116, Feb. 2013.
[97] Z. Xiang, M. Tao, and X. Wang, "Massive MIMO multicasting in noncooperative cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.
[98] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.
[99] A. Lozano, "Long-term transmit beamforming for wireless multicasting," in *Proc. IEEE ICASSP*, 2007, vol. 3, pp. III-417–III-420.
[100] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
[101] N. Jindal and A. Lozano, "A unified treatment of optimum pilot overhead in multipath fading channels," *IEEE Trans. Commun.*, vol. 58, no. 10, pp. 2939–2948, Oct. 2010.
[102] H. Huh, A. M. Tulino, and G. Caire, "Network MIMO with linear zero-forcing beamforming: Large system analysis, impact of channel estimation, reduced-complexity scheduling," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2911–2934, May 2012.
[103] A. Ashikhmin and T. L. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2012, pp. 1137–1141.
[104] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.
[105] Y. Zeng, R. Zhang, and Z. N. Chen, "Electromagnetic lens-focusing antenna enabled massive MIMO: Performance improvement and cost reduction," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.
[106] Y.-H. Nam *et al.*, "Full-dimension MIMO (FD-MIMO) for next generation cellular technology," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 172–179, Jun. 2013.
[107] X. Lu, A. Tolli, O. Piirainen, M. Juntti, and W. Li, "Comparison of antenna arrays in a 3-D multiuser multicell network," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2011, pp. 1–6.
[108] S. Wu, C. Wang, H. Aggoune, M. Alwakeel, and Y. He, "A non-stationary 3D wideband twin-cluster model for 5G massive MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.
[109] *Study on 3D-Channel Model for Elevation Beamforming and FD-MIMO Studies for LTE*, 3GPP TSG RAN Plenary No. 58, Dec. 2012.
[110] A. Kammoun, H. Khanfir, Z. Altman, M. Debbah, and M. Kamoun, "Preliminary results on 3D channel modeling: From theory to standardization," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.
[111] A. Adhikary, E. A. Safadi, and G. Caire, "Massive MIMO and inter-tier interference coordination," in *Proc. ITA*, Feb. 2014, pp. 1–10.
[112] J. Chen and V. Lau, "Two-tier precoding for FDD multi-cell massive MIMO time-varying interference networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.
[113] A. Adhikary *et al.*, "Joint spatial division and multiplexing for mm-wave channels," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.
[114] K. S. Gilhousen *et al.*, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Technol.*, vol. 40, no. 2, pp. 303–312, May 1991.

[115] P. Bender *et al.*, "CDMA/HDR: A bandwidth efficient high speed wireless data service for nomadic users," *IEEE Commun. Mag.*, vol. 38, no. 7, pp. 70–77, Jul. 2000.
[116] J. Li, X. Wu, and R. Laroia, *OFDMA Mobile Broadband Communications: A Systems Approach*. Cambridge, U.K.: Cambridge Univ. Press, 2013.
[117] C. E. Shannon, "A mathematical theory of communication (Part 1)," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, Jul.–Oct. 1948.
[118] A. Ghosh, J. Zhang, J. G. Andrews, and R. Muhamed, *Fundamentals of LTE*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2010.
[119] S. Hong *et al.*, "Frequency and quadrature-amplitude modulation for downlink cellular OFDMA networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.
[120] A. Sahin, I. Guvenc, and H. Arslan, "A survey on multicarrier communications: Prototype filters, lattice structures, implementation aspects," *IEEE Commun. Surveys Tuts.*, to be published.
[121] A. Barbieri, D. Fertonani, and G. Colavolpe, "Time–frequency packing for linear modulations: Spectral efficiency and practical detection schemes," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 2951–2959, Oct. 2009.
[122] F. Rusek and J. Anderson, "Multistream faster than Nyquist signaling," *IEEE Trans. Commun.*, vol. 57, no. 5, pp. 1329–1340, May 2009.
[123] J. Anderson, F. Rusek, and V. Owall, "Faster-than-Nyquist signaling," *Proc. IEEE*, vol. 101, no. 8, pp. 1817–1830, Aug. 2013.
[124] D. Dasalukunte, F. Rusek, and V. Owall, "Multicarrier faster-than-Nyquist transceivers: Hardware architecture and performance analysis," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 58, no. 4, pp. 827–838, Apr. 2011.
[125] B. Farhang-Boroujeny, "OFDM versus filter bank multicarrier," *IEEE Signal Process. Mag.*, vol. 28, no. 3, pp. 92–112, May 2011.
[126] V. Vakilian, T. Wild, F. Schaich, S. ten Brink, and J.-F. Frigon, "Universal-filtered multi-carrier technique for wireless systems beyond LTE," in *Proc. IEEE GLOBECOM Broadband Wireless Access Workshop*, Dec. 2013, pp. 223–228.
[127] G. Wunder *et al.*, "5GNOW: Non-orthogonal, asynchronous waveforms for future mobile applications," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 97–105, Feb. 2014.
[128] N. Cassiau, D. Ktenas, and J.-B. Dore, "Time and frequency synchronization for CoMP with FBMC," in *Proc. 10th Int. Symp. Wireless Commun. Syst.*, Aug. 2013, pp. 1–5.
[129] G. Fettweis, M. Krondorf, and S. Bittner, "GFDM—Generalized frequency division multiplexing," in *Proc. IEEE Veh. Technol. Conf.*, Apr. 2009, pp. 1–4.
[130] R. Datta, D. Panaitopol, and G. Fettweis, "Analysis of cyclostationary GFDM signal properties in flexible cognitive radio," in *Proc. Int. Symp. Commun. Inf. Technol.*, Oct. 2012, pp. 663–667.
[131] N. Benvenuto, R. Dinis, D. Falconer, and S. Tomasin, "Single carrier modulation with nonlinear frequency domain equalization: An idea whose time has come—Again," *Proc. IEEE*, vol. 98, no. 1, pp. 69–96, Jan. 2010.
[132] N. Lee, R. W. Heath, Jr., D. Morales-Jimenez, and A. Lozano, "Base station cooperation with dynamic clustering in super-dense cloud-RAN," in *Proc. IEEE GLOBECOM*, Dec. 2013, pp. 784–788.
[133] Z. Zhu *et al.*, "Virtual base station pool: Towards a wireless network cloud for radio access networks," in *Proc. ACM Int. Conf. Comput. Frontiers*, 2011, p. 34.
[134] Network Function Virtualisation—An Introduction, Benefit, Enablers, Challenges and Call for Action, white paper, Oct. 2012. [Online]. Available: http://portal.etsi.org/NFV/NFV_White_Paper.pdf
[135] Network Function Virtualisation—Network Operator Perspectives on Industry Progress, white paper, Oct. 2013. [Online]. Available: http://portal.etsi.org/NFV/NFV_White_Paper2.pdf
[136] AT&T Domain 2.0 Vision, white paper, Nov. 2013. [Online]. Available: http://att.com
[137] Software-Defined Networking, the New Norm for Networks, Apr. 2012. [Online]. Available: http://www.opennetworking.org
[138] S. Sezer *et al.*, "Are we ready for SDN? Implementation challenges for software-defined networks," *IEEE Commun. Mag.*, vol. 51, no. 7, pp. 36–43, Jul. 2013.
[139] [Online]. Available: http://www.greentouch.org
[140] "Special issue on energy-efficient wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, Sep. 2011.
[141] "Special issue on energy efficiency in communications," *IEEE Commun. Mag.*, vol. 48, no. 11, Nov. 2010.
[142] G. Auer *et al.*, "How much energy is needed to run a wireless network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40–49, Oct. 2011.
[143] C. Saraydar, N. B. Mandayam, and D. Goodman, "Efficient power control via pricing in wireless data networks," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 291–303, Feb. 2002.
[144] F. Meshkati, H. Poor, and S. Schwartz, "Energy-efficient resource allocation in wireless networks," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 58–68, May 2007.
[145] S. Buzzi and H. Poor, "Joint receiver and transmitter optimization for energy-efficient CDMA communications," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 3, pp. 459–472, Apr. 2008.
[146] S. Buzzi, G. Colavolpe, D. Saturnino, and A. Zappone, "Potential games for energy-efficient power control and subcarrier allocation in uplink multicell OFDMA systems," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 2, pp. 89–103, Apr. 2012.
[147] G. Miao, N. Himayat, Y. Li, and D. Bormann, "Energy efficient design in wireless OFDMA," in *Proc. IEEE Int. Conf. Commun.*, May 2008, pp. 3307–3312.
[148] B. L. Ng *et al.*, "Fulfilling the promise of massive MIMO with 2D active antenna array," in *Proc. IEEE GLOBECOM Workshop*, Dec. 2012, pp. 691–696.
[149] L. Venturino, A. Zappone, C. Risi, and S. Buzzi, "Energy-efficient scheduling and power allocation in downlink OFDMA networks with base station coordination," *IEEE Trans. Wireless Commun.*, available on arXiv:1405.2962, to be published.
[150] S. He, Y. Huang, H. Wang, S. Jin, and L. Yang, "Leakage-aware energy-efficient beamforming for heterogeneous multicell multiuser systems," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.
[151] G. Koutitas, "Low carbon network planning," in *Proc. Eur. Wireless Conf.*, Apr. 2010, pp. 411–417.
[152] E. Oh and B. Krishnamachari, "Energy savings through dynamic base station switching in cellular wireless access networks," in *Proc. IEEE GLOBECOM*, Dec. 2010, pp. 1–5.
[153] S. Zhou, J. Gong, Z. Yang, Z. Niu, and P. Yang, "Green mobile access network with dynamic base station energy saving," in *Proc. ACM MobiCom*, 2009, vol. 9, no. 262, pp. 10–12.
[154] Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 74–79, Nov. 2010.
[155] A. Conte *et al.*, "Cell wilting and blossoming for energy efficiency," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 50–57, Oct. 2011.
[156] M. Marsan, G. Bucalo, A. Di Caro, M. Meo, and Y. Zhang, "Towards zero grid electricity networking: Powering BSs with renewable energy sources," in *Proc. IEEE Int. Conf. Commun. Workshop*, Jun. 2013, pp. 596–601.
[157] H. S. Dhillon, Y. Li, P. Nuggehalli, Z. Pi, and J. G. Andrews, "Fundamentals of heterogeneous cellular networks with energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2782–2797, May 2014.
[158] D. Ng, E. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with hybrid energy harvesting base station," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3412–3427, Jul. 2013.
[159] P. Skillermark and P. Frenger, "Enhancing energy efficiency in LTE with antenna muting," in *Proc. IEEE Veh. Technol. Conf.*, May 2012, pp. 1–5.
[160] C.-L. I, C. Rowell, S. Han, Z. Xu, G. Li, and Z. Pan, "Toward green and soft: A 5G perspective," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 66–73, Feb. 2014.
[161] L. Liu *et al.*, "GreenCloud: A new architecture for green data center," in *Proc. Int. Conf. Ind. Session Auton. Comput. Commun. Ind. Session*, New York, NY, USA, 2009, pp. 29–38.
[162] C. Li, J. Zhang, and K. B. Letaief, "Throughput and energy efficiency analysis of small cell networks with multi-antenna base stations," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2505–2517, May 2014.
[163] S. Rangan, "Energy efficient methods for energy efficient methods for millimeter wave picocellular systems," in *Proc. IEEE Commun. Theory Workshop*, Phuket, Thailand, Jun. 2013, pp. 1–25.
[164] R. Coase, "The Federal Communications Commission," *J. Law Econ.*, vol. 2, no. 1, pp. 1–40, Oct. 1959.
[165] Y. Benkler, "Open wireless vs. licensed spectrum: Evidence from market adoption," *Harvard J. Law Technol.*, vol. 26, no. 1, pp. 71–163, Oct. 2012.
[166] M. Matinmikko *et al.*, "Cognitive radio trial environment: First live authorized shared access-based spectrum sharing demonstration," *IEEE Veh. Technol. Mag.*, vol. 8, no. 3, pp. 30–37, Sep. 2013.
[167] J. Khun-Jush, P. Bender, B. Deschamps, and M. Gundlach, "Licensed shared access as complementary approach to meet spectrum demands: Benefits for next generation cellular systems," in *Proc. ETSI Workshop Reconfigurable Radio Syst.*, Cannes, France, Dec. 2013, pp. 1–7.

[168] H. Zhou, R. Berry, M. Honig, and R. Vohra, "Complexity of allocation problems in spectrum markets with interference complementarities," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 489–499, Mar. 2013.

[169] R. Berry, M. Honig, and R. Vohra, "Spectrum markets: Motivation, challenges, implications," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 146–155, Nov. 2010.

[170] T. Nguyen, H. Zhou, R. Berry, M. Honig, and R. Vohra, "The impact of additional unlicensed spectrum on wireless services competition," in *Proc. IEEE Int. Symp. New Frontiers DYSPAN*, May 2011, pp. 146–155.

[171] H. Zhou, R. Berry, M. Honig, and R. Vohra, "Investment and competition in unlicensed spectrum," in *Proc. 46th CISS*, Mar. 2012, pp. 1–6.

[172] "IMT vision towards 2020 and beyond," presented at the ITU-R 2020 Vision Workshop, Ho Chi Minh, Vietnam, Feb., 2014.

[173] "Views on IMT beyond 2020," presented at the ITU-R 2020 Vision Workshop ARIB 2020 Beyond Ad Hoc, Ho Chi Minh, Vietnam, Feb. 2014.

[174] "5G vision and requirements of 5G forum, Korea," presented at the ITU-R 2020 Vision Workshop 5G Forum Korea, Ho Chi Minh, Vietnam, Feb. 2014.

[175] *Scenarios, Requirements and KPIs for 5G Mobile and Wireless System*, METIS ICT-317669-METIS/D1.1, Apr. 2013.

[176] *Summary on Preliminary Trade-Off Investigations and First Set of Potential Network-Level Solutions*, METIS ICT-317669-METIS/D4.1, Sep. 2013.

[177] *Work Plan, Timeline, Process and Deliverables for "IMT-2020" Development*, ITU-R WP5D, Feb. 2014.

[178] A. Scrase, Summit Key Messages, Nov. 2013. [Online]. Available: http://docbox.etsi.org/Workshop/2013/201311_FUTUREMOBILESUMMIT/ZZ_CONCLUSIONS.pdf

[179] *Working Document Towards a Preliminary Draft: New Report ITU-R M.[IMT.above 6 GHz]*, ITU-R WP5D, Feb. 2014.

[180] D.-E. Meddour, T. Rasheed, and Y. Gourhant, "On the role of infrastructure sharing for mobile network operators in emerging markets," *Comput. Netw.*, vol. 55, no. 7, pp. 1576–1591, May 2011.

[181] A. Ghanbari, J. Markendahl, and A. Widaa, "Cooperation patterns in small cell networks—Risks and opportunities to distinguish the win-win model," in *Proc. 24th Eur. Regional ITS Conf.*, Oct. 2013, pp. 1–24.

[182] R. Berry *et al.*, "On the nature of revenue-sharing contracts to incentivize spectrum-sharing," in *Proc. IEEE INFOCOM*, 2013, pp. 845–853.

[183] B. Leng, P. Mansourifard, and B. Krishnamachari, "Microeconomic analysis of base-station sharing in green cellular networks," in *Proc. IEEE INFOCOM*, Toronto, ON, Canada, 2014.

[184] W.-C. Liao *et al.*, "Min flow rate maximization for software defined radio access networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.

[185] Y. Zhou and W. Yu, "Optimized backhaul compression for uplink cloud radio access network," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014.

[186] N. Golrezaei, K. Shanmugam, A. G. Dimakis, A. F. Molisch, and G. Caire, "Femtocaching: Wireless video content delivery through distributed caching helpers," in *Proc. IEEE INFOCOM*, Mar. 2011, pp. 1107–1115.

Jeffrey G. Andrews (S'98–M'02–SM'06–F'13) received the B.S. degree in engineering (with high distinction) from Harvey Mudd College, Claremont, CA, USA, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA. He is the Cullen Trust Endowed Professor (#1) of Department Of Electrical and Computer Engineering (ECE) at the University of Texas at Austin, Austin, TX, USA, Editor-in-Chief of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and Technical Program Co-Chair of IEEE Globecom 2014. He developed CDMA systems at Qualcomm from 1995 to 1997 and has consulted for entities, including Verizon, the WiMAX Forum, Intel, Microsoft, Apple, Samsung, Clearwire, Sprint, and NASA. He is a member of the Technical Advisory Boards of Accelera and Fastback Networks and the coauthor of the books *Fundamentals of WiMAX* and *Fundamentals of LTE*.

Dr. Andrews received the National Science Foundation CAREER award in 2007 and has been the coauthor of ten Best Paper Award recipients: ICC 2013, Globecom 2006 and 2009, Asilomar 2008, European Wireless 2014, the 2010 IEEE Communications Society Best Tutorial Paper Award, the 2011 IEEE Heinrich Hertz Prize, the 2014 EURASIP Best Paper Award, the 2014 IEEE Stephen O. Rice Prize, and the 2014 IEEE Leonard G. Abraham Prize. He is an elected member of the Board of Governors of the IEEE Information Theory Society.

Stefano Buzzi (M'98–SM'07) received the Ph.D. degree in electronic engineering and computer science from the University of Naples "Federico II," Naples, Italy, in 1999. He has had short-term visiting appointments at the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA, in 1999, 2000, 2001, and 2006. He is currently an Associate Professor with the University of Cassino and Southern Latium, Cassino (FR), Italy. He is the author or coauthor of more than 50 journal papers and 90 conference papers. His research and study interests are in the wide area of statistical signal processing and resource allocation for communications, with emphasis on wireless communications. Dr. Buzzi was awarded by the Associazione Elettrotecnica ed Elettronica Italiana (AEI) the "G. Oglietti" scholarship in 1996, and was the recipient of a NATO/CNR Advanced Fellowship in 1999 and of three CNR short-term mobility grants. He is a former Associate Editor of the IEEE COMMUNICATIONS LETTERS and the IEEE SIGNAL PROCESSING LETTERS.

Wan Choi (M'06–SM'12) received the B.Sc. and M.Sc. degrees from Seoul National University, Seoul, Korea, in 1996 and 1998, respectively, and the Ph.D. degree from the University of Texas at Austin, Austin, TX, USA, in 2006. He is currently an Associate Professor with the Department of Electrical Engineering, Korea Advance Institute of Science and Technology, Daejeon, Korea. From 1998 to 2003, he was a Senior Member of the Technical Staff of the R&D Division, KT Freetel, Korea, where he researched 3G CDMA systems.

Dr. Choi was the recipient of IEEE Vehicular Technology Society Jack Neubauer Memorial Award in 2002. He also received the IEEE Vehicular Technology Society Dan Noble Fellowship Award in 2006 and the IEEE Communication Society Asia Pacific Young Researcher Award in 2007. He serves as an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and IEEE WIRELESS COMMUNICATIONS LETTERS.

Stephen V. Hanly (M'98) received the B.Sc.(Hons.) and M.Sc. degrees from the University of Western Australia, Perth, Australia, and the Ph.D. degree in mathematics from Cambridge University, Cambridge, U.K., in 1994. From 1993 to 1995, he was a Postdoctoral Member of Technical Staff at AT&T Bell Laboratories. From 1996 to 2009, he was with the University of Melbourne, Melbourne, Australia, and from 2010 to 2011, he was with the National University of Singapore, Singapore. He now holds the CSIRO-Macquarie University Chair in Wireless Communications at Macquarie University, Sydney, Australia. He has been an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, and a Guest Editor of the *EURASIP Journal on Wireless Communications and Networking*. In 2005, he was the Technical Co-Chair for the IEEE International Symposium on Information Theory held in Adelaide, Australia. He was a corecipient of the Best Paper Award at the IEEE Infocom 1998 Conference and the 2001 Joint IEEE Communications Society and IEEE Information Theory Society Best Paper Award.

Angel Lozano (S'90–M'99–SM'01–F'14) received the M.Sc. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 1994 and 1998, respectively. He is a Professor and the Vice-Rector for Research at the Universitat Pompeu Fabra, Barcelona, Spain. He was with Bell Labs (Lucent Technologies, now Alcatel-Lucent) between 1999 and 2008 and served as an Adjunct Associate Professor at Columbia University, New York, NY, USA, between 2005 and 2008. Prof. Lozano is an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY, the Chair of the IEEE Communication Theory Technical Committee, and an elected member to the Board of Governors of the IEEE Communications Society. His papers have received two awards: ISSSTA 2006 and the 2009 IEEE Stephen O. Rice Prize.

Anthony C. K. Soong (S'88–M'91–SM'02–F'14) received the B.Sc. degree in animal physiology and physics from the University of Calgary, Calgary, AB, Canada, and the B.Sc. degree in electrical engineering, the M.Sc. degree in biomedical physics, and the Ph.D. degree in electrical and computer engineering from the University of Alberta, Edmonton, AB, Canada. He is currently the Chief Scientist for wireless research and standards at Huawei Technologies Co. Ltd., Plano, TX, USA. He served as the Chair for 3GPP2 TSG-C NTAH (the next generation radio access network technology development group) from 2007 to 2009 and the Vice Chair for 3GPP2 TSG-C WG3 (the physical layer development group for CDMA 2000) from 2006 to 2011. Prior to joining Huawei, he was with the systems group for Ericsson, Inc., and Qualcomm, Inc. Dr. Soong has published numerous scientific papers and has over 80 patents granted or pending. He was the corecipient of the 2013 IEEE Signal Processing Society Best Paper Award.

Jianzhong Charlie Zhang (S'96–M'02–SM'09) received the Ph.D. degree from the University of Wisconsin-Madison, Madison, WI, USA. He is currently the Senior Director and Head of Wireless Communications Lab, Samsung Research America, Dallas, TX, USA, where he leads technology development, prototyping, and standardization for beyond 4G and 5G wireless systems. From August 2009 to August 2013, he served as the Vice Chairman of the 3GPP RAN1 working group and led development of LTE and LTE-Advanced technologies such as 3D channel modeling, UL-MIMO and CoMP, carrier aggregation for TD-LTE, etc. Before joining Samsung, he was with Motorola from 2006 to 2007, working on 3GPP HSPA standards, and with Nokia Research Center from 2001 to 2006, working on IEEE 802.16e (WiMAX) standard and EDGE/CDMA receiver algorithms.