AN OVERVIEW OF SUSTAINABLE GREEN 5G NETWORKS

QINGQING WU, GEOFFREY YE LI, WEN CHEN,
DERRICK WING KWAN NG, AND ROBERT SCHOBER

ABSTRACT

The stringent requirements of a 1000× increase in data traffic and 1 ms round-trip latency have made limiting the potentially tremendous ensuing energy consumption one of the most challenging problems for the design of the upcoming 5G networks. To enable sustainable 5G networks, new technologies have been proposed to improve the system energy efficiency, and alternative energy sources are introduced to reduce our dependence on traditional fossil fuels. In particular, various 5G techniques target the reduction of the energy consumption without sacrificing the quality of service. Meanwhile, energy harvesting technologies, which enable communication transceivers to harvest energy from various renewable resources and ambient radio frequency signals for communication, have drawn significant interest from both academia and industry. In this article, we provide an overview of the latest research on both green 5G techniques and energy harvesting for communication. In addition, some technical challenges and potential research topics for realizing sustainable green 5G networks are also identified.

INTRODUCTION

Fifth generation (5G) wireless networks will support up to a 1000-fold increase in capacity compared to existing networks. It is anticipated to connect at least 100 billion devices worldwide with approximately 7.6 billion mobile subscribers due to the tremendous popularity of smartphones, electronic tablets, sensors, and so on, and provide an up to 10 Gb/s individual user experience [1]. Along with the dramatic traffic explosion and device proliferation, 5G wireless networks also have to integrate human-to-machine and machine-to-machine communications in order to facilitate more flexible networked social information sharing aiming for one million connections per square kilometer. Consequently, sensors, accessories, and tools are expected to become wireless communication entities exchanging information, giving rise to the well-known Internet of Things (IoT). Furthermore, the latency on the order of 20 ms of the incumbent mobile networks is insufficient to fulfill the anticipated Tactile Internet applications in 5G networks, such as real-time remote control and physical tactile experience. With such tremendously expanding demand for wireless communications in the future, researchers are currently looking for viable solutions to meet the stringent throughput requirement. In this regard, three paradigms have emerged:
• Exploit unused and unlicensed spectrum: millimeter-wave (mmWave) communications and Long Term Evolution (LTE) in unlicensed spectrum (LTE-U).
• Reduce the transmitter-receiver (Tx-Rx) distance and improve frequency reuse: ultra-dense networks (UDNs) and device-to-device (D2D) communications.
• Enhance spectral efficiency (SE) by deploying a massive amount of antennas: massive multiple-input multiple-output (M-MIMO).

The technologies listed above increase the system throughput from three different angles. However, the performance gains introduced by these technologies do not come for free. For example, M-MIMO exploits a large number of antennas for potential multiplexing and diversity gains at the expense of an escalating circuit power consumption in the RF chains that scales linearly with the number of antennas. In addition, power-hungry transceiver chains and complex signal processing techniques are needed for reliable mmWave communications to operate at extremely high frequencies. In light of this, the network energy consumption may no longer be sustainable, and designing energy-efficient 5G networks is critical and necessary. In fact, in recent years, energy consumption has become a primary concern in the design and operation of wireless communication systems motivated by the desire to lower the operating cost of the base stations (BSs), prolong the lifetime of the user terminals (UTs), and also protect the environment. As a result, energy efficiency (EE), measured in bits per Joule, has emerged as a new prominent figure of merit and has become the most widely adopted green design metric for wireless communication systems [2].

Meanwhile, it should also be noted that EE cannot be improved infinitely by only applying spectrally efficient communication technologies,

Qingqing Wu is with Shanghai Jiao Tong University and Georgia Institute of Technology.

Geoffrey Ye Li is with Georgia Institute of Technology.

Wen Chen is with Shanghai Jiao Tong University.

Derrick Wing Kwan Ng is with the University of New South Wales.

Robert Schober is with the Institute for Digital Communications, Friedrich-Alexander University Erlangen-Nürnberg.

due to the constraint imposed by Shannon capacity bound as well as the non-negligible circuit power consumption. Consequently, even with the aforementioned 5G paradigms for improving EE, the power consumption will still grow because of the explosive data rate requirements in the future. Therefore, improving EE can only alleviate the power consumption problem to a certain extent and is insufficient for enabling sustainable 5G communications. Hence, energy harvesting technologies, which allow BSs and devices to harvest energy from renewable resources (solar, wind, etc.) and even RF signals (television signals, interference signals, etc.), have received significant attention recently. They provide green energy supply solutions for powering different components of wireless communication networks. Therefore, integrating energy harvesting technologies into future wireless networks is imperative [3]. Gaining a better understanding of the specific challenges of energy harvesting technologies and alternative energy sources will pave the way to incorporating them smoothly into the upcoming 5G wireless network planning, design, and deployment. This article focuses on the state of the art of green 5G and energy harvesting technologies, which will create an ecosystem of 5G wireless networks as shown in Fig. 1.

The rest of this article is organized as follows. After briefly introducing the trade-off between EE and SE, we discuss green 5G technologies from the perspectives of enlarging spectrum availability, shortening Tx-Rx distances, and enhancing the spatial degrees of freedom, respectively. Then energy harvesting technologies for 5G systems are discussed, and conclusions are drawn.

GENERAL EE-SE TRADE-OFF

In general, the system achievable throughput or SE can be expressed as

$$SE = K \times B \times N \times \log_2(1 + SINR(d)), \quad (1)$$

where K, B, N, and d are the frequency reuse factor, the signal bandwidth, the number of spatial beams (spatial multiplexing factor), and a single link distance, respectively. The signal-to-interference-plus-noise ratio (SINR) at the receiver increases with decreasing d. Correspondingly, the system EE can be expressed as

$$EE = \frac{K \times B \times N \times \log_2\big(1 + SINR(d)\big)}{P_t + P_c}, \quad (2)$$

where P_t and P_c are the consumed transmit and circuit powers, respectively. From Eq. 1, the SE of wireless networks can be improved by increasing the frequency reuse factor (reducing the Tx-Rx distance), the signal bandwidth, and/or the number of spatial beams, which are discussed in the subsequent three sections, respectively. Table 1 compares various performance aspects of several 5G technologies. From the table, most of the 5G technologies enable the reduction of the transmit power at the expense of incurring additional circuit power consumption due to the required hardware expansion, sophisticated signal processing, and so on.

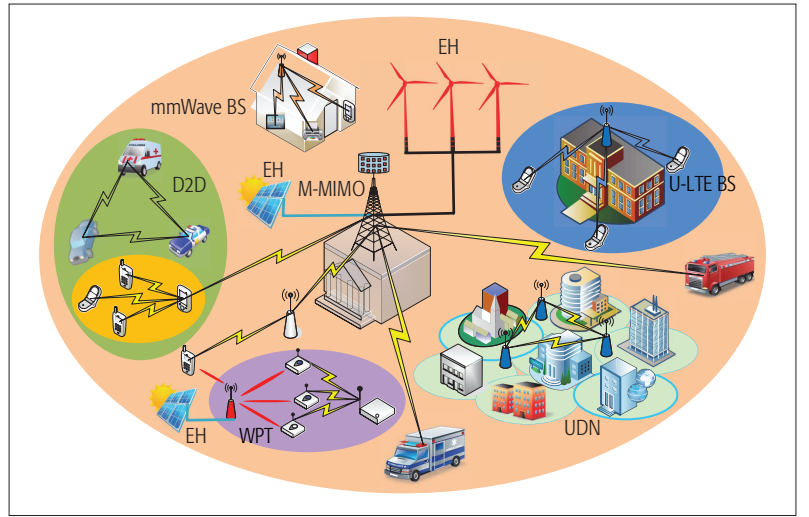For point-to-point link-level communications, where K, B, N, and d in Eqs. 1 and 2 are fixed, the

FIGURE 1. An ecosystem of disruptive technologies for 5G wireless networks.

Technology	High EE at	Coverage	Transmit power	Circuit power	Challenges
mmWave	BS and UT	200 m	Low	High	Transceiver design
LTE-U	BS and UT	500 m	Moderate	Moderate	Coexistence design
UDNs	UT	10–200 m	Low	High	Interference management, backhaul, user association, on-off switching
D2D	BS	2–100 m	Low	Low	Interference management, cooperative design
M-MIMO	UT	1000 m	Low	High	Signal processing, CSI, antenna cost

TABLE 1. Comparison of green 5G technologies.

relation between EE and SE can be analyzed using the approach in [2], which is illustrated in Fig. 2. From the figure, one can observe that:
- The EE decreases monotonically with SE if the circuit power consumption is ignored (i.e., $P_c = 0$).
- The EE increases with increasing SE when the SE is below a threshold and decreases with increasing SE when the SE is beyond the threshold if the circuit power consumption is considered (i.e., $P_c > 0$).
- The EE eventually converges to that of $P_c = 0$ as the SE increases, due to the domination of the transmit power.
- Reducing the circuit power will enlarge the EE-SE trade-off region.

Although observed from a single communication link, the fundamental trade-off between EE and SE in Fig. 2 carries over to more complicated systems employing the aforementioned 5G technologies. Besides, the EE and SE trade-off can be achieved via different system configurations such as spectrum management, frequency reuse, spatial multiplexing, and power allocation. In addition, energy harvesting technologies can provide green energy, which allows 5G networks to possibly operate at higher SEs compared to the conventional energy limited networks.

MmWave communications exploit the enormous chunks of spectrum available at mmWave frequencies and are expected to push the mobile data rates to gigabits-per-second for 5G wireless networks. However, the propagation conditions at the extremely high frequencies of the mmWave bands lead to several fundamental issues.
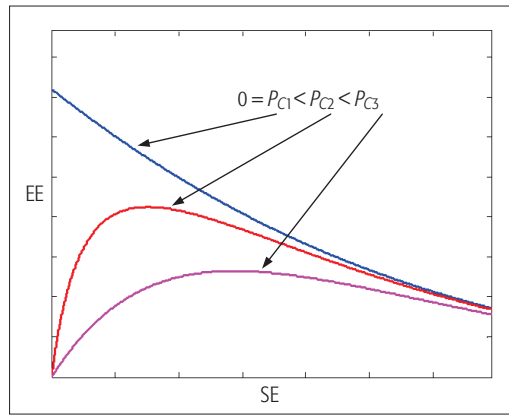

FIGURE 2. Fundamental trade-off between EE and SE.

GREEN 5G TECHNOLOGIES: ENLARGING SPECTRUM AVAILABILITY

Given the extreme shortage of available spectrum at traditional cellular frequencies, both mmWave (30 to 300 GHz) and LTE-U (5 GHz) communications aim to boost the throughput by directly expanding the available radio spectrum for cellular networks (i.e., increasing B in Eqs. 1 and 2).

MILLIMETER-WAVE

MmWave communications exploit the enormous chunks of spectrum available at mmWave frequencies and are expected to push the mobile data rates to gigabits per second for 5G wireless networks. However, the propagation conditions at the extremely high frequencies of the mmWave bands lead to several fundamental issues including high path attenuation, severe blockage effects, and atmospheric and rain absorption, which pose new challenges for providing guaranteed quality of service (QoS). Thanks to the small wavelengths of mmWave bands, a large number of antenna elements can be integrated in a compact form at both the BSs and the UTs and leveraged to synthesize highly directional beams leading to large array gains. Energy-efficient mmWave communications comprises the following aspects.

Energy-Aware Transceiver Architectures: For mmWave frequencies, the conventional fully digital architecture employed for microwave frequencies, where each antenna is connected to a dedicated RF chain, is unsustainable. In particular, excessive power consumption arises from the processing of massive parallel gigasamples per second data streams, leading to an energy-inefficient and expensive system. Thus, a direct approach to reduce power consumption in practice is to adopt analog beamforming. Specifically, one RF chain is connected to multiple antennas, and the signal phase of each antenna is controlled by a network of analog phase shifters (PSs). However, pure analog beamforming with a single beamformer only supports single-user and single-data-stream transmission, which does not fully exploit the potential multiplexing gains provided by multiple antennas. As such, hybrid architectures have been proposed as energy-efficient alternatives for mmWave communications [4]. Specifically, the signal processing operations are divided between the analog and digital domains such that digital beamforming is utilized in addition to analog beamforming. Basically, two hybrid structures have been developed:
- Fully connected architecture: Each RF chain is connected to all antennas.
- Partially connected architecture: Each RF chain is connected only to a subset of antennas.

From the energy-aware hardware design perspective, the fully connected architecture requires thousands of PSs and introduces additional energy consumption, such as the high power compensation required for the giant phased radar and the insertion loss of the PSs. In contrast, for the partially connected architecture, the number of PSs needed decreases by a factor equal to the number of RF chains, and all signal processing is conducted at the subarray level. Hence, the system complexity as well as the circuit power consumption are significantly reduced, although at the expense of less flexibility in utilizing the spatial degrees of freedom. Nevertheless, for these structures, the hybrid precoding and combining schemes, the number of RF chains and antennas, and the transmit power can be separately or jointly optimized for customizing the system performance depending on the needs and requirements. Another emerging technique is to use lens antenna arrays with energy-focusing capability, which is passive and thus reduces the power consumption [5].

Low Resolution Design: Besides hybrid structures, employing low-resolution analog-to-digital converters (ADCs) at the receivers is an alternative approach toward energy-efficient designs. The theoretical motivation is that the power dissipation of an ADC scales linearly with the sampling rate and exponentially with the number of bits per sample. Furthermore, the data interface circuits connecting the digital components to the ADCs are usually power hungry and highly depend on the adopted resolution. Thus, the power consumption of high-speed and high-resolution ADCs becomes a critical concern and system performance bottleneck in mmWave systems employing a large number of antennas and very wide bandwidth. This motivates the configuration of low-resolution ADCs in practical systems, especially for battery-constrained handheld devices. However, a resolution reduction naturally causes a loss of data rate, which suggests that an optimal ADC resolution for striking a balance between the data rate and the power consumption exists such that the system EE is maximized. However, the characterization of the channel capacity taking into account low-resolution ADCs is in its infancy for mmWave communications and requires extensive research efforts. In addition, deploying a mix of high-resolution and low-resolution ADCs (i.e., mixed ADCs) is also a promising direction for power savings.

LTE-U

LTE-U further increases the LTE capacity by enabling the LTE air interface to operate in unlicensed spectrum. The unlicensed spectrum available in the 5 GHz band/WiFi band (≥400 MHz bandwidth) provides LTE-U with additional spectrum resources and was launched in Third Generation Partnership Project (3GPP) Release 13.

Harmonious Coexistence: The fundamental challenge of LTE-U is the coexistence of the LTE system and the incumbent unlicensed systems, such as WiFi systems. LTE employs a scheduling-based channel access mechanism, where multiple users can be served simultaneously by assigning them different resource blocks. In contrast, WiFi adopts contention-based channel access with a random backoff mechanism, where users are allowed to only access channels that are sensed as idle. Multi-user transmission with centralized scheduling enables LTE to make better use of the unlicensed band and also to achieve higher EE than WiFi since the channel sensing and backoff time in WiFi lead to a waste of spectrum resources. However, if left unconstrained, LTE-U transmission may generate continuous interference to WiFi systems such that the channel is detected as busy most of the time [6]. This will lead to unceasing backoff times for the WiFi links and incur high energy consumption and low EE for the WiFi users even though no WiFi data is transmitted. Therefore, intelligent modifications to the resource management in the unlicensed band becomes critical to harmonious coexistence. So far, two coexistence methods have been proposed in contributions to 3GPP:

- Listen before talk (LBT): LTE-U devices are required to verify whether the channel is occupied by WiFi systems before transmitting.
- Duty cycling: The LTE signal is periodically turned on and off, occupying and vacating the channel for periods of time without verifying the channel occupancy before transmitting.

Clearly, contention-based WiFi systems inherently follow the LBT protocol, which is deemed to facilitate fair coexistence. However, an increasing number of contending devices will lead to high transmission collision probabilities and thereby limit the system performance. In light of this, the duty cycling protocol seems to enable more efficient utilization of the unlicensed band. However, duty cycling assumes that the carriers are in charge of on-off scheduling, which contradicts the common property nature of the unlicensed spectrum. This is a drawback as the carriers are under no real obligation to provide a decent time window for WiFi. Although there is no globally accepted coexistence protocol yet, one popular consensus is that "LTE-U is often a better neighbor to Wi-Fi than Wi-Fi itself."

In summary, although both mmWave communication and LTE-U improve the system EE and SE by exploiting extra spectrum, the challenges facing these two technologies are different due to different operating frequencies. In particular, efficient transceiver design and coexistence protocol design are the main problems of mmWave communication and LTE-U, respectively.

GREEN 5G TECHNOLOGIES: SHORTENING TX/RX DISTANCE

Both UDNs and D2D communications boost the SE via shortening the distances between Txs and Rxs. Short-range communication has the dual benefits of providing high-quality links and a high spatial reuse factor compared to the current wireless networks (i.e., increasing K and decreasing d in Eqs. 1 and 2).

ULTRA-DENSE NETWORKS

UDNs are obtained by extensively deploying diverse types of BSs in hotspot areas and can be regarded as a massive version of heterogenous networks [7]. Energy-efficient UDNs involve the following aspects.

User-Centric Design: One of the green networking design principles of UDNs in 5G is the *user-centric* concept where signaling and data as well as uplink and downlink transmissions are decoupled [8]. A well-known architecture for the decoupling of signaling and data is the *hyper-cellular structure* where macrocells provide umbrella coverage, and microcells (utilizing, e.g., mmWave BSs) aim to fulfill the high capacity demands. Thereby, the significant signaling overheads caused by mobile device roaming in small size cells are reduced, which decreases the associated energy consumption at both transceivers substantially. Furthermore, the macrocell BSs, which are only capable of providing reliable coverage, can be replaced by more energy-efficient types of BSs rather than the conventional energy-consuming ones. In addition, the separation of signaling and data also eases the integration of other radio access technologies (RATs), such as WiFi and future mmWave RATs, which may help to realize further potential EE gains. Meanwhile, decoupling downlink and uplink enables more flexible user association schemes, which also leads to substantial energy savings for both BSs and UTs. For example, for two neighbor BSs where BS 1 and BS 2 are heavily loaded and with limited available spectrum left in the downlink and uplink, respectively, a UT can dynamically connect to BS 1 in the uplink and to BS 2 in the downlink. As such, both uplink and downlink transmit power consumption can be reduced, which is significant for 5G applications with ultra-high data rate requirements.

BS On/Off Switching: Separating signaling and data also enables efficient discontinuous transmission/reception (i.e., BS sleeping) to save energy via exploiting the dynamics of the wireless traffic load across time and geographical locations [9]. It has been shown that today, 80 percent of BSs are quite lightly loaded for 80 percent of the time but still consume almost their peak energy to account for elements such as cooling and power amplifying circuits. Therefore, BS sleeping is deemed to be an effective mechanism for substantial energy savings, especially for UDNs with highly densified BSs [10]. Specifically, data BSs (DBSs) can be densely deployed in order to satisfy the capacity requirement during peak traffic times, while a portion of the DBSs can be switched off or put into sleep mode if the traffic load is low in order to save energy.

Interference Management: In general, cellular networks are exposed to two major sources of interference, namely, intra-cell interference and intercell interference. The former is not a significant issue in today's cellular networks due to the use of orthogonal frequency-division multiple access (OFDMA) technology and BS controlled scheduling. However, the latter is a critical concern for UDNs due to the high frequency reuse factor in multi-tier and heterogeneous networks. For instance, because of interference, increases of

Although both mmWave communication and LTE-U improve the system EE and SE by exploiting extra spectrum, the challenges facing by these two technologies are different due to different operating frequencies. In particular, efficient transceiver design and coexistence protocol design are the main problems of mmWave communication and LTE-U, respectively.

With the unprecedented growth of the number of mobile devices and the data traffic, another benefit of D2D communication is the possibility of active user cooperation, which facilitates energy savings in 5G networks, especially with regard to extending the battery lifetime of handheld devices.

the transmit powers of two neighboring BSs will cancel each other out without improving the system throughput, which leads to low system EE. In addition, femtocell BSs may create "dead zones" around them in the downlink, especially for cell edge users associated with other BSs. Therefore, efficient interference management schemes such as power control, resource partitioning and scheduling, cooperative transmission, and interference alignment are needed for the successful deployment of energy-efficient UDNs. Although completely eliminating interference is overly optimistic for practical systems, it is expected that removing the two or three strongest interferers still brings an order-wise network EE improvement.

D2D Communications

D2D communications [11] enable densified local reuse of spectrum and can be regarded as a special case of ultra-dense networks with the smallest cell consisting of two devices as the transmitter and the receiver. Energy-efficient D2D communications involve the following aspects.

Mode Selection and Power Control: In D2D communication, there are two modes of communication: the cellular mode and the D2D mode. In the cellular mode, UTs are treated as regular cellular devices and communicate via BSs. In contrast, in the D2D mode, UTs communicate directly with each other without going through BSs by reusing the spectrum that has already been assigned to a cellular user (underlay communication) or has not been assigned to a cellular user (overlay communication). Underlay D2D communication generates co-channel interference and may switch to overlay D2D communication when the generated co-channel interference is strong. It has been shown that underlay D2D communication is preferable for EE-oriented designs, while overlay D2D communication tends to be more efficient for SE-oriented designs. This is mainly due to the interference mitigation characteristics of EE-oriented designs via limiting the transmit power.

Active User Cooperation: With the unprecedented growth of the number of mobile devices and the data traffic, another benefit of D2D communication is the possibility of active user cooperation, which facilitates energy savings in 5G networks, especially with regard to extending the battery lifetime of handheld devices. In particular, D2D devices can:

•Act as mobile relays for assisting cellular transmissions or other pairs of D2D transmissions. For example, for the uplink transmission of cell-edge users, the channel conditions are generally severely degraded, and direct uplink transmission to the BS incurs exceedingly high energy consumption, which would heavily affect the user experience and satisfaction. With proper devices as relays, significant energy consumption can be saved, and it can also be further extended into multihop, two-way, and multiple relay scenarios.

•Act as cluster heads for multicast transmission (e.g., for synchronous video streaming). It is known that the data rate of multicast transmission is limited by the worst channel among targeted members. However, with D2D functionality, any member (e.g., the member with the best channel) can be selected as the cluster head and further

multicasts its received data to other members, which achieves multiuser spatial diversity. Alternatively, the content may be divided into multiple chunks, and each member that receives a subset of them can share its data with others.

•Act as local caching devices for content exchange (e.g., for asynchronous video streaming). Wireless caching is appealing since the data storage at the devices is often underutilized, which provides a way to exploit the inherent content reuse characteristic of 5G networks while coping with asynchronism of the demands. Taking into account content popularity, deciding what content to cache plays a important role in alleviating backhaul burden as well as reducing power consumption. Although active user cooperation facilitates the efficient use of spectrum, it is still of practical interest to study how self-interested devices can be motivated to relay, share, and cache data for other devices at the cost of sacrificing their own limited energy.

As discussed, both UDNs and D2D communication bring the transmitter closer to the receiver. Thus, interference management has to be carefully addressed for these two technologies due to frequency reuse. Furthermore, the newly added BSs in UDNs will complicate problems, such as user association, backhaul deployment, and transceiver on/off switching issues. For D2D communication, rewarding and pricing schemes from economic theory may be resorted to when designing efficient D2D cooperative protocols and mechanisms.

Green 5G Technologies: Enhancing Spatial Degrees of Freedom via M-MIMO

M-MIMO is expected to provide at least an order of magnitude improvement in multiplexing gain and array gain via deploying a large number of antennas at the BSs while serving a large number of users with the same time-frequency resource (i.e., by increasing N in Eqs. 1 and 2). Energy-efficient M-MIMO involves the following aspects.

How Many Antennas Are Needed for Green M-MIMO Systems?

For an M-MIMO system with M antennas equipped at the BS and K single-antenna users, it has been shown in [12] that each user can scale down its uplink transmit power proportional to M and $\sqrt{M}$ with perfect and imperfect channel state information (CSI), while achieving the same performance as a corresponding single-input single-output (SISO) system. However, only reducing the transmit power consumption at the users is not sufficient for improving system EE, since the overall power consumption also includes the circuit power consumption, which increases linearly with the number of hardware components. Basically, a general guideline for determining the number of antennas needed for achieving green M-MIMO is: *when the transmit power largely dominates the overall power consumption, deploying more antennas yields a higher EE; while when the circuit power largely dominates the overall power consumption, further deploying*

more antennas is no longer energy-efficient. This is due to the fundamental interplay between the achievable throughput gain and the additional circuit power consumed as a result of deploying more antennas. In light of this, a realistic power consumption model is established in [13], where it was shown that deploying 100–200 antennas to serve a relatively large number of terminals is the EE-optimal solution for today's circuit technology.

SIGNAL PROCESSING AND GREEN HARDWARE IMPLEMENTATION

Centralized digital signal processing enables large savings in both hardware and power consumption. The favorable propagation conditions arising from deploying massive antennas significantly simplifies the precoding and detection algorithms at the BS. Linear signal processing techniques, such as maximum ratio combining (MRC) and maximum ratio transmission (MRT), can achieve near-optimal system throughput performance. Compared to existing sophisticated signal processing methods such as dirty paper coding (DPC) and successive interference cancellation, the simplifications facilitated by M-MIMO reduce the dissipated energy required for computations. Besides, a large number of antennas also permits green hardware implementation. It has been shown that M-MIMO systems require a significantly lower RF transmit power (e.g., in the milliwatt range), which will result in substantial power savings in power amplifier operations. However, the mobile hardware at UTs is still the system performance bottleneck due to more stringent requirements on power consumption and limited physical size, which needs further investigation.

PILOT DESIGN

For M-MIMO systems, massive timely and accurate CSI acquisition will also lead to significant pilot power consumption since the accuracy of the channel estimation will directly affect the achievable throughput. It is known that the required pilot resources for CSI acquisition are proportional to the number of transmit antennas, which makes frequency-division duplexing (FDD) unaffordable for practical implementation of M-MIMO systems, and novel schemes, such as pilot beamforming and semi-orthogonal pilot design, have to be exploited. Current research efforts mainly focus on time-division duplexing (TDD) systems, where users first send uplink pilots, and then the BSs estimate the required CSI via pilots for downlink data transmission. Furthermore, for multi-cell scenarios, pilot contamination caused by reusing the same pilot resources in adjacent cells also affects the EE and SE significantly. Therefore, how to balance the use of the time-frequency resource for pilot training and data transmission in both uplink and downlink as well as designing low-complexity pilot mitigation schemes are important challenges for achieving high EE in M-MIMO systems.

In summary, the hardware circuit power consumption, sophisticated signal processing, and accurate CSI acquisition are the main obstacles for realizing energy-efficient M-MIMO systems.

Technology	Energy source controllability	Efficiency	Distance	5G applications
Renewable energy harvesting	Non-controllable	Low	Long	Hybrid BSs
RF energy harvesting	Controllable	Low	Long	IoT, sensor networks
Inductive coupling	Controllable	Very high	Very short	Cell phone charging
Magnetic resonant coupling	Controllable	High	Short	Cell phone charging
Laser beaming	Controllable	High	Long	–

TABLE 2. Comparison of energy harvesting technologies.

ENABLER OF SUSTAINABLE 5G COMMUNICATIONS: ENERGY HARVESTING

Besides the above techniques for improving the EE, the energy harvesting technologies drawing energy from renewable resources or RF signals are important enablers of sustainable green 5G wireless networks. In Table 2, energy harvesting from renewable resources and RF energy harvesting as well as other energy harvesting technologies are compared from different perspectives.

ENERGY HARVESTING FROM RENEWABLE RESOURCES

Renewable resources energy harvesting transforms natural energy sources such as solar and wind into electrical energy, and has been recognized as a promising green technology to alleviate the energy crisis and to reduce energy costs. The following research directions for communication system design based on renewable energy have emerged.

Distinctive Energy-Data Causality and Battery Capacity Constraints: A fundamental feature that distinguishes energy harvesting communication systems from non-energy harvesting communication systems is the existence of an energy causality constraint [3]. This constraint imposes a restriction on the energy utilization, that is, energy cannot be utilized before it is harvested, which is in essence attributed to intermittent availability of renewable resources. In order to fully embrace the potential benefits of energy harvesting, carefully incorporating this constraint into the design of 5G networks and gaining deep understanding of its impacts on 5G networks are vital. Two other relevant constraints are the data causality constraint and the battery capacity constraint, which depend on the data traffic variations and energy storage limitations. Specifically, the data causality constraint accounts for the fact that a data packet cannot be delivered before it has arrived at the transmitter. The battery capacity constraint models the limited capacity of the battery for storing harvested energy. Thus, the conventional research studies assuming infinitely backlogged data buffers and unlimited energy storage can be regarded as special cases of the above constraints. However, in upcoming 5G networks, these assumptions are no longer valid due to the extremely divergent traffic and device types. Thus, new QoS concerns on energy harvesting communication naturally arise in 5G systems.

It is evident that the diverse applications and heterogeneous user requirements of sustainable green 5G networks cannot be satisfied with any particular radio access technology. Instead, an ecosystem of interoperable technologies is needed such that the respective technological advantages of the different components can be exploited together.

Offline vs. Online Approaches: Depending on whether the energy arrivals, CSI, data traffic, and so on are predictable or not, resource allocation algorithms for communication systems with energy harvesting can be categorized into *offline* and *online* policies. Specifically, based on non-causal knowledge information, optimal offline solutions can be developed by exploiting tools from optimization theory. Although this approach is impractical due to highly dynamic network variations in 5G communication networks, offline solutions provide performance upper bounds for online algorithms. In contrast, based on causal information, online schemes can be developed and further categorized into two branches. The first branch assumes that statistical knowledge information is available at the transmitter, and the optimal solutions thereby can be numerically obtained by exploiting dynamic programming theory. However, such an approach lacks analytical insight and cannot be exploited for handling large-scale systems due to the "curse of dimensionality." Alternatively, some suboptimal online algorithms are desirable in practice, which can be designed based on the insights observed from the optimal offline results. However, in some of the 5G application scenarios, statistical characteristics may change over time, or such information may not be available before deployment. This leads to a more aggressive branch that does not require any a priori information at the transmitter. The optimal solutions in this branch rely on the learning theory, where the transmitter learns the optimal transmission policy over time by iteratively performing actions and observing their immediate rewards.

Hybrid Energy Supply Solutions and Energy Cooperation: Due to the uncertain and dynamic changing environmental conditions, the harvested energy is generally intermittent by nature. This poses new challenges for powering 5G systems that require stable and satisfactory QoS. A pragmatic approach is to design a hybrid energy harvesting system, which uses different energy sources, such as solar, wind, diesel generators, or even the power grid, in a complementary manner to provide uninterrupted service. As an example, for 5G communication scenarios such as UDNs with a massive number of BSs, the integration of energy harvesting not only provides location diversity to compensate for the additional circuit power consumption caused by deploying more BSs, but also reduces the power grid power consumption significantly. In addition, densified BSs employing energy harvesting also facilitate possible energy cooperation between BSs, which has huge potential for realizing green 5G networks. For example, the BSs located in "good" spots will harvest more energy and can correspondingly serve more users or transfer their superfluous energy through power lines to BSs that harvest less energy to realize energy cooperation.

RF ENERGY HARVESTING

RF energy harvesting, also known as wireless power transfer (WPT), allows receivers to harvest energy from received RF signals [14, 15]. Compared to opportunistic energy harvesting from renewable resources, WPT can be fully controlled via wireless communication technologies, which has been therefore regarded as a promising solution for powering the massive number of small UTs expected in 5G applications such as IoT. The following research directions for RF energy harvesting have been identified.

Typical System Architectures and Extensions: In the literature, two canonical lines of research can be identified for implementing WPT. The first line focuses on characterizing the fundamental trade-off between the achievable rate and the harvested energy by exploring the dual use of the same signal. This concept is known as simultaneous wireless information and power transfer (SWIPT). Specifically, wireless devices split the received signal sent by a BS into two parts, one for information decoding and the other for energy harvesting. Thus, a fundamental issue of this line of work is to optimize the power splitting ratio at the receiver side to customize the achievable throughput and the harvested energy. The other line of research separates energy harvesting and information transmission in time (i.e., harvest and then transmit), leading to wireless powered communication networks (WPCNs). Specifically, wireless devices first harvest energy from signals sent by a power station and then perform wireless information transmission (WIT) to the target BS/receiver using harvested energy. Thus, the WPT and WIT durations have to be carefully designed, since more time for WPT leads to more harvested energy for using in WIT but leaves less time for WIT. *These two canonical architectures as well as their extensions serve as a foundation for wireless-powered UDNs, M-MIMO, mmWave networks, and so on, which widely utilize RF energy flow among different entities of 5G networks.*

Improving the Efficiency of WPT: Despite the advantages of WPT, the system performance is basically constrained by the low energy conversion efficiency and the severe path loss during WPT. Generally, the 5G techniques that are able to improve the performance of a wireless communication link can also be exploited to improve the efficiency of WPT. For example, narrow beams can be realized by employing multiple antennas at the transmitter side with optimized beamforming vectors, which is known as *energy beamforming* [15], and this fits the applications of M-MIMO and mmWave well. In addition, short-range-communication-based 5G techniques such as D2D communications and UDNs are capable of improving the efficiency of WPT by reducing the energy transfer distance. Besides improving the amount of harvested energy, the circuit power consumed by information decoding for SWIPT is also an important design issue since it reduces the net harvested energy that can be stored in the battery for future use. In particular, the active mixers used in conventional information receivers for RF to baseband conversion are substantially power-consuming. This motivates additional efforts on designing novel receiver architectures that consume less power by avoiding the use of active devices. Finally, efficient CSI acquisition schemes are also imperative for improving the WPT efficiency in practical scenarios.

Rethinking Interference in Wireless Networks with WPT: Another advantage of WPT is that not only dedicated signals but also co-channel interference signals can be exploited for energy harvesting. This fundamentally shifts our traditional viewpoint of "harmful" co-channel interference,
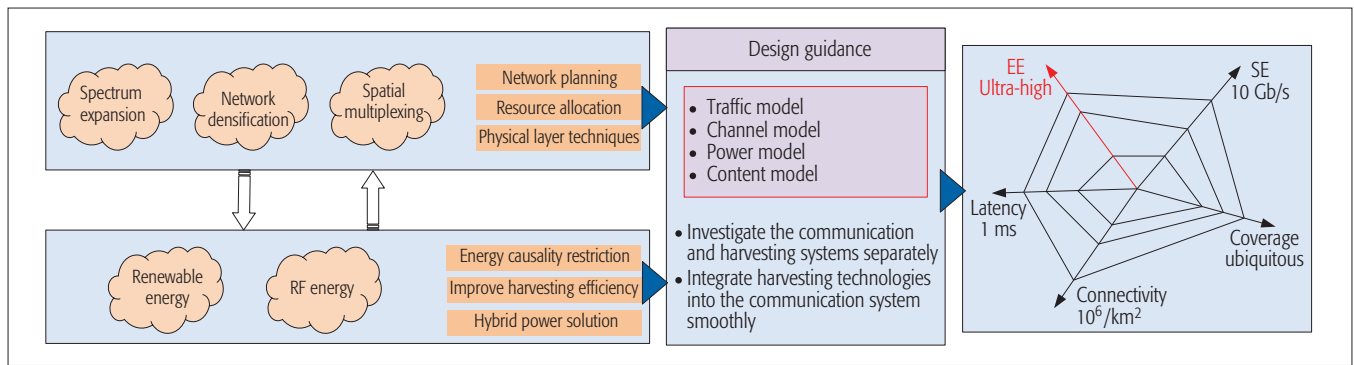
FIGURE 3. An overview of components of sustainable green 5G networks.

which now becomes a potential source of energy. In this regard, UDNs and D2D underlaying communications, where the spectrum is heavily reused across the whole network, provide plentiful opportunities for WPT to extract the potential gain from co-channel interference. In practice, WPT enabled devices can harvest energy when co-channel interference is strong and decode information when co-channel interference is relatively weak. Besides, one may inject artificial interference deliberately into communication systems, which may be beneficial for the overall system performance, especially when receivers are hungry for energy to support normal operations. As such, this paradigm shift provides new degrees of freedom to optimize the network interference levels for achieving the desired trade-off between information transmission and energy transfer.

In summary, renewable resource energy harvesting with conventional power grid and RF energy harvesting with conventional battery are envisioned to power BSs and devices, respectively, in upcoming 5G networks.

CONCLUSIONS

In this article, we have surveyed the advanced technologies that are expected to enable sustainable green 5G networks. A holistic design overview can be found in Fig. 3. Energy harvesting underpins the green expectation of 5G networks, while promising spectrum-efficient 5G technologies can be tailored to realize energy-efficient wireless networks. Facing the highly diversified communication scenarios of the future, user traffic, channel, power consumption, and even content popularity models need to be taken into account jointly for improving the system EE. Therefore, it is evident that the diverse applications and heterogeneous user requirements of sustainable green 5G networks cannot be satisfied with any particular radio access technology. Instead, an ecosystem of interoperable technologies is needed such that the respective technological advantages of the different components can be exploited together, pushing toward the ultimate performance limits. This, however, also poses new challenges for system designers.

REFERENCES

[1] J. G. Andrews et al., "What Will 5G Be?," IEEE JSAC, vol. 32, no. 6, June 2014, pp. 1065–82.
[2] Y. Chen et al., "Fundamental Trade-offs on Green Wireless Networks," IEEE Commun. Mag., vol. 49, no. 6, June 2011, pp. 30–37.
[3] D. Gunduz et al., "Designing Intelligent Energy Harvesting Communication Systems," IEEE Commun. Mag., vol. 52, no. 1, Jan. 2014, pp. 210–216.
[4] S. Han et al., "Large-Scale Antenna Systems with Hybrid Analog and Digital Beamforming for Millimeter Wave 5G," IEEE Commun. Mag., vol. 53, no. 1, Jan. 2015, pp. 186–94.
[5] Y. Zeng and R. Zhang, "Millimeter Wave MIMO with Lens Antenna Array: A New Path Division Multiplexing Paradigm," IEEE Trans. Commun., vol. 64, no. 4, Apr. 2016, pp. 1557–71.
[6] R. Zhang et al., "LTE-Unlicensed: the Future of Spectrum Aggregation for Cellular Networks," IEEE Wireless Commun., vol. 22, no. 3, June 2015, pp. 150–59.
[7] S. Samarakoon et al., "Ultra Dense Small Cell Networks: Turning Density into Energy Efficiency," IEEE JSAC, vol. 34, no. 5, Apr. 2016, pp. 1267–80.
[8] I. Chih-Lin et al., "Toward Green and Soft: A 5G Perspective," IEEE Commun. Mag., vol. 52, no. 2, Feb. 2014, pp. 66–73.
[9] S. Zhang et al., "How Many Small Cells Can Be Turned Off via Vertical Offloading Under a Separation Architecture?," IEEE Trans. Wireless Commun., vol. 14, no. 10, Oct. 2015, pp. 5440–53.
[10] S. Cai et al., "Green 5G Heterogeneous Networks Through Dynamic Small-Cell Operation," IEEE JSAC, vol. 34, no. 5, May 2016, pp. 1103–15.
[11] D. Feng et al., "Device-to-Device Communications Underlaying Cellular Networks," IEEE Trans. Commun., vol. 61, no. 8, Aug. 2013, pp. 3541–51.
[12] H. Q. Ngo, E. Larsson, and T. Marzetta, "Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems," IEEE Trans. Commun., vol. 61, no. 4, Apr. 2013, pp. 1436–49.
[13] E. Björnson et al., "Optimal Design of Energy-Efficient Multi-User MIMO Systems: Is Massive MIMO the Answer?," IEEE Trans. Wireless Commun., vol. 14, no. 6, June 2015, pp. 3059–75.
[14] X. Lu et al., "Wireless Networks with RF Energy Harvesting: A Contemporary Survey," IEEE Commun. Surveys & Tutorials, vol. 17, no. 2, 2nd qtr. 2015, pp. 757–89.
[15] S. Bi, C. K. Ho, and R. Zhang, "Wireless Powered Communication: Opportunities and Challenges," IEEE Commun. Mag., vol. 53, no. 4, Apr. 2015, pp. 117–25.

BIOGRAPHIES

QINGQING WU [S'13, M'16] (wu.qq@sjtu.edu.cn) received his B.Eng. and Ph.D. degrees in electronic engineering from South China University of Technology and Shanghai Jiao Tong University (SJTU), China, in 2012 and 2016, respectively. From 2015 to

2016, he worked as a visiting research scholar at the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta. He received the IEEE WCSP Best Paper Award in 2015 and Exemplary Reviewer of IEEE Communications Letters in 2016. He was also the recipient of outstanding Ph.D. thesis funding at SJTU in 2016. He served as a TPC member of IEEE VTC 2017, GLOBECOM 2016, and ICCC 2016, among others. His research interests include convex and nonconvex optimization, energy-efficient wireless communications, and unmanned aerial vehicle communications.

GEOFFREY YE LI [S'93, M'95, SM'97, F'06] (liye@ece.gatech.edu) is a professor with the School of ECE at Georgia Tech. His general research interests include wireless communications and statistical signal processing. In these areas, he has published around 400 articles with over 26,000 citations. He has been listed as a Highly-Cited Researcher by Thomson Reuters. He won the 2010 Stephen O. Rice Prize Paper Award from the IEEE Communications Society, the 2013 James Evans Avant Garde Award and 2014 Jack Neubauer Memorial Award from the IEEE Vehicular Technology Society, and the 2015 Distinguished ECE Faculty Achievement Award from Georgia Tech.

WEN CHEN [M'03, SM'11] (wenchen@sjtu.edu.cn) is a full professor with the Department of Electronic Engineering, SJTU, where he is also the director of the Institute for Signal Processing and Systems. From 2014 to 2015, he was the Dean of the School of Electronic Engineering and Automation, Guilin University of Electronic Technology, China. Since 2016, he has been the chairman of the Intellectual Property Management Corporation, SJTU. He has authored 80 papers in IEEE journals and transactions and over 100 papers in IEEE conference publications. His research interests include multiple access,and cooperative communications. He is an Editor of *IEEE Transactions on Wireless Communications* and an Associate Editor of *IEEE Access.*

DERRICK WING KWAN NG [S'06, M'12] (w.k.ng@unsw.edu) received his Bachelor's (Hons.) and M.Phil. degrees in electronic engineering from the Hong Kong University of Science and Technology in 2006 and 2008, respectively, and his Ph.D. degree from the University of British Columbia in 2012. He was a senior postdoctoral fellow with the Institute for Digital Communications, FAU. He is currently a lecturer with the University of New South Wales, Sydney, Australia. His research interests include convex and nonconvex optimization, physical layer security, wireless information and power transfer, and green (energy-efficient) wireless communications.

ROBERT SCHOBER [M'01, SM'08, F'10] (robert.schober@fau.de) received his Diplom and Ph.D. degrees in electrical engineering from the Friedrich-Alexander-University of Erlangen-Nuermberg (FAU) in 1997 and 2000, respectively. From 2002 to 2011, he was a professor and Canada Research Chair at the University of British Columbia, Vancouver, Canada. Since January 2012 he is an Alexander von Humboldt Professor and the Chair for Digital Communication at FAU. From 2012 to 2015, he served as Editor-in-Chief of *IEEE Transactions on Communications*. Since 2014, he is the Chair of the Steering Committee of *IEEE Transactions on Molecular, Biological and Multiscale Communication.*