5G Air Interface System Design Principles

Naga Bhushan, Tingfang Ji, Ozge Koymen, John Smee, Joseph Soriaga, Sundar Subramanian, and Yongbin Wei
Qualcomm Technologies, Inc.

The fifth generation (5G) of mobile networks is on the horizon and promises to be a new scalable framework that can efficiently multiplex envisioned and future 5G services. This kind of network will not only interconnect people, but also interconnect and control machines, objects, and devices. 5G should be a platform for innovations that can enable new services, empower new user experiences, and connect new industries. The design of 5G is focused not only on enhancing and lowering the cost per bit for mobile broadband services, but also enabling mission-critical control through ultra-reliable, low-latency communication (URLLC) links and connecting the massive Internet of Things (IoT).

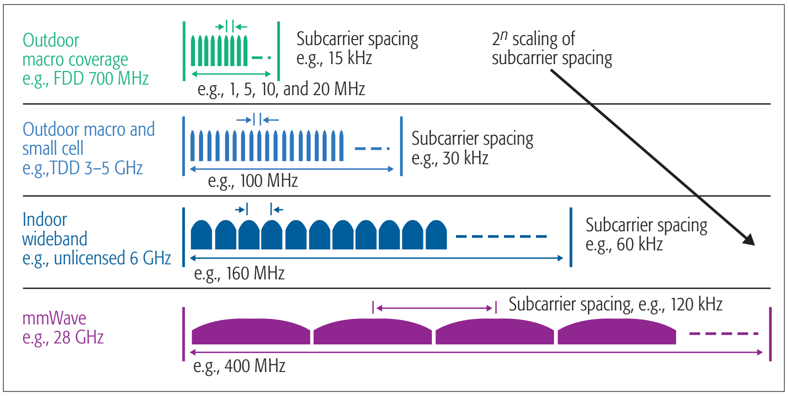

FIGURE 1. Example usage models and channel bandwidths.

At the heart of 5G is a new, more capable, unified air interface — 5G New Radio (NR). It adopts an optimized orthogonal frequency-division multiplexing (OFDM)-based family of waveforms and multiple access techniques, as well as a common, flexible framework that enables efficient multiplexing of various services and provides the forward compatibility required to future-proof 5G. This article explores these key foundational elements that are integral to the 5G NR air interface design. 5G NR will also incorporate a plethora of advanced wireless technologies that will bring new levels of performance and efficiency. We explore some of these advanced techniques from our viewpoint, with a focus on the key features targeted for Third Generation Partnership Project (3GPP) 5G NR Release 15 and beyond.

Designing 5G NR Unified Air Interface

One of the foremost decisions for designing a unified 5G air interface is the choice of radio waveforms and multiple access techniques. Not only will they need to deliver high performance at low complexity, but they must also be capable of supporting (and multiplexing) the initial and future 5G use cases efficiently.

Based on extensive studies [1], we believe that the OFDM family is the right choice for 5G enhanced mobile broadband (eMBB) and beyond. OFDM allows for enhancements such as waveform windowing/filtering, which can effectively minimize in-band and out-of-band spurious emissions — critical for 5G service multiplexing. The OFDM family can also efficiently coexist with other waveforms and multiple access schemes in the same framework, such as supporting asynchronous, grant-free transmissions (e.g., RSMA[1]) for connecting IoT devices and enabling mission-critical control communications.

Scalable OFDM Numerology with Scaling of Subcarrier Spacing

Today, LTE supports carrier bandwidths up to 20 MHz with mostly a fixed OFDM numerology — 15 kHz spacing between OFDM tones or subcarriers.[2] 5G NR, on the other hand, will introduce scalable OFDM numerology to support diverse spectrum bands/types and deployment models. For example, 5G NR must be able to operate in millimeter-wave (mmWave) bands that have extremely wide channel widths (e.g., hundreds of megahertz). It is critical that the OFDM subcarrier spacing is able to scale with the channel width, so the fast Fourier transform (FFT) size scales such that processing complexity does not increase exponentially for wider bandwidths.

In addition to supporting different channel widths for different deployment types, as seen in Fig. 1, 5G NR is also being designed to accommodate different/scalable numerologies, allowing services that use different bandwidths to efficiently multiplex in the same network. For example, using scalable numerologies, smaller subcarrier spacing provides larger cyclic prefix, which can be used to support broadcast service (Fig. 2), in the same carrier. The 5G NR unified air interface will also allow carrier aggregation across numerologies, such as aggregating mmWave and sub-6 GHz carriers to bring more robust and higher-performance connectivity.

A Flexible Framework with Forward Compatibility

Supporting the wide range of 5G services and devices requires more than optimized waveforms and multiple access. 5G NR is being designed with a flexible frame structure to efficiently multiplex diverse 5G services and provide forward compatibility for future ones. This equates to flexibility not only in the frequency domain, as discussed earlier with scalable OFDM numerology, but also in the time domain. The 5G NR framework will be able to support the diverse services, features, and deployment scenarios envisioned for 5G. Some high-level features of this new flexible framework are described in this section.

Scalable Transmission Time Interval

In the time domain, 5G NR will enable scalability to latencies much lower than what is possible in today's LTE networks. Today, LTE supports a fixed transmission time interval (TTI) of 1 ms. In the LTE evolution path, there is an ongoing Work Item on latency reduction in 3GPP. Although the technical details are still under discussion, the design target for hybrid automatic repeat request (HARQ) retransmission time is likely greater than 8 shortened TTIs (sTTIs) for 1.14 ms with 143 μs sTTI. To better support service with short latency requirement, the 5G NR flexible frame structure is being designed with TTI that will scale up and down, depending on specific service requirements. This flexibility allows the air interface to optimize for lower latencies using shorter TTI (e.g., hundreds of microseconds) or trading off for higher spectral efficiency for delay-tolerant use cases with a longer TTI. Scalable frame structure is also designed to cover all the frequency bands, from sub-6GHz all the way to mmWave bands.

In addition to a scalable TTI, 5G NR will also support service-aware TTI multiplexing on the same frequency, which allows transmissions with different TTIs to start on integer symbol boundaries instead of a subframe boundary (i.e., 1 ms). For instance, a high-quality of service (QoS) mobile broadband service may

---

[1] Resource-spread multiple access.

[2] Some exceptions, such as narrowband IoT (NB-IoT) as defined in 3GPP Release 13, can support transmissions with 3.75 kHz subcarrier spacing.

choose to utilize a 500 µs TTI instead of a standard, LTE-compatible 1 ms TTI/subframe, while another latency-sensitive service further shortens its TTI to ~140 µs with scaled numerology. Instead of requiring the second, latency-sensitive transmission to wait until the beginning of the next subframe (i.e., 500 µs later), it can begin as soon as the previous transmission is completed on the symbol boundary, thereby eliminating a waiting period.

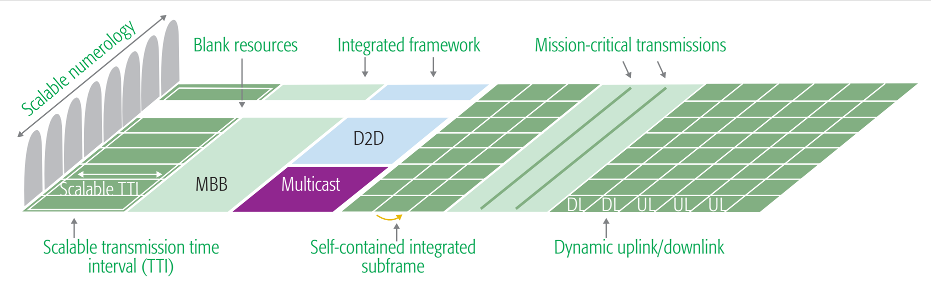

FIGURE 2. Scalable framework that can efficiently multiplex envisioned and future 5G services.

SELF-CONTAINED INTEGRATED SUBFRAME

The self-contained integrated subframe is another key enabler for lower latency, forward compatibility, and many new 5G NR features. The lower latency is achieved by having the data transmission and its acknowledgment all contained in the same subframe. Figure 3 shows an example of a time-division duplex (TDD) downlink-centric subframe, where data transmission is from the network to the device, and the acknowledgment is sent by the device back to the network in the same subframe. With the 5G NR self-contained integrated subframe, each TTI is now a modular transaction (e.g., DL grant → DL data → guard period → UL ACK) that gets completed within that time period.

The modular aspect of the self-contained integrated subframe design also allows for different types of subframes to be multiplexed for new services that are introduced in the future. This, along with the ability for the 5G NR framework to support blank subframes and blank frequency resources, enables a forward-compatible 5G NR design for easily adding future features/services to be deployed in the same frequency in synchronous and asynchronous manners.

The subframe can also contain additional precursor headers that can be used to provide additional information for the transmission. For example, operating in unlicensed or shared spectrum typically requires the support of Listen-Before-Talk (LBT) to ensure fair sharing across different users. These LBT headers can include measurement gaps used by downlink devices to assess channel availability for data transmission as well as precursor transmissions to signal channel occupancy. And when used in device-to-device communications, these headers can indicate the link direction and provide scheduling information for the directly connected devices.

The self-contained subframe also plays an integral role in enabling advanced 5G NR antenna techniques, such as massive multiple-input multiple-output (MIMO). On the downlink, in order for a transmitting cell to more efficiently direct RF energy to a device, it needs to continuously evaluate link quality and make necessary beamforming adjustments. The feedback mechanism is provided by the transmission of common uplink burst, which carries uplink control information (e.g., the ACK) and the uplink sounding reference signals. More accurate and timely knowledge of downlink channel can be obtained at the base stations thanks to channel reciprocity, which enables the use of uplink sounding reference signal for downlink channel estimate in a TDD system. Note that in addition to the uplink control information and sounding signals, common uplink burst can also be used to carry time critical data (e.g., TCP ACK).

An uplink-centric subframe has a similar structure, with similar downlink and uplink control bursts at the beginning and the end of the subframe, but with an uplink data burst that follows a guard period. This design reduces downlink and uplink control channel interferences by requiring all control bursts to be transmitted in the same direction across neighboring cells, thereby allowing more robust link direction switching. The dynamic configuration of downlink- or uplink-centric subframes increases overall network efficiency and capacity by allowing faster switching based on network traffic conditions.

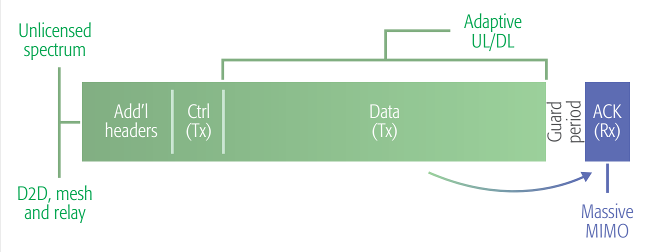

FIGURE 3. Self-contained integrated subframe design (e.g., TDD downlink).

MASSIVE MIMO AND MMWAVE FOR EXTREME CAPACITY AND UNIFORM USER EXPERIENCE

One key innovation area in wireless communications is advanced antenna technologies. By using more antennas intelligently, one can improve both network capacity and coverage. That is, more spatial data streams can significantly increase spectral efficiency (e.g., with MU-MIMO[3]), allowing more bits to be transmitted per Hertz, and smart beamforming can extend the reach of base stations by focusing RF energy in specific directions. In LTE today, networks are evolving from 2 × 2 to 4 × 4 MIMO, with even more antennas in sight. However, there is an intrinsic limitation on how many antennas one can realistically fit onto a device, especially at low frequencies where the antennas are large due to long wavelength. One way to further increase capacity without adding more device antennas is to have more antennas at the base stations.

By continuing to evolve FD-MIMO,[4] which was first introduced in LTE Advanced Pro, 5G NR will support massive MIMO [2, 3], which can utilize an even larger number of antenna elements, supporting up to 256 as currently defined. The 2D antenna arrays are capable of 3D beamforming in both the azimuth and elevation planes, and test results have shown significant gains in both capacity and cell edge user throughput — key to bringing a more uniform mobile broadband user experience.

Massive MIMO is also a key enabler for opening up the high-frequency bands in the sub-6 GHz spectrum. With intelligent beamforming and beam-tracking, it is possible to reuse existing cell sites (e.g., at 2 GHz) and transmit power for new macrocell networks that operate at higher frequencies (e.g., at 4 GHz). Simulation results demonstrated more than 4× improvement in average capacity going from a 2 × 4 MIMO to a 24 × 4 massive MIMO setup, and almost the same level of throughput enhancements for users at the cell edge.

5G NR will not only enable the use of higher frequencies in the 3 to 6 GHz band for macro/small cell deployments, but will also open up new mmWave opportunities for mobile broadband. The abundant spectrum available at these high frequencies is

---

[3] Multi-user MIMO.

[4] Full-dimensional MIMO.

capable of delivering extreme data speeds and capacity that will reshape the mobile experience. However, mobilizing the mmWave comes with its own set of challenges. Transmissions in these higher bands suffer from significantly higher path loss as well as susceptibility to blockage, while meeting the power and form-factor requirements of mobile devices has also proven to be challenging. Further, antennas for a given aperture get progressively more directive with frequency. Thus, traditional mmWave implementations have been limited to mostly stationary applications such as shorter-range wireless docking, enabled by technologies like 802.11ad that operates in the 60 GHz band.

With the recent advancements in signal processing, mmWave components, and antenna technologies, the idea of mobilizing the mmWave is no longer out of reach. By utilizing a large number of antenna elements in both the base station and the device, along with intelligent beamforming and beam-tracking algorithms [4], 5G mmWave can provide increased coverage, reduced interference, and a continuous connectivity experience even for non-line-of-sight (NLOS) communications and device mobility.

In order to provide continuous connectivity with high data rate in mmWave bands, it is imperative to have mechanisms that can provide the mmWave base station and user equipment (UE) with the right beamforming patterns at any given time. The beams chosen should provide near-optimal link gain as well as be robust to sudden changes or mobility at the UE or in the channel. However, traditional techniques of learning the full channel matrix (NT × NR) and computing the right Eigenmode beamformers for the channel do not work well for mmWave channels as:

- Learning the full channel matrix is time- and resource-intensive.
- The Eigenmode-based beamformers are very sensitive to channel changes.
- The channel is likely to be primarily composed of a few specular/strong directional components [8].

Initial directional search based on periodic directional sweeps and further refinement and on-demand tracking around primary directions identified in periodic sweeps have been observed to be more robust and close to optimality. In [4], a few simulation results and analytical results comparing different approaches to beam search and tracking are provided. Moreover, 5G NR will also leverage multi-connectivity with 5G sub-6 GHz and/or gigabit LTE to improve overall link robustness and to help achieve faster system acquisition. Additional design considerations have also been incorporated to enable easy deployment of dense mmWave small cells, such as with the support for integrated access and backhaul.

Advanced Channel Coding

Along with the scalable numerology and flexible framework for 5G NR services, the physical layer design should include an efficient channel coding scheme that can provide robust performance and flexibility. Although turbo codes were the basis for 3G and 4G technologies, low-density parity check (LDPC) codes [5] have demonstrated advantages from both the complexity and implementation standpoints when scaling to very high throughputs and larger blocklengths. In particular, an advanced LDPC channel coding scheme with a flexible parity check matrix structure can provide 5G NR with full rate compatibility including incremental redundancy (IR) HARQ, and flexibility for blocklength scaling and fine rate granularity, all while allowing for an easily parallelizable decoder design that can be scaled to achieve higher throughputs at low complexity. Such codes can also provide an effective solution for mission-critical and massive IoT traffic, efficiently reusing the hardware from mobile broadband services.

For very small blocklengths such as 80 bits or less, which are often used in physical control channels, polar codes [6, 7] provide an encoding structure that can allow 5G NR to achieve performance gains over the existing LTE tail biting convolutional code (TBCC). In fact, the encoding structure of these codes can be thought of as generalizations of Reed Muller codes, which are also employed at the smaller blocklengths.

Carrier Aggregation across Spectrum Bands and Types

5G will build on the solid LTE carrier aggregation (CA) foundation to support aggregation across frequency bands (e.g., low band FDD, mid-band TDD, and mmWave TDD) and types (e.g., licensed, shared, and unlicensed).

CA delivers many benefits; for example, adding a supplemental low-band FDD uplink to mid-band TDD can improve system performance for both NSA and SA[5] deployments. For NSA, the low-band LTE or supplemental NR uplink can enhance the data rate and range of the deployment. For SA, using low-band NR to carry uplink control and data for cell edge users can improve robustness of the mid-band system.

Device-Centric Mobility to Improve Energy and Overhead Efficiency

In addition to downlink-centric mobility, 5G NR will consider uplink- or device-centric mobility, which allows the device to send out periodic reference signals for the access network to monitor; with this information, the network can trigger cell reselection or handover based on the uplink signal strength measurement. This offloads the device from monitoring and processing reference signals from all nearby cells, thereby decreasing both signaling and processing overhead. This is in contrast to LTE, where device mobility is driven by a more overhead-heavy process that involves the device measuring the signal strength of a downlink reference signal sent by the access network, which requires overhead processing on the device. To optimize even further, the 5G network will also reduce the amount of broadcasts. Instead of sending system information regardless of device presence, the 5G network will only send out minimum system information periodically, and on-demand system information upon requests from devices. The reduction in signaling and processing overhead will enable longer battery life on the devices as well as make the network more energy-efficient; in addition, the new device-centric mobility design also contributes to supporting seamless handovers at higher speeds (e.g., up to 500 km/h).

Conclusion

In summary, 5G is a unifying connectivity fabric that will expand the value of mobile networks to take on a much larger role than previous generations, empowering many new connected services across an array of world-changing use cases. At the heart of 5G is the new 5G New Radio unified air interface that is being designed to meet the expanding connectivity needs in the next decade and beyond.

References

[1] "5G Waveform & Multiple Access Techniques Whitepaper," Qualcomm Technologies, Inc.; https://www.qualcomm.com/documents/5g-research-waveform-and-multiple-access-techniques

[2] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas." *IEEE Trans. Wireless Commun.*, Sept. 11,, 2010, pp. 3590–600.

[3] C. Shepard *et al.*, "Argos: Practical Many-Antenna Base Stations," *Proc. 18th ACM Annual Int'l. Conf. Mobile Computing and Net.*, 2012.

[4] V. Raghavan *et al.* "Beamforming Tradeoffs for Initial UE Discovery in Millimeter-Wave MIMO Systems," *IEEE J. Selected Topics Signal Proc.,* Oct. 3, 2016, pp. 543–59.

[5] T. Richardson and U. Rüdiger, "Multi-Edge Type LDPC Codes," workshop honoring Prof. Bob McEliece on his 60th Birthday, Caltech, Pasadena, California, 2002.

[6] E. Arıkan, "Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels," *IEEE Trans. Info. Theory*, vol. 55, no. 7, July 2009, pp. 3051–73.

[7] I. Tal and A. Vardy, "List Decoding of Polar Codes," *IEEE Trans. Info. Theory*, vol. 61, no. 5, May 2015, pp. 2213–26.

[8] 3GPP Tech. Spec. Group Radio Access Network, "Channel Model for Frequency Spectrum above 6 GHz, TR 38.900, June 2016.

---

[5] Non-standalone (NSA) 5G NR utilizes LTE as an anchor (e.g., to support mobility management and other signaling). Standalone (SA) 5G NR operates as an independent system without an LTE anchor.