Application Notes

Toward Real-Time Software-Defined Radios for Ultrabroadband Communication Above 100 GHz

■ **Viduneth Ariyarathna, Arjuna Madanayake, and Josep M. Jornet**

The advancement of wireless technologies for mobile broadband communications, health care, robotics, and other application verticals demands reduced latency, increased capacity, and denser penetration in both urban and rural environments. Wireless data rates have doubled every 18 months over the last 30 years. It is therefore expected that faster connections will be a pressing requirement in the near future. Terabits-per-second (Tbps) data rates will need to be a manifest realty in the next five years and, presumably, a prime design objective in 6G networks [1].

IMAGE LICENSED BY INGRAM PUBLISHING

A solution that furnishes both high data rate and high aggregate network capacity calls for an increase in bandwidth. The search for large, continuous bandwidth entails climbing the frequency axis. Although optical communication (100–750 THz) is a solution, the (sub-) terahertz (THz) band (100 GHz–10 THz) presents a more obvious extension to current networks. The THz band offers favorable propagation characteristics compared to optics and is currently underutilized (only a small subset of frequencies are allocated to science experiments) [2], [3].

Much of the progress in THz-band systems has been made at the device level through electronic, photonic, and plasmonic approaches [4]. In parallel, considerable work has been reported in modeling and measuring the THz channel [5]. Although there has been significant progress toward the development of digital signal processing (DSP), communication, and networking solutions, this work has been mostly theoretical in nature due to the lack of testbeds that can support experiments beyond channel sounding [6], [7].

The majority of existing testbeds that have carrier frequencies above 100 GHz and that are able to support ultrabroadband (approaching 10 GHz or more) physical-layer (PHY) experimentation do not operate in real time [6], [8]. Instead, such testbeds rely on offline processing using a high-end arbitrary waveform generator

Viduneth Ariyarathna (v.ariyarathna@northeastern.edu) was with Northeastern University Boston, MA 02115 USA. He currently works in the wireless semiconductor industry. Josep M. Jornet (j.jornet@northeastern.edu) is with the Institute for the Wireless Internet of Things and the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115 USA. Arjuna Madanayake (amadanay@fiu.edu) is with the Department of Electrical and Computer Engineering, Florida International University, Miami, FL 33199 USA.

Digital Object Identifier 10.1109/MMM.2023.3277362
Date of current version: 13 July 2023

at the transmitter and a digital storage oscilloscope (DSO) at the receiver that can handle ultrabroadband bandwidths, thanks to cutting-edge data converters exceeding 100 gigasamples-per-second (GSps) speeds. Although these platforms can be utilized to test new waveforms and PHY algorithms, they do not support networking solutions or dynamic scenarios due to a lack of real-time processing capabilities.

A couple of testbeds have been reported to operate in real time above 100 GHz [9], but the bandwidths of such systems do not exceed 2 GHz. There are commercial software-defined radios (SDRs) that also exist with real-time digital backends (e.g., National Instruments' millimeter-wave (mm-wave) system [10]), but the baseband bandwidths supported by them are again limited to less than 2 GHz and hence do not provide experimental capabilities for emerging above 100-GHz technologies. The development of 6G systems requires SDRs that are able to process basebands bandwidths with tens of GHz while tackling the characteristics of (sub-) THz channels. In this article, we identify the challenges in the real-time DSP of ultrabroadband signals and discuss possible solutions (see the "Challenges in Ultrabroadband DSP" section), such as frequency-multiplexed multichannel backends (see the "Solutions for Ultrabroadband DSP" section). As an early proof of concept, we design a Xilinx RF system-on-chip (SoC)-based real-time DSP engine (see the "Multi-GHz Multichannel DSP on RF SoCs" section), which we utilize to provide first-of-a-kind experimental results of an 8-GHz bandwidth, real-time SDR platform in the D-band (110–170 GHz) (see the "RF SoC-Based Real-Time Prototype Supporting 8 GHz of Bandwidth at the D-Band" section).

Much of the progress in THz-band systems has been made at the device level through electronic, photonic, and plasmonic approaches.

Challenges in Ultrabroadband DSP

Despite advances in 100-plus-GHz analog front ends, software-defined ultrabroadband DSP solutions remain uncertain. Figure 1 shows the architectures of conventional RF front ends for a single-input, single-output system. Following the Nyquist sampling theorem, the DSP has to happen at a frequency f_s that is greater than twice the bandwidth B of the signal. Figure 1(a) and (b) shows a heterodyne architecture where a low-intermediate frequency (IF) signal is generated by the digital front end at the transmitter and a signal at a low-IF frequency is sampled at the receiver, respectively. For such systems, the digital front-end circuits (at least the blocks closer to data converters) operate at an effective rate of $2\times(f_{\text{IF}}+B/2)$ for the case of first-Nyquist-zone operation in the data converters [depicted in Figure 1(a) and (b)]. Figure 1(c) and (d) shows direct upconversion/downconversion architectures where both the transmitter and receiver use two data converters for the in-phase (I) and quadrature

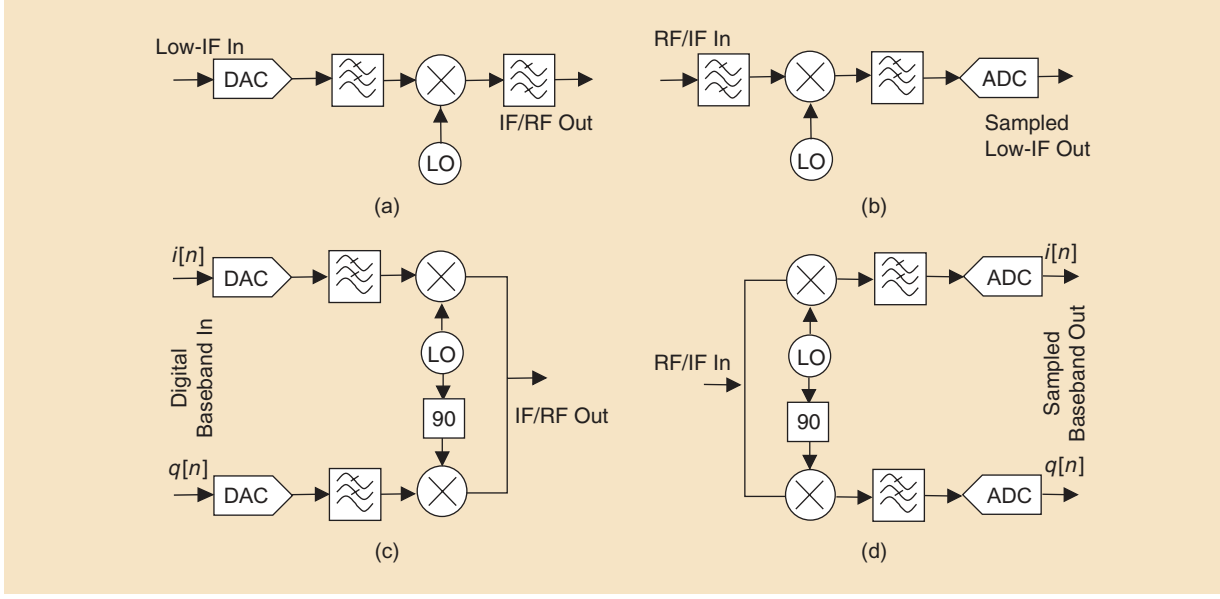

Figure 1. *Conventional transmit-receive RF front-end architectures. (a) Heterodyne transmitter where a low-IF is generated from digital signal. (b) Heterodyne receiver where a low-IF is sampled into digital. (c) Direct upconversion from baseband. (d) Direct downconversion to baseband. IF: intermediate frequency; LO: local oscillator; ADC: analog-to-digital converter; DAC: digital-to-analog converter.*

The adoption of systolic-array processors may allow 100% processor utilization and the maximum possible speedup afforded by Amdahl's law.

(Q) components. In these, the data converters need to be clocked at a B sampling frequency at a minimum. In practice, oversampling (with interpolation/decimation) is common. Oversampling is needed at the data converters to reduce sharpness of the transition bands in the anti-imaging and antialiasing filters; for such cases, the aforementioned sample rates have to be scaled by the oversampling factor (which is ignored in this discussion).

The desired signal bandwidth B for above-100-GHz systems would be at least a few tens-of-GHz wide, especially in the race to achieve Tbps data rates. Evolving data-converter designs incorporating interleaved analog-to-digital converter (ADC)/digital-to-analog converter (DAC) architectures are capable of handling sample rates in excess of 256 GSps [11]. Although these capabilities are encouraging, the high cost of such devices make them unlikely to be adopted in commercial systems. With Moore's law slowing down, DSP clock frequencies have stagnated at a few GHz, which means that although the data converters support very high sampling rates, modern digital hardware is not capable of rendering such high sample rates for real-time processing. Currently, the maximum clockable frequencies of digital application-specific integrated circuits (ASICs) reach a few GHz. Field-programmable gate arrays (FPGAs) have a slower fabric (a maximum of ~500 MHz) and their maximum realizable digital circuit speeds are much lower than what is possible with an ASIC. On one hand, this entails faster digital fabric technologies, and on the other, innovative solutions exploiting both massive parallelism and multirate processing to perform DSP of ultrabroadband signals [12].

A possible alternative to using custom digital hardware for DSP is to conduct all the processing in software, like in traditional SDRs (such as legacy Universal Software Radio Peripherals), where all PHY (and up) processing happens on a host using a general purpose processor. But for ultrabroadband bandwidths this becomes challenging as time-critical functions have to be performed at GSps rates. Even nontime-critical functions like beam tracking, channel estimation, and others become much more challenging when using software processing at these frequencies. For example, the update rate for channel estimation is directly tied to the coherence time of the channel, which depends mainly on carrier frequency. In a typical 2.4-GHz Wi-Fi-like channel, the update rate may be ~1,000 updates per second, which is slow enough for implementation on a fast embedded processor inside a traditional SDR or an SoC (e.g., ARM core or RISC-V processor). However, as the carrier frequency increases 100–1,000-times beyond 100 GHz, the channel coherence times are also expected to scale down by the same factors. This assumes that all other conditions, such as user motion and multipath, are similar, leading to 100–1,000-times speedup requirements for the corresponding estimation algorithms, which makes software processing of nontime-critical functions impossible for ultrabroadband communication.

On the flip side, the recently initiated Open Radio Access Network (O-RAN) [13] is a push toward softwarization of 5G and beyond networks and aims at developing next-generation open, software-defined, and intelligent-virtualized cellular networks. The O-RAN architecture builds upon disaggregation of monolithic base stations into different components for the radio (radio unit) and the processing of different layers of the protocol stack [distributed unit (DU) and centralized unit]. But, even in this current O-RAN architecture, the DU relies on FPGAs or GPU-based accelerators [13] to perform the PHY functions of 5G New Radio, which employs 800 MHz of bandwidth. Fully GPU-based processing is another option (e.g., Nvidia's Arial framework for 5G O-RAN). However, GPU-based systems are still unlikely to be able to meet the computation speeds demanded by future ultrabroadband systems. Moreover, processor-based computations generally consume more power than custom-silicon/FPGA-based solutions.

Additionally, shifting from traditional SDRs to fully custom VLSI architectures based on non-Von Neumann computing algorithms becomes important when processing ultrabroadband signals and implementing computationally intensive algorithms. The adoption of systolic-array processors may allow 100% processor utilization and the maximum possible speedup afforded by Amdahl's law. Systolic arrays are modular, regular, and locally interconnected, with the DSP algorithm being hardware wired into the topology of the interprocessor connections. Systolic arrays are perfectly suited for exploiting the massive parallelism available in deep-nano-CMOS technology nodes, such as gate-all-around (GAA) and fin field-effect transistor (FinFET) approaches going down to the 3-nm CMOS node. Systolic-array processors can be efficiently used to partition several dozens of GHz of real-time bandwidth across multiple parallel channels that work independently of each other using multiphase clocks, as described by polyphase signal processing theory. The use of systolic arrays also enables the fastest possible computational throughput from a given VLSI technology node.

Solutions for Ultrabroadband DSP

Given the maxed-out clock rates in digital CMOS, the adoption of massive parallelism with multirate processing seems a viable option. We envisage systolic arrays for real-time DSP of ultrabroadband signals. Parallel sample processing digital architectures have already been adopted in commercial ASICs that involve signal bandwidths comparable to or exceeding maximum digital clock frequency limits. If the total data-converter sampling rate is f_s and the maximum digital clock rate is f_{clk}, then, $\lceil f_s/f_{clk} \rceil$ of parallel digital hardware would be required to process the bandwidth. Figure 2 shows a finite-impulse response filtering application at the receiver, which uses M phases to process M samples at each clock cycle of f_{clk}, leading to M output samples per cycle. The number of phases satisfies $f_s = Mf_{clk}$. There exists a design choice between the maximum bandwidth and resource utilization plus power consumption, given that the DSP operates at the maximum possible clock rate, with no serialized components, to furnish maximum throughput following Amdahl's law. Sampling of ultrabroadband signals can take the following two approaches:

1) *Time multiplexing* (*TM*): In this approach, a time-interleaved ADC (an array of polyphase ADCs that are time-interleaved to capture the entire bandwidth) where the entire full bandwidth signal (either at IF or baseband) is sampled and interfaced to a DSP through $\lceil f_s/f_{clk} \rceil$ phases where $f_s > 2B$. TM requires precise time-interleaved data converters and a costly ultrastable clock source.
2) *Frequency multiplexing* (*FM*): The ultrabroadband baseband is split into multiple bands across analog channels, where each channel is downconverted, low-pass filtered, and sampled using parallel data converters, with each supporting a relatively low bandwidth. Here, the RF bandwidth B is determined by the number of channels multiplexed (N_{ch}) and the bandwidth of each channel (B_{ch}) such that $B = N_{ch}B_{ch} + B_g$, where B_g represents the total guard-band bandwidth.

The former approach has the advantage of not requiring baseband microwave filter banks in the analog domain but suffers from the need for stable clock sources and for extremely precise time-interleaved data converters that come at the highest cost. The digital processing in such a setup would require a large number of phases in the circuit to capture the whole bandwidth.

The latter is more complex in terms of analog microwave circuitry and requires multiple low-end data converters utilized per channel but also needs simpler single-phase or a relatively low number of parallel phases per channel in systolic-array digital realizations. This approach has an advantage over the dynamic range; for example, in an orthogonal frequency-division multiplexing (OFDM)-based PHY, the peak-to-average power ratio is a function of the fast Fourier transform (FFT) size, and having a channelized approach provides an advantage over the dynamic range in data converters for maintaining the same frequency selectivity in the FFT bins over the full band. This fact benefits automatic gain control (AGC) at the receiver side, and AGC can happen per channel as well, obviating the need for the sophisticated AGC algorithms that can be required for ultrabroadband signals. Although it is much harder to build ADCs with ultrabroadband bandwidths that maintain a high effective number of bits (ENoBs) over the full band, using a channelized approach obtains a higher ENoB using stable technologies over the multiple narrow channels covering the full band. But such a channelized system would require sufficient guard band to be employed among the channels, which reduces the spectral efficiency compared to the former approach, which utilizes the full, contiguous bandwidth.

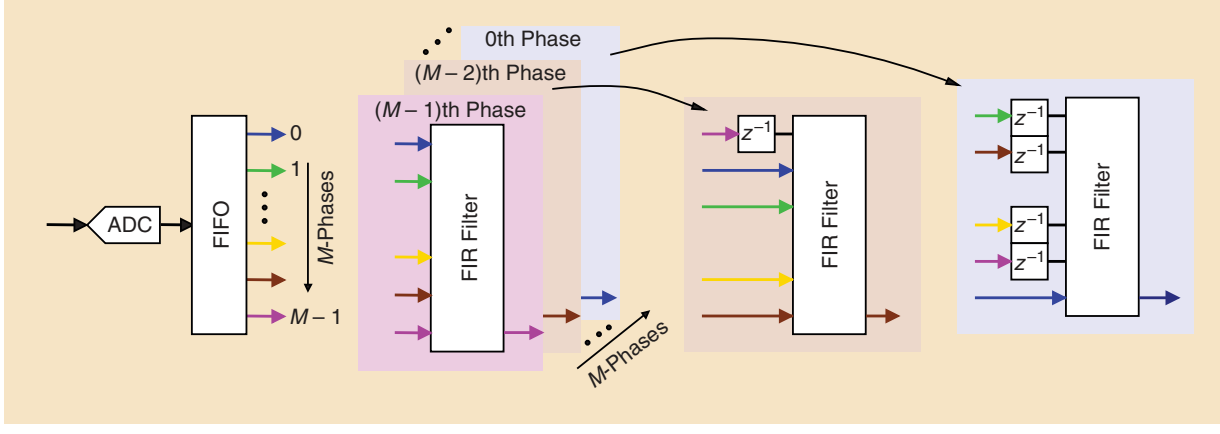

Figure 2. *(a) A rate adjustment using a first-in, first-out (FIFO)buffer in the digital domain for parallel processing of multiple samples simultaneously, and an illustration of how the parallel samples (phases) are processed using multiple copies of digital circuits to process the full bandwidth. FIR: finite-impulse response.*

RF SoCs have been used for phased-array solutions for 5G applications to interface multiple antennas for fully digital beamforming.

For either approach, the complexity of the DSP hardware involved is determined by the overall information-bearing bandwidth. As multiphase digital architectures involve complex hardware, it is imperative to explore novel digital computing strategies that transcend the conventional, fixed-point computing architectures (e.g., residue number systems, Dirichlet number systems, and posits) for 6G radio design. Such approaches may lead to better computational throughput and energy efficiency in DSP in future SDRs.

Multi-GHz Multichannel DSP on RF SoCs

RF SoC

In 2017, Xilinx released the first RF SoC [14] to integrate data converters on the same SoC as the programmable logic (PL). Modern RF SoCs come with integrated data converters and support up to a few GHz of bandwidth per converter. RF SoCs bring integration of the data converters and elimination of external interfaces/components; this reduces system power consumption by removing the need for JESD204-like serializer interfaces. RF SoCs provide multiple (eight or 16) data converters integrated into single chip. For example, the Xilinx ZU28DR chip supports eight DACs at 6.554 GSps and eight ADCs at 4.096 GSps; the chip supports a total of ≈ 16 GHz of bandwidth across all ADC channels and ≈ 25 GHz of bandwidth across DACs. Recently, Xilinx released Gen-3 chips, which can support up to 10-GSps sample rates at the DAC side, and 5-GSps Gen-3 chips with 16 channels that can support up to 80 GHz of bandwidth. RF SoCs have been used for phased-array solutions for 5G applications to interface multiple antennas for fully digital beamforming [15].

Multichannel DSP Using RF SoCs

The large overall bandwidth supported by RF SoCs can be leveraged in a channelized DSP engine that can aggregate multi-GHz-wide channels for interfacing with ultrabroadband front ends at 100-plus GHz [16]. Figure 3 shows an SDR that uses high-bandwidth data converters in the RF SoCs to process in parallel wideband channels corresponding to an aggregate ultrabroadband system.

The system, with the architecture of the transmitter and receiver shown in Figure 3(a) and (b), comprises two subsystems: DSP baseband processors and an analog IF circuit that implements the FM. Parallel baseband DSPs on RF SoCs with DAC outputs generate multiple analog baseband IQ waveforms, which are upconverted to different IF frequencies to form an aggregated multichannel IF signal using an analog multiplexing circuit. Multiple RF SoC systems can be used to aggregate more channels to the IF signal. The number of aggregated channels is adjustable adaptively by digitally turning on/off the channels to fit the bandwidth requirement at the transmission window. The receiver side acts reciprocally.

The analog transmit/receive chains can be either heterodyne or homodyne. Although both architectures enable the use of the full bandwidth supported by the RF SoC, using a heterodyne architecture would require a dedicated mixer per converter that entails a dedicated oscillator to push each low-IF channel from the DAC output to a unique IF frequency (and vice versa at the receiver side). Using a homodyne architecture allows the use of one oscillator per every two DAC channels, which is half the number of oscillators compared to a heterodyne architecture. The cost is the need for quadrature local oscillator (LO) signals. which demands higher-output power oscillators (which can be overcome by amplifying the LO output) and performance degradation due to IQ imbalance (which can be compensated for in DSP). The example design in Figure 3 employs the homodyne architecture. The homodyne/direct-conversion configuration for each baseband channel can be realized by employing an IQ mixer, which can upconvert the analog baseband IQ channels to a desired IF center frequency ω_k, $k \in \{1, 2, 3, 4\}$ (for the case illustrated in Figure 3). If the kth digital baseband channel is denoted as $p_k(t) = i_k(t) + j \cdot q_k(t)$, then the baseband bandwidth of $p_k(t)$ can be up to 4 GHz when using Gen-1 RF SoC devices.

An IQ mixer associated with each channel acts as a quadrature modulator realizing $\mathrm{Re}\{p(t) \cdot e^{j\omega_0 t}\}$. The ω_k frequencies must be chosen such that enough guard band is accommodated between the channel such that interchannel harmonic distortions and overlapping leakage are minimized. The upconverted frequency-multiplexed channels are combined to form an aggregated IF signal, which is applied to the front ends. At the receiver side, a quadruplexer (for a four-channel design, or similarly, a diplexer for a two-channel design) can be used to extract the different frequency-multiplexed channels. A simple power splitter can also be used (which will suffer from additional power-division loss based on the number of channels used.) The quadruplexer must be custom designed based on the correct channel frequency bands [17], and the implementation can be challenging depending on selection of the IF frequency bands.

The use of combiners to aggregate channels at different IF frequencies can pose multiple challenges:

combining N_{ch} channels needs a total of $N_{ch}-1$ two-port combiners (assuming N_{ch} is a power of two). The insertion losses in the combiner network can lead to reduced link margins. The realization of a wideband combiner in excess of 8 GHz of bandwidth is rather nontrivial. Such wideband combiners can possess gain variations in the frequency response, which in turn affect frequency selectivity of the input channels. Isolation among the input ports is another important factor that must be considered during design. Multiway combiners can be implemented as either passive planar or hybrid transformer-based circuits, which can at times be bulky. A large number of channels in a printed circuit board-level implementation can be physically large

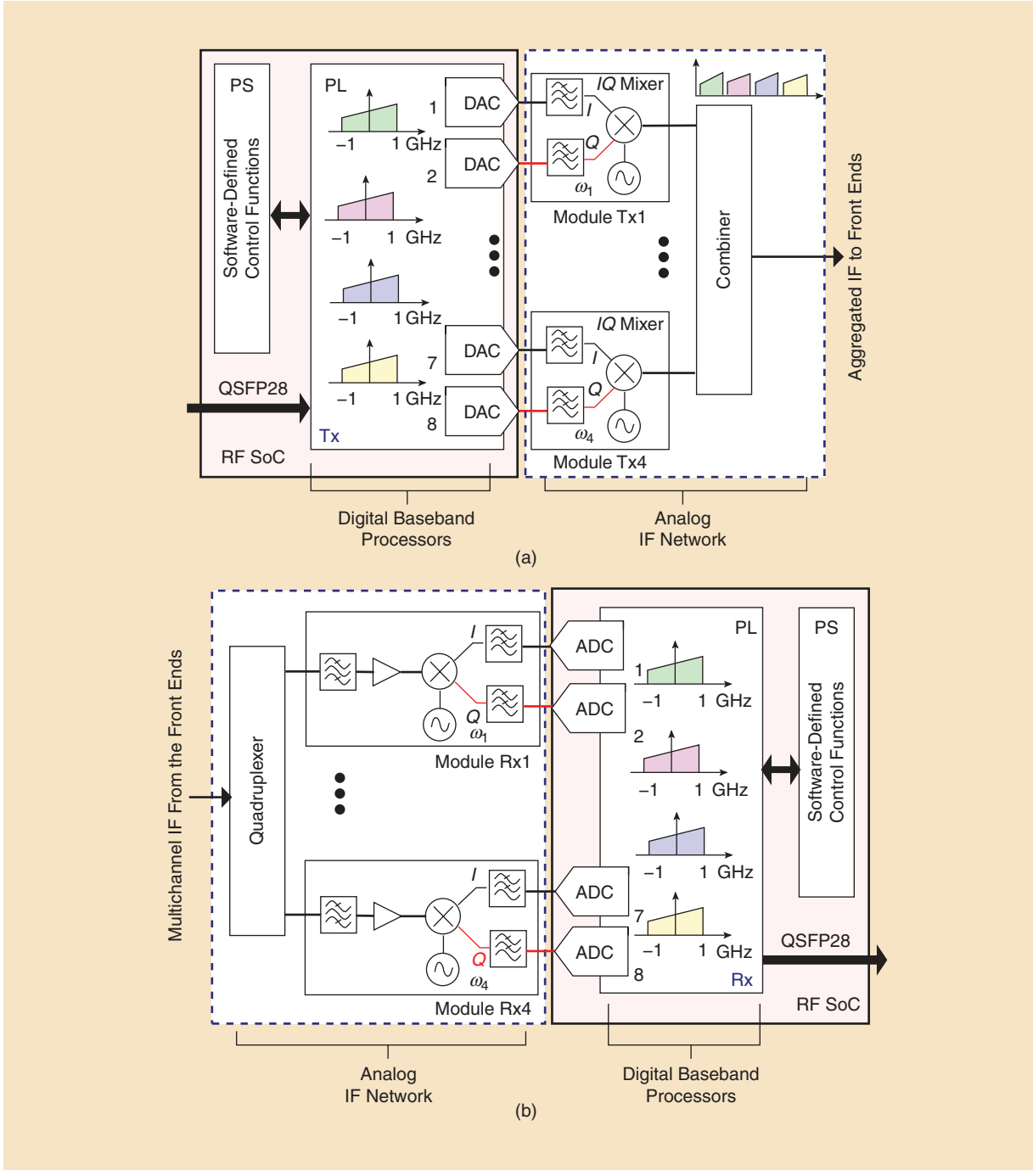

Figure 3. *(a) An overview of the Xilinx RF SoC-based real-time digital back-end architecture and the IF circuitry at the transmitter side. (b) A possible realization of the receiver side involving the IF demultiplexing circuitry and the baseband receiver logic for simultaneous processing of each baseband channel. PS and PL stand for processor subsystem and programmable logic (FPGA fabric), respectively. Tx: transmitter; Rx: receiver: I: in phase.*

and thus take up significant real estate within a system. Miniaturization using custom-designed, active combiner circuits can be considered for integrated circuit implementations, which can in turn lead to improved performance at reduced size, weight, and power. Moreover, resistive elements in wideband combiners (e.g., Wilkinson type) can be lossy when fed by RF signals at different frequencies, which can further degrade the loss. Further, we note that scaled-impedance designs can be considered for integrated circuits where custom impedance levels can be maintained at various locations on the circuit.

The simultaneous baseband digital processors are implemented in PL. The use of RF SoCs allows the processor subsystem (PS) to be used for a slow update-rate dynamic configuration, and to perform software-defined control of the PHY, as indicated in Figure 3. Such functions include 1) adapting the modulation index (per channel or per subcarrier) and 2) dynamically turning on/off channels/subcarriers depending on the THz channel-state information. Optical Ethernet interfaces such as Quad SFP28, supporting data rates up to 100 Gb/s, can be utilized for high-speed transfer of data to and out of the multichannel processors, thus meeting data rate demands.

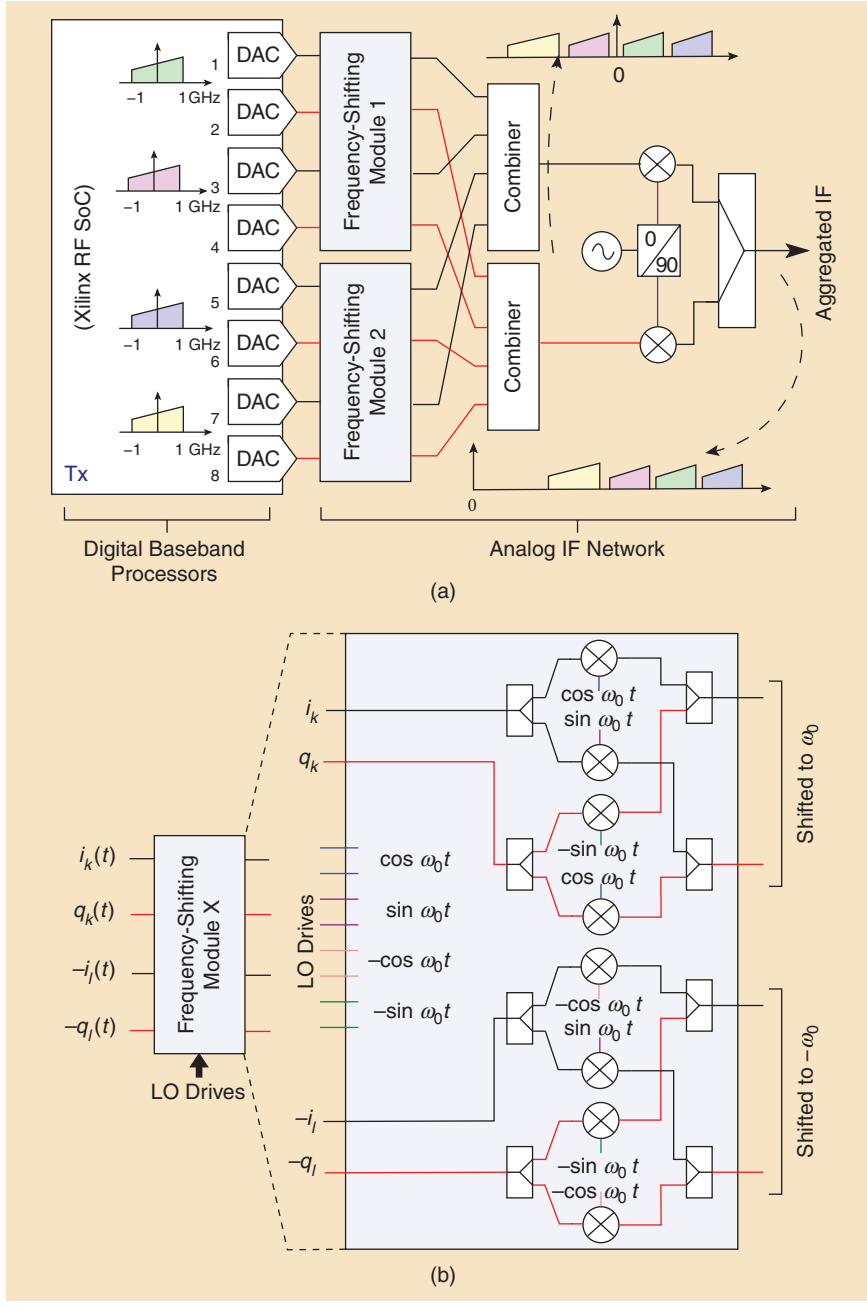

Figure 4. *A reduced oscillator design for frequency-multiplexed signal processing backends. (a) An overview of the architecture of such a four-channel system and (b) the frequency-shifting module architecture.*

Low-Complexity Design Using Fewer Oscillators

The use of IQ mixers in an FM-channelized system allows one oscillator per channel, thereby halving the number of oscillators compared to a low-IF heterodyne approach. The number of oscillators required can be further reduced by another factor of two by maintaining both IQ signal paths in analog throughout the IF stage. This comes at the cost of increased circuit complexity and high-precision electronics to maintain Q signal paths. As illustrated in Figure 4, two baseband channels can be shifted to ω_0 or $-\omega_0$ in analog by realizing $p_l(t) \cdot e^{\pm j\omega_0 t}$ by exploiting access to the Q components, thus enabling the same oscillator to upconvert two baseband channels to $+\omega_0$ and $-\omega_0$ frequencies, respectively. This leads to a reduced number of oscillators in the design. Once the channels are frequency shifted, the I and Q components need to be combined separately. The combined I and Q signals can then be upconverted to the desired RF center frequency using a wideband Q modulator (or an IQ mixer).

Figure 4(a) shows the architecture of a four-channel reduced oscillator backend where two DACs are used for each baseband channel. The "frequency-shifting modules" can be used to implement the $p_k(t) \cdot e^{+j\omega_0 t} + p_l(t) \cdot e^{-j\omega_0 t}$ operation, and the internal architecture is shown in Figure 4(b). Here, the different phases of the LO signals that are needed to shift two channels to $\pm\omega_0$ must be externally generated. This can be achieved in the digital domain (e.g., using RF SoCs, especially because the required LO tones will only be a few GHz) or using a passive, multiway microwave splitter. Note that realization via splitters and combiners needs careful design for low insertion loss and high phase coherency from input to output. A similar approach can be used at the receiver side to use a single oscillator to downconvert corresponding channels at $\omega_{IF} \pm \omega_0$ to baseband.

RF SoC-Based Real-Time Prototype Supporting 8 GHz of Bandwidth at the D-Band

We engineered a real-time four-channel transmitter (following the system in Figure 3) capable of handling 8 GHz of bandwidth at a maximum bit rate of 12 Gb/s per channel. The design is based on the HTG-ZRF8-R2 board [18], which features a Xilinx ZU28DR RF SoC. Polyphase digital cores were developed to generate 2.048 GHz of bandwidth per baseband channel, supporting an OFDM-based PHY.

Choice of PHY Parameters

The 100-plus-GHz wireless channel itself is not necessarily frequency selective as the channel is mostly sparse (due to less multipath propagation). Our 100-plus-GHz front ends [8] operate in between absorption lines, and thus, there is no significant frequency selectivity. Most of the frequency selectivity comes from the frequency response of ultrabroadband front-end devices and the IF RF chain electronics due to the large bandwidths. An OFDM-based PHY was selected for the prototype, thereby simplifying equalization at the receiver.

The maximum sample rates supported by the RF SoC chip used are different for DACs (6.554 GSps) and ADCs (4.096 GSps). The minimum of the maxima is chosen as the system's sample rate. This allows 2.048 GHz of real bandwidth per data converter (4.096 GHz of complex baseband bandwidth). But use of the full Nyquist rate is challenging as it increases the demands on the analog anti-imaging and anti-aliasing filters, therefore, two-times interpolation at the transmitter and two-times decimation at the receiver was used. This configuration allows a total of 2.048 GHz × 4, which equals a 8.192-GHz bandwidth. The number of subcarriers in the OFDM setup was set at 64 for each 2-GHz-wide channel, which corresponds to a subcarrier spacing of 32 MHz, wide enough to consider the channel to be frequency flat for our front ends. The frame-structure used resembles the 802.11a PHY due to the use of 64-point FFTs.

To avoid the sideband overlapping at the receiver, a high-side LO scheme was employed at the downconverter by setting the LO to generate an effective drive of 146 GHz.

Experimental Prototype

Figure 5(a) shows the experimental setup, consisting of an RF SoC-based DSP and the D-band experimental setup at 120–140 GHz. The LO frequencies ω_k were selected as 3.5, 6, 9.5, and 12.5 GHz for the four channels. The 120–140-GHz upconverter employs a four-times LO multiplier, and therefore, the upconverter LO input was set to 32.5 GHz, which translates to a 130-GHz LO. The upconverter does not incorporate an RF bandpass filter, and thus, it transmits both the sidebands where the upper sideband is placed from 132.5 GHz to 143.5 GHz; whereas the lower sideband is placed from 116.5–127.5 GHz. At the receiver, the 120–140-GHz downconverter employs a four-times LO multiplier. To avoid the sideband overlapping at the receiver, a high-side LO scheme was employed at the downconverter by setting the LO to generate an effective drive of 146 GHz. Although the front ends are named as 120–140 GHz, the device responses are within 1–2 dB of maximum performance through 120–150 GHz [7].

Verification of the Transmitter-Side Implementation

Random data bits generated in real time on the PL were used for the transmission at all four channels. Although the DSP was designed to support up to 64 quadrature amplitude modulations per channel, quaternary phase-shift keying (QPSK) modulation was employed on all the channels for the experiment shown in Figure 5(a). The spectrum of the transmitted aggregated IF is shown in Figure 5(b). Modulation above QPSK is limited by the noise level in the commercially available off-the-shelf analog front end. The downconverted signals at the 120–140-GHz receiver heads were captured in the DSO and were then processed offline for validation of proper transmission at each channel. Figure 5(c) and (d) shows the software-processed constellation outputs of channel 1 and channel 4, respectively, out of the four frequency-multiplexed channels. The error vector magnitude (EVM) of the shown constellations is 11.5 and 12.4 dB, for Figure 5(c) and (d), respectively. The prototype supports 10-times-more bandwidth over current 5G

The learning/inference process necessitates stream processing of multichannel input data at hundreds of gigabits per second.

bandwidth targets and four times over IEEE 802.11 ad/ay specifications.

Open Issues and Research Opportunities

Up to this point, we have discussed the complexity associated with the processing of one data stream. Obviously, such complexity increases when multiple input/multiple output (MIMO) or even massive MIMO systems are considered. As mentioned previously, at lower frequencies or for systems with lower bandwidth requirements, RF SoCs have been leveraged to implement digital beamforming systems. Although the use of multiple data converters in FM systems prevents RF SoC ADC/DAC channels from being used for digital beamforming, this can be compensated through high-gain antennas and using passive-lens integrated antennas for fixed links, especially for backhaul applications. Connecting with approaches like low-power, passive reflect arrays can be used for applications where steerable beams are desired [19]. Still, in the ultimate reconfigurable platform, an ultrabroadband SDR and a programmable front end would be co-designed to support dynamic allocations of channels, including 1) all for FM, 2) all for MIMO, or 3) any combination in between.

Whether in a single channel or in a MIMO system, the expansion of baseband bandwidths to dozens of GHz opens new computational challenges for wireless algorithms that, conventionally, are assumed to be operating on-edge software applications. One example is channel estimation and equalization. Because the adoption of upper-mm-wave and (sub-) THz carriers imply half wavelengths on the order of a millimeter, it leads to coherence times on the order of a few microseconds. It is expected that custom accelerators in the form of edge-computing digital

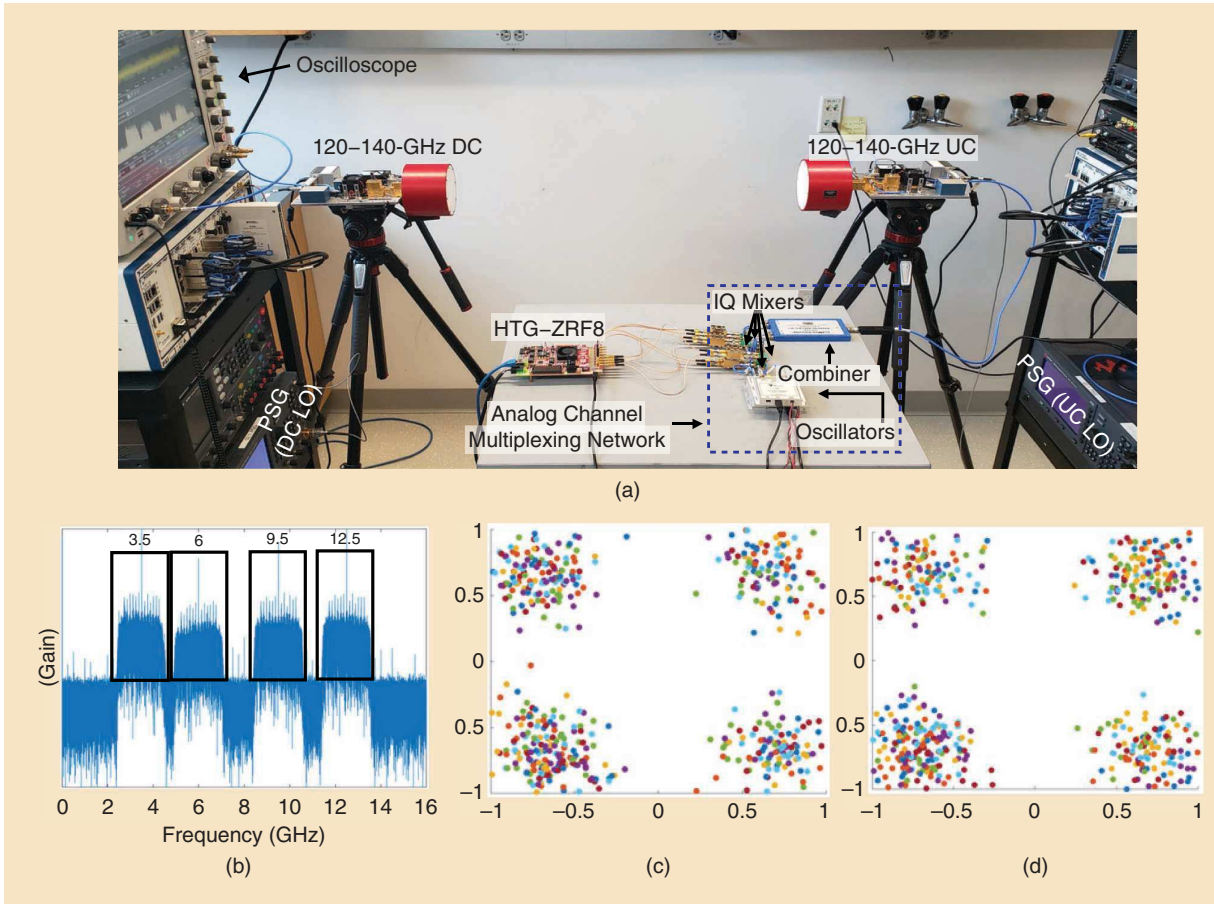

Figure 5. *(a) The Xilinx RF SoC-based prototype at the D-band consisting of the real-time transmitter supporting 8 GHz of bandwidth. (b) The transmitted four channels in the frequency domain. (c) and (d) The recovered constellations corresponding to channels 1 and 4, respectively, generated by recording the IF output from the 120–140-GHz downconverter using a DSO. UC: upconverter; DC: downconverter.*

systems may be required to meet throughput demands as purely SDR approaches may not suffice.

Another challenge is artificial intelligence (AI) algorithms that require both machine learning inference and lifelong learning as the next-generation systems (e.g., 6G) operate with real data. The learning/inference process necessitates stream processing of multichannel input data at hundreds of gigabits per second. These data must be processed in a stream processor to apply them to deep learning engines such as deep belief networks, convolutional neural networks, and/or transformer-based algorithms. The real-time computational throughput will likely exceed the capacity of available AI-at-edge accelerators, which are traditionally aimed at applications such as robotics, vision, and so on. In fact, we believe that the wireless industry can benefit from custom digital hardware accelerators that are deployed at the 6G radio edge.

Conclusion

Although next-generation wireless networks eye 100-plus-GHz carrier frequencies when searching for large contiguous bandwidths, there has not been much discussion on the DSP of such ultrabroadband signals. This article discussed the issues pertaining to real-time DSP backends that process ultrabroadband signals, along with their associated challenges. The article also presented one possible approach to real-time SDRs for ultrabroadband systems, one which uses a multichannel FM strategy and a Xilinx RF SoC to process parallel channels in real time. Using a prototype of the proposed system, real-time data transmission over 4 × 2 GHz of bandwidth at the D-band was demonstrated, which is capable of achieving a ≈12 Gb/s transmission rate per channel (48 Gb/s across all channels). This platform opens the door to experimental testing of the innovative communication and networking solutions that are needed to overcome the challenges and exploit opportunities for wireless communications beyond 100 GHz.

Acknowledgment

This work was supported by U.S. Air Force Research Laboratory Grant FA8750-20-1-0200. Viduneth Ariyarathna thanks Michele Polese (Institute for the Wireless Internet of Things, Northeastern University) for the insightful discussions about O-RAN.

References

[1] I. F. Akyildiz et al., "6G and beyond: The future of wireless communications systems," *IEEE Access*, vol. 8, pp. 133,995–134,030, Jul. 2020, doi: 10.1109/ACCESS.2020.3010896.

[2] I. F. Akyildiz et al., "Terahertz band communication: An old problem revisited and research directions for the next decade," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 4250–4285, Jun. 2022, doi: 10.1109/TCOMM.2022.3171800.

[3] T. S. Rappaport et al., "Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond," *IEEE Access*, vol. 7, pp. 78,729–78,757, Jun. 2019, doi: 10.1109/ACCESS.2019.2921522.

[4] K. Sengupta et al., "Terahertz integrated electronic and hybrid electronic–photonic systems," *Nature Electron.*, vol. 1, no. 12, pp. 622–635, Dec. 2018, doi: 10.1038/s41928-018-0173-2.

[5] S. Ju et al., "Millimeter wave and sub-terahertz spatial statistical channel model for an indoor office building," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1561–1575, Jun. 2021, doi: 10.1109/JSAC.2021.3071844.

[6] J. M. Eckhardt et al., "Measurements in a real data centre at 300 GHz and recent results," in *Proc. 13th Eur. Conf. Antennas Propag. (EuCAP)*, 2019, pp. 1–5.

[7] P. Sen and J. M. Jornet, "Experimental demonstration of ultrabroadband wireless communications at true terahertz frequencies," in *Proc. IEEE 20th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, 2019, pp. 1–5, doi: 10.1109/SPAWC.2019.8815595.

[8] P. Sen, D. A. Pados, S. N. Batalama, E. Einarsson, J. P. Bird, and J. M. Jornet, "The TeraNova platform: An integrated testbed for ultra-broadband wireless communications at true terahertz frequencies," *Comput. Netw.*, vol. 179, Oct. 2020, Art. no. 107370, doi: 10.1016/j.comnet.2020.107370.

[9] S. Abu-Surra et al., "End-to-end 140 GHZ wireless link demonstration with fully-digital beamformed system," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, 2021, pp. 1–6, doi: 10.1109/ICCWorkshops50388.2021.9473600.

[10] "Introduction to the NI mmWave transceiver system hardware," National Instruments, Austin, TX, USA, White Paper, 2019. [Online]. Available: http://www.ni.com/product-documentation/53095/en/

[11] "M8199A arbitrary waveform generator," Keysight Technologies, Santa Rosa, CA, USA, 2020. Accessed: Nov. 30, 2020. [Online]. Available: https://www.keysight.com/us/en/assets/3120-1465/data-sheets/M8199A-128-256-GSa-s-Arbitrary-Waveform-Generator.pdf

[12] D. Siafarikas and J. L. Volakis, "Toward direct RF sampling: Implications for digital communications," *IEEE Microw. Mag.*, vol. 21, no. 9, pp. 43–52, Sep. 2020, doi: 10.1109/MMM.2020.2998422.

[13] L. Bonati et al., "Open, programmable, and virtualized 5G networks: State-of-the-art and the road ahead," *Comput. Netw.*, vol. 182, Dec. 2020, Art. no. 107516, doi: 10.1016/j.comnet.2020.107516.

[14] "ZYNQTM Ultrascale+™ product selection guide," Xilinx, San Jose, CA, USA, 2020. Accessed: Jan. 2020. [Online]. Available: https://www.xilinx.com/support/documentation/selection-guides/zynq-usp-rfsoc-product-selection-guide.pdf

[15] M. Fosberry and M. Livadaru, "Digital synthetic receive beamforming with the Xilinx ZC1275 evaluation board," in *Proc. IEEE Int. Symp. Phased Array Syst. Technol. (PAST)*, 2019, pp. 1–2, doi: 10.1109/PAST43306.2019.9020971.

[16] V. Ariyarathna et al., "Real-time digital baseband system for ultrabroadband THz communication," in *Proc. IEEE 45th Int. Conf. Infrared, Millimeter, Terahertz Waves (IRMMW-THz)*, 2020, pp. 1–2, doi: 10.1109/IRMMW-THz46771.2020.9370733.

[17] S.-J. Zeng et al., "Compact and high-isolation quadruplexer using distributed coupling technique," *IEEE Microw. Wireless Compon. Lett.*, vol. 21, no. 4, pp. 197–199, Apr. 2011, doi: 10.1109/LMWC.2011.2109702.

[18] "HTG-ZRF8: Xilinx Zynq® UltraScale+™ RFSoC development platform," HiTech Global, San Jose, CA, USA, 2020. Accessed: Aug. 2020. [Online]. Available: http://www.hitechglobal.com/Boards/Zynq_RFSoc.htm

[19] N. M. Monroe et al., "Electronic THz pencil beam forming and 2D steering for high angular-resolution operation: A 98 × 98-unit 265GHz CMOS reflectarray with in-unit digital beam shaping and squint correction," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, 2022, vol. 65, pp. 1–3, doi: 10.1109/ISSCC42614.2022.9731671.