Towards 6G Hyper-Connectivity: Vision, Challenges, and Key Enabling Technologies

Howon Lee, Byungju Lee, Heecheol Yang, Junghyun Kim, Seungnyun Kim,
Wonjae Shin, Byonghyo Shim, and H. Vincent Poor

Abstract—Technology forecasts anticipate a new era in which massive numbers of humans, machines, and things are connected to wireless networks to sense, process, act, and communicate with the surrounding environment in a real-time manner. To make the visions come true, the sixth generation (6G) wireless networks should be *hyper-connected*, implying that there are no constraints on the data rate, coverage, and computing. In this article, we first identify the main challenges for 6G hyper-connectivity, including terabits-per-second (Tbps) data rates for immersive user experiences, zero coverage-hole networks, and pervasive computing for connected intelligence. To overcome these challenges, we highlight key enabling technologies for 6G such as distributed and intelligence-aware cell-free massive multi-input multi-output (MIMO) networks, boundless and fully integrated terrestrial and non-terrestrial networks, and communication-aware distributed computing for computation-intensive applications. We further illustrate and discuss the hyper-connected 6G network architecture along with open issues and future research directions.

Index Terms—6G hyper-connectivity, cell-free massive MIMO, non-terrestrial networks, over-the-air distributed computing, scalable deep learning, terahertz communications.

I. INTRODUCTION

MOVING toward the post-pandemic world, we are experiencing a new, previously inconceivable change in our lives. A large part of our daily activities are going online, and we are witnessing numerous use cases in shopping, mobility,

H. Lee is with the School of Electronic and Electrical Engineering and IITC, Hankyong National University, Korea, email: hwlee@hknu.ac.kr.

B. Lee is with the School of Electronic Engineering, Incheon National University, Korea, email: bjlee@inu.ac.kr.

H. Yang is with the Division of Computer Convergence, Chungnam National University, Korea, email: hcyang@cnu.ac.kr.

J. Kim is with the Department of Artificial Intelligence, Sejong University, Korea, email: j.kim@sejong.ac.kr.

W. Shin is with the Department of Electrical and Computer Engineering, Ajou University, Korea, and Princeton University, NJ, USA, email: wjshin@ajou.ac.kr.

S. Kim and B. Shim are with the Department of Electrical and Computer Engineering and the Institute of New Media and Communications, Seoul National University, Korea. email: snkim@islab.snu.ac.kr, bshim@snu.ac.kr.

H. V. Poor is with the Department of Electrical and Computer Engineering, Princeton University. USA. email: poor@princeton.edu.

Co-first authors: H. Lee, B. Lee, H. Yang, and J. Kim.

W. Shin and B. Shim are the corresponding authors.

healthcare, and manufacturing. As a platform to embrace this dramatic change, the fifth generation (5G) wireless systems have been standardized and commercialized in 2019, and as of now, 5G services have been rolled out in more than 70 countries. Notwithstanding the many technological advances brought by 5G, there are still many hurdles and challenges in realizing an envisioned boundless and hyper-connected society where intelligent things such as machines, vehicles, sensors, and robots are seamlessly connected both in physical and virtual spaces without limitation on the data rates and/or transmission delays [1], [2]. To back up these disruptive changes and therefore make the slogan a reality, we need to prepare the sixth generation (6G) of wireless networks by identifying key enabling technologies that cannot be achieved through a simple evolution of 5G.

One important expectation in the 6G era is that machines and things will be the main consumers of mobile data traffic, and thus there will be more than 55 billion devices connected to the Internet by 2025. These devices will continuously sense, process, act, and communicate with the surrounding environment, generating more than 73 zettabytes of data per year [3]. Since 6G wireless networks should support a variety of form factors with different service requirements and hardware capabilities, enabling technologies should be dearly distinct from those of 5G in various aspects and ultra-wideband contiguous bandwidth will be required [4], [5].

The primary purpose of this article is to discuss key challenges arising from these considerations. These include the provision of terabits-per-second (Tbps) wireless connectivity, zero coverage-hole networks, and pervasive computing for connected intelligence, along with enabling technologies that can be used as a core for the 6G system. The key concepts of 6G hyper-connectivity are illustrated in Fig. 1.

After the introduction, the rest of this article is organized as follows. In Section II, we address the challenges facing 6G hyper-connectivity. In Section III, we present the key enabling technologies for realizing this 6G vision. In Section IV, we discuss future research issues and provide our concluding remarks.

II. CHALLENGES AND PROSPECTS FACING 6G HYPER-CONNECTIVITY

A. Tbps for Immersive User Experience

Since the target of 5G is to support 20 gigabits-per-second (Gbps) peak data rate and 0.1 Gbps user-experienced
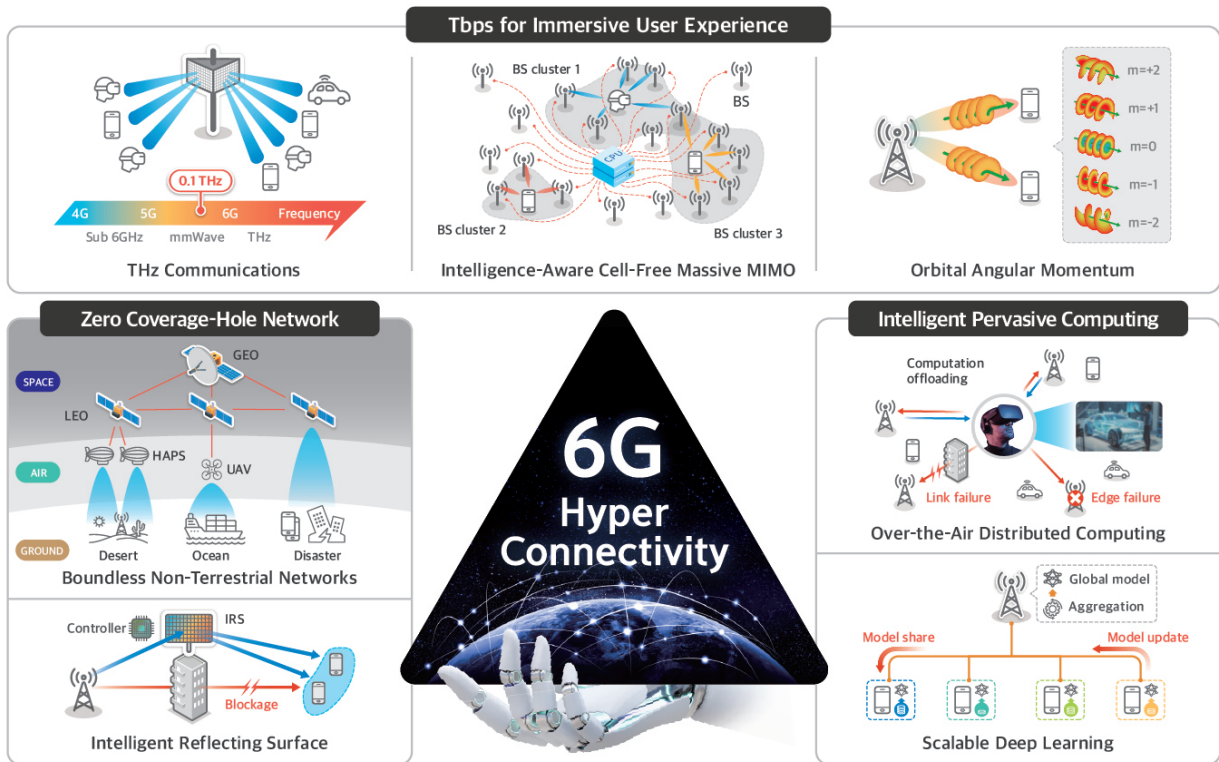
Fig. 1. Key concepts of 6G hyper-connectivity.

data rate, it is not too difficult to support advanced multimedia services such as 8K augmented reality (AR) streaming, 16K virtual reality (VR) streaming, and extended reality (XR) requiring beyond-gigabit data rates. It is however not easy to support truly immersive mobile services such as digital twin or metaverse since the virtual model accurately reflecting the physical world using XR devices or high-fidelity mobile holographic displays requires a rate of up to Tbps [2]. For example, a full-view 24K VR transmission with a 64 pixels per degree (PPD) minimum resolution, a 200 Hz minimum refresh rate, and a 300:1 compression rate requires at least 6.37 Gbps. Clearly, this level of rate cannot be supported anytime anywhere in the current 5G systems since the typical user-experienced data rate of 5G is in the range of 0.1–1 Gbps.

One way to support emerging services requiring tremendous data rates is to exploit the terahertz (THz) frequency band. The THz band offers a colossal spectrum but the THz waves suffer high attenuation due to the attendant significant path loss, molecular absorption, and atmospheric attenuation. One straightforward option to overcome these problems is to densify the network, meaning that we reduce the distance between the base station (BS) and mobile device. In this ultra-dense scenarios, pencil beamforming via ultra-massive multi-input multi-output (UM-MIMO) plasmonic nano-antenna arrays can be useful in providing 10 to $100\times$ improvement in data rate [6]. While the densified small cells equipped with massive multi-input multi-output (MIMO) systems are a good fit for THz communication systems, their deployment will make the notions of cell and handover obsolete.

Thus, we need to design the cell-free networks from scratch by considering various technical challenges, such as beam tracking/management, user association and BS coordination, and synchronization and initial access. Other key enabling technologies to support Tbps data rates include orbital angular momentum (OAM) multiplexing [7], full-duplex, and spatial modulation-MIMO (SM-MIMO).

B. Zero Coverage-Hole Networks

Another more important goal of 5G systems is to provide advanced wireless connectivity for a wide variety of vertical applications including self-driving cars and drones, robot-aided smart factories and healthcare, and remote control for disaster scenarios. To accommodate such applications, wireless access should be available anywhere, meaning that there should be no coverage holes. Although many network operators claim nationwide coverage for 5G, there are still many coverage holes in places such as tunnels, basements, elevators, mountains, and oceans. Moreover, higher frequencies (3.5 GHz at FR1 and 28 GHz at FR2) used in 5G systems tend to be vulnerable to blockage owing to their high directivity and severe path loss. For these reasons, more often than not we experience link failure in our daily use of 5G services.

Recently, several new approaches to enhance coverage and ensure universal and seamless connectivity have been proposed. An intelligent reflecting surface (IRS), an approach that provides an artificial path in the absence of line-of-sight (LOS) links, is a viable candidate to eliminate coverage holes, particularly in dense urban areas [8]. Wave-

controlled holographic MIMO surfaces that flexibly shape electromagnetic waves using sub-wavelength metallic or dielectric scattering particle structures is another good option [9]. In remote areas such as deserts, oceans, and mountain areas, in which the deployment of terrestrial networks (TNs) is almost impossible, non-terrestrial networks (NTNs) with satellites, high-altitude platform stations (HAPSs), and unmanned aerial vehicles (UAVs) can be an attractive solution because they ensure coverage without relying on the ground infrastructure. By the deliberate combination of TNs and NTNs [10], one can dramatically improve the coverage and thus generate the network with virtually no coverage hole.

C. Pervasive Computing for Connected Intelligence

With the advent of truly immersive mobile services, such as XR and metaverse, computational overhead has become a serious concern since these services require high energy consumption, end-to-end latency, and processor and memory overhead [11]. In 5G, mobile edge computing (MEC) has been widely used to delegate computation-intensive tasks to the edge (a local server near a BS). However, MEC would not be enough to support future applications. In a VR service, for example, a mobile VR headset must provide photorealistic and 6 degrees of freedom (6-DoF) rendered images within 20 ms of motion-to-render-to-photon (M2R2P) latency. To control this overhead solely in a mobile device or offload it to a nearby BS will be by no means easy so that new paradigm to deal with the surge of computation is required. Fortunately, as the density of wireless networks increases and beyond-gigabit transmissions can be supported using mmWave and THz band transmissions, one can exploit computational resources in nearby BSs and mobile devices that are equipped with high-performance computing processors. For instance, a vehicle equipped with a high-performance battery (1000 Wh/L) and powerful graphics processing units (GPUs) can perform environment sensing, learning, and computation-oriented tasks and send the results to the destination device. Particularly, computation offloading is useful for artificial intelligence (AI) applications, such as object detection and classification, speech recognition, and metaverse control, since these tasks usually require heavy deep neural network (DNN) computation, but the inference results are often very simple.

In order to facilitate the boundless use of computation resources over the network, one should use advanced wireless distributed computing techniques. Decentralized learning obtaining a global model from localized data via peer-to-peer communication among mobile devices can be one good choice. When security matters, federated learning where the mobile devices jointly train the global model without exchanging the local dataset with the central server would be a proper choice. Yet another option to support large-scale models is split learning, in which large-scale DNNs are divided into several pieces and then trained separately on mobile devices. For the data transmitted by mobile devices to BS, federated learning transfers the entire neural network model obtained locally, while split learning transfers the intermediate output of the divided parts of the neural network model.

III. KEY ENABLING TECHNOLOGIES FOR 6G HYPER-CONNECTIVITY

In this section, we discuss key enabling technologies to achieve beyond Gbps for immersive user experience, zero coverage-hole networks, and pervasive computing for connected intelligence.

A. Distributed & Intelligence-Aware Cell-Free Massive MIMO Network

In the 6G era, a network will be heavily densified and thus, the areal density of BSs will be comparable to, or even surpass, the density of active mobile devices. As the coverage of a cell gets smaller, the distance between BS and mobile will also be shorter. To make the most of the proximity between BSs and mobile devices, a network should be designed such that a group of small cells coherently serve users in a user-centric manner (see in Fig. 2(a)) [12]. In most urban scenarios, user-centric cell association will generate multiple LOS links between BSs and mobile so that the mobile can easily achieve a high degree of macro-diversity gain and spectral efficient improvement [13], [14].

To evaluate the efficacy of the cell-free massive MIMO network, we plot cumulative distribution functions (CDFs) of user signal-to-interference-and-noise ratio (SINR) in mmWave (100 GHz) and THz (1 THz) communication scenarios. In this figure, we consider the cell-free massive MIMO systems where $M = 128$ small cells, each of which is equipped with $N = 4$ antennas, cooperatively serve $K = 20$ mobiles using the same time-frequency resource. For comparison, we use the small cell systems where the small cells do not cooperatively serve the whole users but serve only the users in its cell coverage and the centralized massive MIMO systems where a single macro cell equipped with $N_{macro} = 512$ antennas serve the users. In every scenario under test, we fix the total transmit power of the network so that the overall energy consumption does not increase in the densified network. One can see from Fig. 2(b) that the SINR variation of the cell-free massive MIMO network is much smaller, meaning that the cell-free massive MIMO network can well suppress the inter-cell interference through the cooperative beamforming of BSs. For example, in the THz systems, the cell-free massive MIMO network can provide at least 7 dB SINR to more than 95% of users while the conventional small cell network guarantees only -2.5 dB. We also observe that the SINR gain of the THz band is higher than that for the mmWave band, meaning that it can effectively compensate for the huge path loss of the THz systems. To fully enjoy the benefits of cell-free massive MIMO, we need to address the following challenges.

Intelligence-aware network architecture: In order to provide the user-centric coverage through the joint beamforming of cooperating BSs, a cell control information, channel state information (CSI), and transmit data of local BSs should be shared with the central unit (CU) through the fronthaul link. Sharing such large information among all network entities will cause a significant fronthaul overhead, not to mention the tremendous computational overhead required for the channel
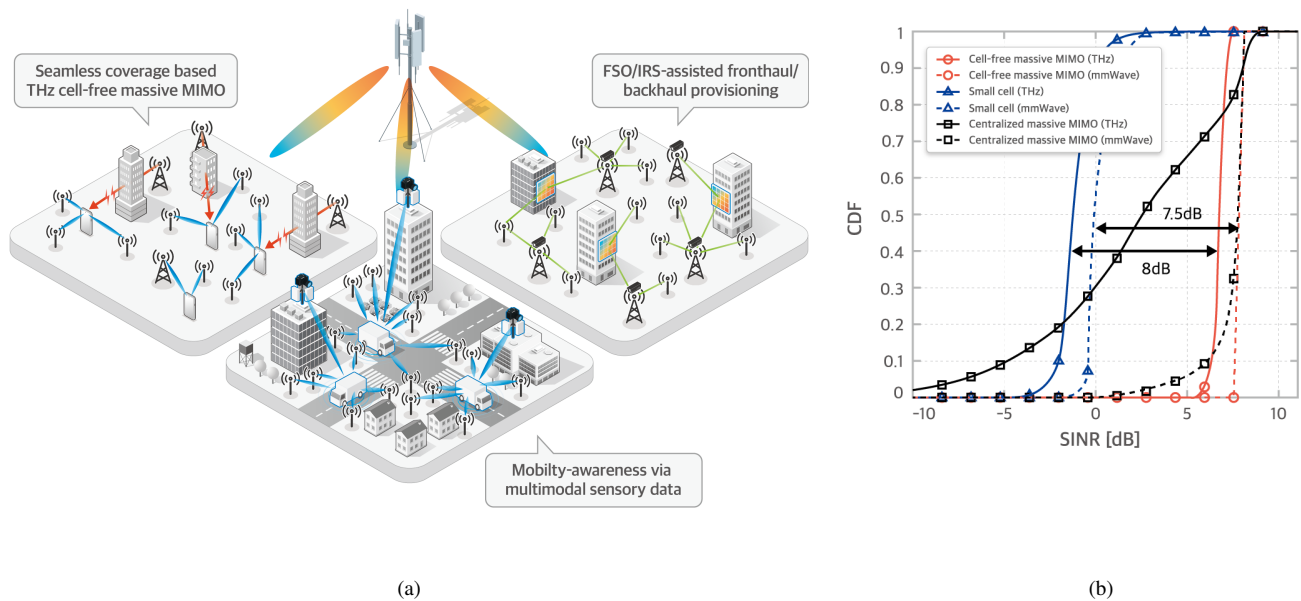
Fig. 2. (a) Illustrations of 6G cell-free massive MIMO network (the blue lines are the reliable links between the mobile and the densely deployed small BSs and the red lines are the blocked links between the mobile and the macro cells) and (b) CDFs of user SINR in cell-free massive MIMO networks, small cell, and centralized massive MIMO networks.

estimation, data transmission/reception, and resource allocation [15]. A promising approach to mitigate the overhead is decentralized processing. One such example is multi-agent deep reinforcement learning (MA-DRL) where the operations like beamforming, resource allocation, and user association are performed locally and individually by the DRL agent in each small cell.

Multi-sensory data fusion for channel prediction and beamforming: In order to realize a truly zero coverage-hole cell-free massive MIMO network, the network should have mobility-aware functions such as fast-fading channel estimation, mobile trajectory tracking, blockage prediction, and dynamic BS clustering and beamforming. One promising direction is to use the multi-modal sensory data (e.g., RGB depth vision information of camera, range information of LiDAR, and SLAM-based user location) along with the computer vision techniques to identify the 3D shape of nearby wireless environments (e.g., obstacle locations, LOS and non-line-of-sight (NLOS) paths, and mobility patterns). In this scheme, using a set of panoramic images obtained from the stereo camera, the DL-based object detection can extract the location (3D Cartesian coordinate (x, y, z)) of target objects in an image and then determine its class. In acquiring the location information of a mobile device from input images, a convolutional neural network (CNN) that extracts the spatially correlated feature using the convolution filter can be employed. Once the location of the mobile is identified, by measuring the angle and distance from the BS to the target object (e.g., mobile phone, vehicle, or urban air mobility (UAM)), the mmWave and THz beam can be generated without a complicated beam management process. Indeed, from the numerical evaluations based on the processing latency and power consumption of the latest Qualcomm

Snapdragon 888, recently introduced AI-focused system-on-a-chip (SoC), it has been shown that the image processing latency and the power consumption of deep learning (DL)-based beamforming scheme are much smaller than those of the conventional beamforming scheme in 5G NR [16]. By employing dedicated AI processors designed with a few nano-scale CMOS technology, we expect that the DNN processing latency can be further reduced down the road.

Computation-oriented end-to-end cell-free massive MIMO network: In order to support the computation-intensive and/or latency-critical applications, the future 6G network should have the computation offloading functionality [17]. Computation offloading can be realized by utilizing the local computing power at the BSs as well as cloud computing at the CU. For example, latency-sensitive computation tasks will be processed at the BSs in proximity to the user, and the computation-intensive tasks can be handled via the cooperation of BSs and CU.

To ensure successful computation offloading, an integrated cell-free network design exploring both communication and computational perspectives needs to be suggested. In general, it should be guaranteed that computing operations can be delivered by reliable and low-latency communication links through a user, BSs, and CU. From a computation perspective, edge servers located at the BSs may have limited computing capabilities in terms of computing power, memory size, and energy consumption compared to the central server at the CU. In addition, since the user needs to offload computation tasks to nearby BSs, the randomness in computation latency at each BS should be taken into account to guarantee low-latency and reliable computing services.

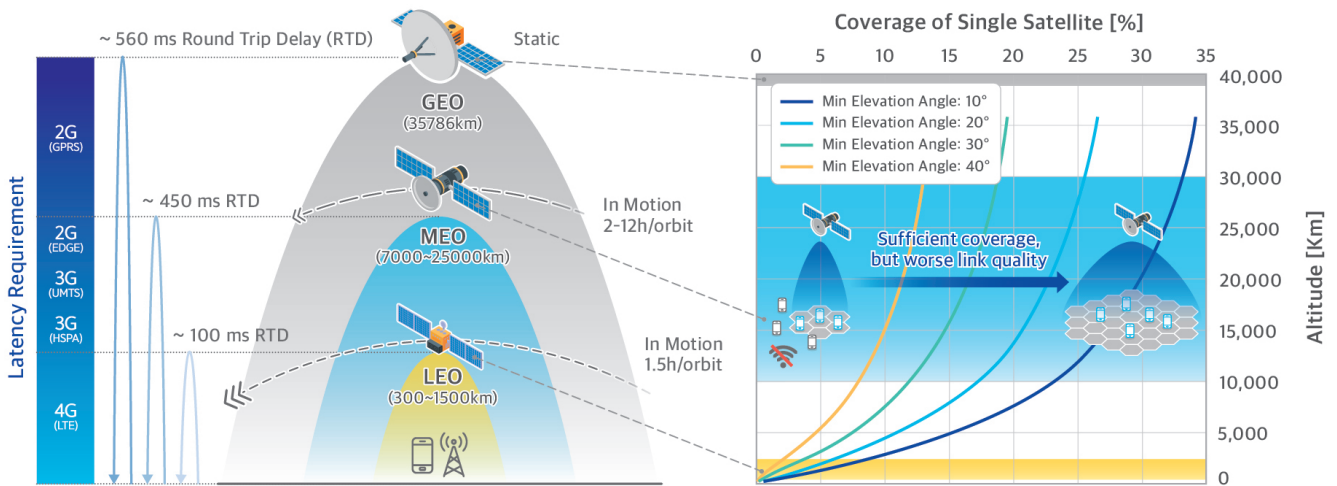**Integrated RF and optical X-haul provisioning:** In the

Fig. 3. Satellite communication systems with different altitudes: Round-trip delay and single satellite coverage as a fraction of the earth's surface.

cell-free massive MIMO network, fronthaul and backhaul overhead is a serious yet underrepresented problem. Recently, advanced fronthaul/backhaul techniques have been proposed. Replacing the current CPRI interface with the optical fiber-based fronthaul/backhaul is a simple option but will increase the CAPEX significantly. A radio stripe network-based fronthaul and backhaul where the BSs are serially connected using daisy-chain architecture to the same cables would be a cost-effective option for the densely deployed BSs. Also, an integrated access and backhaul (IAB) network where only a few BSs are connected to traditional fiber infrastructures while other BSs wirelessly relay the fronthaul/backhaul traffic would be a promising option. Yet another simple solution is free-space optical (FSO) communication where the modulated optical beams are transmitted through the free space. When the LOS link is guaranteed, FSO-based backhaul can offer several Gbps data rates with a much lower deployment cost than the optical fiber-based backhaul. When the LOS link is absent, optical IRS, a planar metasurface that controls the optical propagation environment by adjusting the phase shifts of reflecting elements can be used to offer the virtual LOS link between the BSs and CU.

Rank-sufficient THz LOS communications using cell-free massive MIMO network: In the THz band, due to the severe path loss and high directivity, the conventional MIMO techniques exploiting the rich scattering of multipath environments will not work. In the case where the LOS path is the dominant, the cell-free massive MIMO systems can be an appealing solution to address the rank-deficiency issue of THz communications. Specifically, through the joint transmission of small BSs, the cell-free massive MIMO systems can generate a virtual multipath propagation. Further, the shortened communication distance along with the extremely large number of transmit antennas incur a spherical wavefront which creates a rich non-linear pattern of phase variations and therefore, increases the channel rank.

B. Boundless and Fully Integrated Terrestrial & Non-Terrestrial Networks

In remote and rural areas where the existing terrestrial cellular network is unreachable [18], NTNs will be a useful means of providing cost-effective and high-speed internet services. Depending on the flying altitude and coverage area, NTN devices can be classified as UAV, HAPS, and satellite. Satellite stations are further divided into geostationary orbit (GEO), medium earth orbit (MEO), and low earth orbit (LEO) based on their orbital altitude.

In Fig. 3, we plot the round-trip delay (RTD) and single satellite coverage on the earth's surface for various altitude levels. When a satellite revolves in orbit close to the surface of the earth, a faster response time can be guaranteed, but the satellite's decreases owing to its proximity to the ground. For example, only three equally spaced GEO satellites are needed to can provide near-global coverage because each GEO satellite covers almost 35% of the earth's area, but RTD in this case exceeds 0.5 s. Conversely, the RTD of the LEO satellite is less than 0.1 s, but the coverage will be just 0.2% at an altitude of 550 km. One can infer from this discussion that a large number of LEO satellites, often called LEO mega-constellation, would be a good option to guarantee the latency in the order of a few tens of milliseconds and global coverage.

We note that there exists a trade-off between the ground observer's minimum elevation angle and link quality (see Fig. 3). If we increase the minimum elevation angle, so is the LOS probability of the satellite link; however, the beam coverage and visible time of the satellite will be reduced. To balance this trade-off, the ground station's minimum elevation angle should be carefully determined in practice. Recently, SpaceX decided to change the minimum elevation angle of Starlink LEO satellites from 40° to 25° (granted by the FCC in April 2021) to improve coverage with a small number of LEO satellites [19].

In analyzing the NTN communication of satellite and aerial devices, areal capacity ($bits/s/m^2$) is no longer use-

ful; therefore, a new metric called volumetric traffic capacity ($bits/s/m^3$) that measures the quality of user experience in 3D space is needed. To obtain a sufficient volumetric traffic capacity to ensure universal global connectivity for 6G systems, the following issues must be addressed.

Mobility-aware channel modeling and tracking in 3D space: Since the NTN terminals are located at a high altitude and are moving fast with respect to the ground stations, LOS propagation paths are guaranteed, but the wireless signals suffer significant Doppler shift. The first step in addressing the distortions caused by high mobility and a large coverage area is an accurate 3D channel model that accounts for the Doppler shift and propagation delay. In addition, dynamic channel prediction based on the real-time location tracking of NTN terminals is of great importance to achieve agile and effective NTN management.

Boundless multi-layered TN/NTN architecture: To provide sufficient and continuous coverage everywhere, the large LEO satellite constellations should be combined with GEO/MEO using the coverage-aware protocol and NTN-backhaul connection. In fact, a boundless multi-layered architecture fully integrating TNs and NTNs is essential to guarantee the reliable connection in the ocean and desert, as well as high-speed trains and airplanes [20]. UAV and HAPS play a key role in improving the connection quality because they can be used as relays between the LEO/MEO/GEO satellite and ground stations. In addition, an integrated network protocol that includes multi-layer routing, network mobility management, and resource allocation should be introduced.

Seamless inter-beam/satellite/orbit handover: Owing to the high-speed movement of the LEO satellites, the maximum visible time of the satellites at the ground station is at most 10–20 m, causing a frequent handover even for a fixed ground station. To mitigate this problem, LEO satellites should be connected to nearby satellites in the same or adjacent orbits through laser or radio frequency (RF) links. This so-called inter-satellite link (ISL) connectivity supports relatively low end-to-end latency even for the long distance communication scenario. Candidate techniques to establish an efficient ISL connectivity include spatial-mode diversity and multiplexing, wavelength division multiplexing, and adaptive beam control using pointing/acquisition/tracking (PAT) process. Another challenge of NTNs is that the cell size varies according to the elevation angle of the NTN terminals; therefore, conventional handover techniques might not be suitable. To address this challenge, a forecast-based handover decision on the basis of satellite position estimation is required.

3D cell planning and fast beam management: Since the cell size varies according to the altitude and elevation angle of the satellite, we need to introduce 3D cell planning based on the orbits of satellites. A satellite with a higher altitude and lower elevation angle tends to form a larger cell size owing to its wider beam footprint on the ground. Specifically, the GEO and LEO satellites have beam widths of up to 3,500 km and 1,000 km at an altitude of approximately 600 km, respectively. In addition, 3D cell types can be classified into earth-moving and earth-fixed cells with respect to the earth's surface, de-
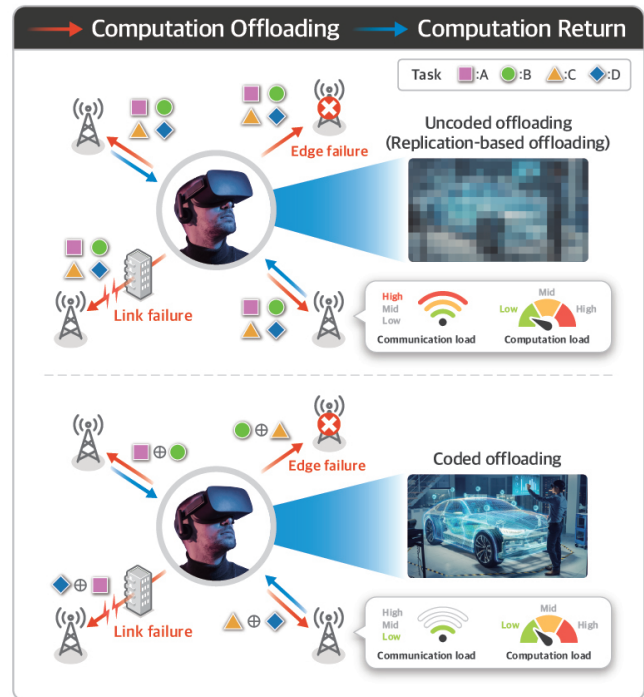

Fig. 4. Over-the-air distributed computing scenario where a VR headset assigns linear computation tasks to four edges with uncoded/coded offloading.

pending on whether the satellite beams can be fixed or steered. The earth-fixed cell should be designed to be highly robust to satellite beam pointing errors since even a very small error makes the satellite beam footprint placed tens of kilometers away from the initially designated area on the earth; whereas, the earth-moving cell may require more frequent handovers because of the dynamics of beam movement on the ground. As for the cell planning in multi-layered TN/NTN architecture, a truncated octahedron-shaped cell planning distinguished from the hexagon-shaped 2D cell layout is desired when considering the volumetric quotient in 3D space.

Intelligent spectrum management and NTN system optimization: To achieve seamless and broad coverage in integrated terrestrial and non-terrestrial networks, powerful spectrum management scheme is needed. To this end, intelligent and dynamic cooperation between primary and cognitive users is preferred over the semi-static spectrum sharing. Note that because the S-band (i.e., 2170–2200 MHz) is adjacent to the 3G/4G LTE band (i.e., 2110–2170 MHz), the large Doppler shift of the high-speed satellites causes severe interference in the adjacent band, deteriorating the service quality of 3G/4G TN services. A possible option to mitigate the interference is the on-board pre-compensation of the Doppler shift at the center of the beam on the ground while the residual Doppler shift can be corrected at each receiver side. To further improve spectral efficiency, we should have an appropriate frequency reuse pattern among multiple beams, available channels, and antenna gains.
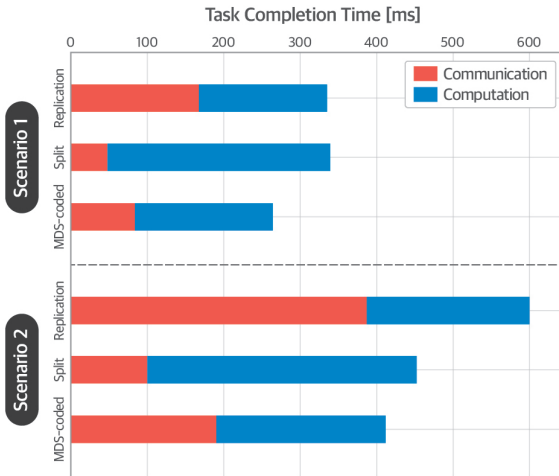
Fig. 5. Task completion times when a user offloads linear computation tasks to four edges over two different channel conditions (Scenario 1: LOS channel at a distance of 15 m; Scenario 2: NLOS channel at a distance of 300 m).

C. Over-the-air Distributed Computing for Artificial Intelligence Applications

A future trend in dealing with the relentless growth in computing-intensive applications is pervasive computing, in which mobile devices offload computational overhead to network edges or nearby devices. Notwithstanding the rosy outlook, a number of issues have to be addressed since there is a fundamental trade-off between computation and communication. In fact, when one tries to offload computational overhead, one should investigate the overhead of the communication link. Fig. 4 depicts computation offloading scenarios where a VR headset assigns linear computations to four edges with various offloading schemes [21]. Since the reduction in the communication load over the network increases the computation overhead at the edges, we need to come up with a proper offloading scheme optimized for the computation capability and network conditions. One thing to note is that the wireless channels experienced by devices and their computing capabilities are quite different among devices so that there is a large variation in the task completion time. Particularly, this problem is critical for the ultra-reliable and low-latency communications (URLLC) services since the task completion time is determined by the device with the worst channel condition or straggling/malfunctioning devices (e.g., edge failure in Fig. 4). For the seamless integration of communication and computing, the following approaches can be considered.

Communication-aware pervasive computation design: To alleviate the difficulties caused by the channel uncertainty, we need to exploit the mmWave and THz frequency bands. If sufficient bandwidth is provided and LOS channel conditions are guaranteed for every computing node, wireless links will not be a serious bottleneck. Obviously, this ideal condition cannot be guaranteed always, so we need to come up with an environment-aware and communication-friendly computing mechanism. For example, the communication load can be reduced by quantization, compression, and sparsification of the transmit information as long as it does not severely degrade the task quality. For special cases where the (weighted) average of the computation tasks is needed, we can design the system such that the computation results are added during the wireless transmission of the same time/frequency resources. This so-called over-the-air computation can significantly reduce the communication and computation overhead.

Collaborative and robust coded computation offloading: A naïve approach to exploit multiple computing nodes is to split the whole computation task into multiple pieces and then distribute them equally to multiple nodes. This split-based offloading will not be effective in the presence of link failure and straggling/malfunctioning devices, since the whole task can be completed only when all subtasks are finished. One way to avoid this problem is to replicate a part of computation tasks and assign them to the multiple devices.

One option to overcome difficulties caused by the channel heterogeneity, link failure, and straggling/malfunctioning is to apply coding techniques to distributed task allocation. By introducing redundancy in distributed tasks, coded distributed computing can be more resilient to the link and edge failures. For example, if computation tasks are designed using maximum distance separable (MDS) codes, one can recover the straggling tasks from other completed tasks [22]. In Fig. 5, we show the task completion times for offloading linear computation tasks over two different channel scenarios. We adopt the computation time and communication time models in [23] with a task (matrix) size $l = 1 \times 10^7$, computation intensities $\eta = [300, 500]$ (CPU cycles/bits), CPU frequency $f = 1 \times 10^9$ (Hz), transmission bandwidth $W = 2$ (MHz), channel noise power density $N_0 = 1 \times 10^{-19}$ (W/Hz), transmit power $P = 0.1$ (W), and an exponential computation delay with a straggling parameter $\mu = [1.8 \times 10^7, 1.5 \times 10^7]$, for two different scenarios, respectively. As illustrated, one can achieve the better computation-communication trade-off by assigning coded computation sub-tasks to computing nodes.

Scalable and over-the-air deep learning: To perform a large-scale deep learning (DL) task over wireless networks, we must decide how to allocate the tasks among the cloud, network edges, and the mobile devices [24]. For instance, network edges perform DNN training, and the mobile device performs the inference using the trained model from edges. To perform the flexible DL using abundant network resources, we can consider the following: 1) Federated learning, which trains local models on each device and then passes updated parameters to the central cloud for a global model; 2) distributed learning, in which each device communicates directly with its neighbors; and 3) split learning, in which the device or the central cloud trains each split model individually.

These DL techniques can be combined with over-the-air techniques [25] to effectively aggregate large-scale distributed data in ultra-low-latency scenarios. In addition, a scalable approach can be applied to resource-constrained mobile devices to reduce communication costs by leveraging gradient clustering, correction, and quantization methods.
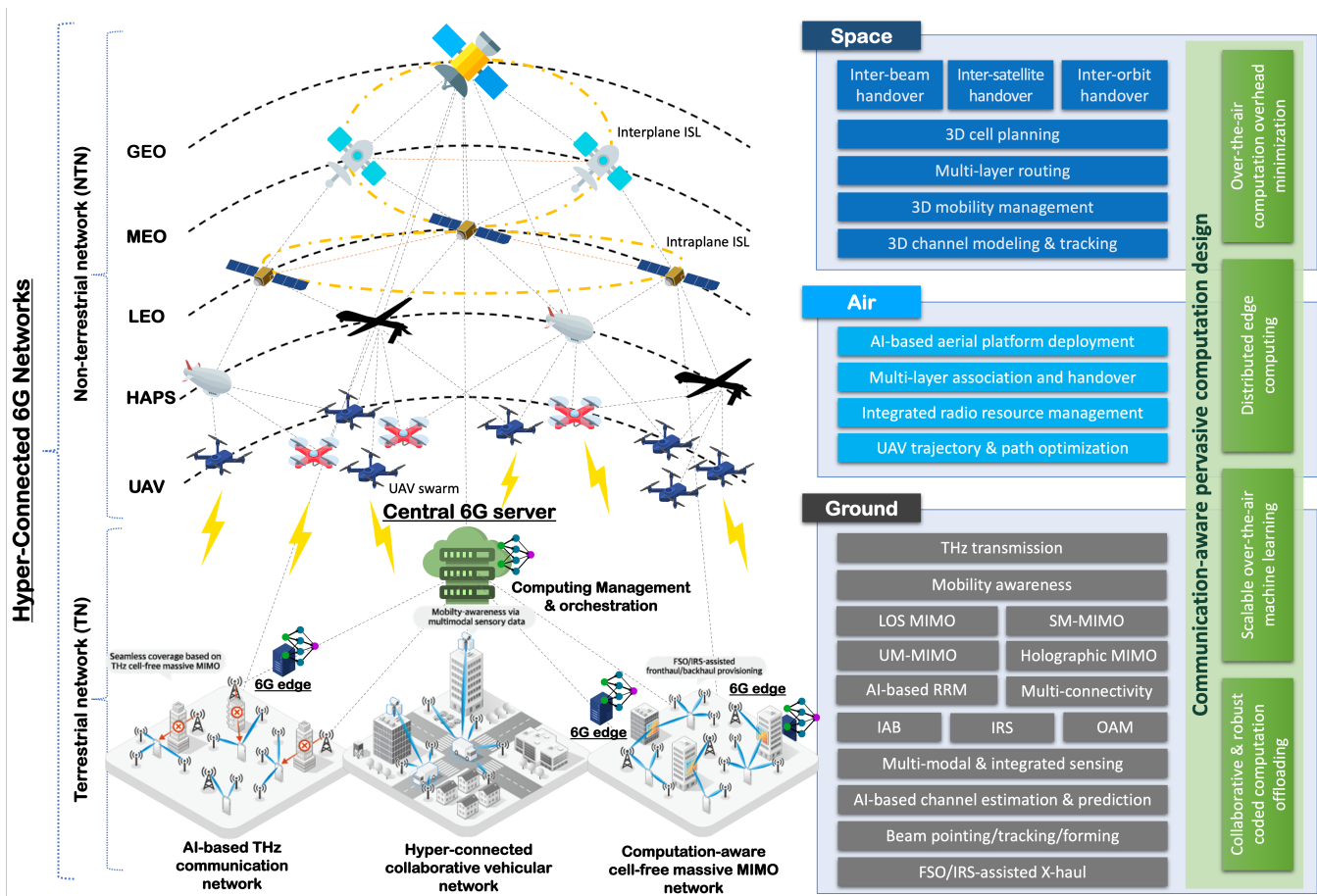
Fig. 6. Hyper-connected 6G network architecture.

IV. SUMMARY AND OUTLOOK

In this article, we discussed the vision and challenges to achieve the hyper-connected society in 6G. We have identified several key enabling technologies and their roles in supporting ultra-broadband, ubiquitous 3D, and computation-aware connectivity in 6G. With these disruptive technologies, 6G is envisioned to collapse the existing boundaries of space/air/ground networks, communication/computing/intelligence capabilities, as well as physical/virtual worlds. We anticipate that this will bring the truly *6G hyper-connectivity* society to humans, machines, and sensors in ubiquitous 3D global coverage (beyond ground-level 2D connectivity for 5G) even in space, oceans and desserts. In Fig. 6, we summarized the hyper-connected 6G network architecture by highlighting the key enabling technologies along with other remaining major challenges.

Going forward, achieving Tbps data rates, zero coverage holes, and pervasive computing for connected intelligence will contribute to reducing differences in social and regional infrastructures and economic opportunities, thereby ameliorating many social issues such as the digital divide, regional polarization, and education inequality. Our hope is that this article will spark further interest and expedite the research activities for 6G.

REFERENCES

[1] 6G Flagship, "Key drivers and research challenges for 6G ubiquitous wireless intelligence," Sep. 2019. Available: http://jultika.oulu.fi/files/isbn9789526223544.pdf

[2] Samsung Research, "6G white paper: The next hyper-connected experience for all," Jul. 2020. Available: https://research.samsung.com/

[3] IDC, "IoT growth demands rethink of long-term storage strategies, says IDC," Jul. 2020. Available: https://iotbusinessnews.com/2020/07/29/20898-iot-growth-demands-rethink-of-long-term-storage-strategies-says-idc/

[4] Docomo, White paper 5G evolution and 6G (Version 4.0)," Jan. 2022. Available: https://www.docomo.ne.jp/english/corporate/technology/whitepaper_6g/

[5] Samsung Research, "6G spectrum: expanding the frontier," May 2022. Available: https://cdn.codeground.org/nsr/downloads/researchareas/2022May_6G_Spectrum.pdf

[6] S. Tarboush *et al.,* "TeraMIMO: A channel simulator for wideband ultra-massive MIMO terahertz communications," *IEEE Veh. Technol. Mag.,* vol. 70, no. 12, pp. 12325–12341, Dec. 2021.

[7] D. Lee, H. Sakaki, H. Fukumoto, K. Hiraga, and T. Nakagawa, "Orbital angular momentum (OAM) multiplexing: An enabler of a new era of wireless communications," *IEICE Trans. Commun.,* vol. 100, no. 7, pp. 1044–1063, Jul. 2017.

[8] Y. Chen, Y. Wang, J. Zhang, P. Zhang, and L. Hanzo, "Reconfigurable intelligent surface (RIS)-aided vehicular networks : Their protocols, resource allocation, and performance," *IEEE Veh. Technol. Mag.,* vol. 17, no. 2, pp. 26–36, Jun. 2022

[9] C. Huang *et al.,* "Holographic MIMO surfaces for 6G wireless networks: Oportunities, challenges, and Trends," *IEEE Wireless Commun.,* vol. 27, no. 5, pp. 118–125, Oct. 2020.

[10] H. Guo, J. Li, J. Liu, N. Tian, and N. Kato, "A survey on space-air-ground-sea integrated network security in 6G," *IEEE Commun. Surveys Tuts.,* vol. 24, no. 1, pp. 53–87, 2022.

[11] L. Chang *et al.*, "6G-enabled edge AI for metaverse: Challenges, methods, and future research directions," *J. Commun. Inf. Netw.*, vol. 7, no. 2, pp. 107–121, Jun. 2022.

[12] S. Elhoushy, M. Ibrahim, and W. Hamouda, "Cell-free massive MIMO: A survey," *IEEE Commun. Surv. Tut.,*, vol. 24, no. 1, pp. 492–523, 2022.

[13] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[14] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4445–4459, Jul. 2017.

[15] S. Kim, J. Choi, and B. Shim, "Downlink pilot precoding and compressed channel feedback for FDD-based cell-free systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3658—3672, Jun. 2020.

[16] Y. Ahn *et al.*, "Towards intelligent millimeter and terahertz communication for 6G: Computer vision-aided beamforming," to appear in *IEEE Wirel. Commun.,* 2023.

[17] M. Zeng, W. Hai, O. A. Dobre, Z. Ding, and H. V. Poor, "Massive MIMO-assisted mobile edge computing: Exciting possibilities for computation offloading," *IEEE Veh. Technol. Mag.*, vol. 15, no. 2, pp. 31–38, Jun. 2020.

[18] D. Wang, M. Giordani, M.-S. Alouini, and M. Zorzi, "The potential of multilayered hierarchical non-terrestrial networks for 6G: A comparative analysis among networking architectures," *IEEE Veh. Technol. Mag.*, vol. 16, no. 3, pp. 99–107, Sep. 2021.

[19] S. Cakaj, "The parameters comparison of the "Starlink" LEO satellites constellation for different orbital shells," *Frontiers Commun. Netw.*, vol. 2, pp. 1–7, May 2021.

[20] 3GPP TR 38.821, v16.1.0, "Solutions for NR to support non-terrestrial networks (Rel. 16)," May 2021.

[21] K. Lee, M. Lam, R. Pedarsani, D. Papailiopoulos, and K. Ramchandran, "Speeding up distributed machine learning using codes," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1514–1529, Mar. 2018.

[22] J. Ng *et al.*, "A comprehensive survey on coded distributed computing: Fundamentals, challenges, and networking applications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1800–1837, 2021.

[23] Y. He and X. He, "Location privacy-aware coded offloading for distributed edge computing," *in Proc. IEEE INFOCOM*, 2022.

[24] Z. Wang *et al.*, "AI-based cloud-edge-device collaboration in 6G space-air-ground integrated power IoT," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 16–23, Feb. 2022.

[25] X. Fan, Y. Wang, Y. Huo, and Z. Tian, "Joint optimization of communications and federated learning over the air," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4434–4449, Jun. 2022.

Howon Lee received the B.S., M.S., and Ph.D. degrees in Electrical and Computer Engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2003, 2005, and 2009, respectively. From 2009 to 2012, he was a Senior Research Staff/team Leader of the knowledge convergence team at the KAIST Institute for Information Technology Convergence (KI-ITC). Since 2012, he has been with the School of Electronic and Electrical Engineering and the Institute for IT Convergence (IITC) at Hankyong National University (HKNU), Anseong, South Korea. He has also experienced as a Visiting Scholar with the University of California, San Diego (UCSD), La Jolla, CA, USA, in 2018. His current research interests include B5G/6G wireless communications, ultra-dense distributed networks, in-network computations for 3D images, cross-layer radio resource management, reinforcement learning for UAV networks, unsupervised learning for wireless communication networks, and Internet of Things.

He is the recipient of the Joint Conference on Communications and Information (JCCI) 2006 Best Paper Award and of the Bronze Prize at Intel Student Paper Contest in 2006. Dr. Lee is also the recipient of the Telecommunications Technology Association (TTA) Paper Contest Encouragement Award in 2011, of the Best Paper Award at the Korean Institute of Communications and Information Sciences (KICS) Summer Conference in 2015, of the Best Paper Award at the KICS Fall Conference in 2015, of the Honorable Achievement Award from 5G Forum Korea in 2016, of the Best Paper Award at the KICS Summer Conference in 2017, of the Best Paper Award at the KICS Winter Conference in 2018, of the Best Paper Award at the KICS Summer Conference in 2018, of the Best Paper Award at the KICS Winter Conference in 2020, and of the Haedong Grand Prize at the KICS Summer Conference in 2022. He received the Minister's Commendation by the Minister of Science and ICT in 2017.

Byungju Lee received the B.S. and Ph.D. degrees from the School of Information and Communication, Korea University, Seoul, South Korea, in 2008 and 2014, respectively. From 2014 to 2015, he was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, Seoul National University, Seoul, South Korea. From 2015 to 2017, he was a Post-Doctoral Scholar with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA. From 2017 to 2020, he was a Senior Engineer with Samsung Research, Seoul, South Korea. He is currently an Assistant Professor with Department of Information and Telecommunication Engineering, Incheon National University, Incheon, South Korea. Prior to joining Incheon National University, he was a Faculty Member with Kumoh National Institute of Technology, Gumi, South Korea, from 2020 to 2022. His research interests include the physical layer system design of the future wireless communications, such as integrated terrestrial and non-terrestrial network and machine learning for wireless networks. He was awarded the 2020 Fred W. Ellersick Prize from the IEEE Communications Society co-recipient of the Bronze Prize in Samsung Best Paper Award Contest in 2018 and announced as a Qualcomm fellowship awardee in 2010.

Heecheol Yang received the B.S. degree and the Ph.D. degree in Electrical and Computer Engineering from Seoul National University, South Korea, in 2013 and 2018, respectively. He was a Staff Engineer of Standard Research team in Samsung Research at Samsung Electronics, Co., Ltd. from 2018 to 2019, and was an Assistant Professor with the School of Electronic Engineering, Kumoh National Institute of Technology, South Korea, from 2019 to 2021. He is currently an Assistant Professor with the Division of Computer Convergence, Chungnam National University, South Korea. His research interests include wireless communication systems for beyond 5G and 6G, distributed storage systems and distributed computing, and data security and privacy in distributed computing and learning systems. He was awarded the Best Ph.D. Dissertation Award from Seoul National University in 2018, the Bronze Prize from the 23th Samsung HumanTech Paper Contest, and the Gold Prize from the IEEE Student Paper Contest (Seoul Section) 2017.

Junghyun Kim received the B.S., M.S., and Ph.D. degrees in the Department of Electrical and Electronic Engineering from Yonsei University, South Korea, in 2006, 2008, and 2017, respectively. From 2010 to 2013, he was an Engineer with Electronics and Telecommunications Research Institute (ETRI), Daejeon, South Korea. From 2017 to 2019, he was a Senior Engineer with Samsung Research, Seoul, South Korea. He is currently an Assistant Professor with the Department of Artificial Intelligence, Sejong University, Seoul, South Korea. Prior to joining Sejong University, he was a Faculty Member with Soonchunhyang University, Asan, South Korea, from 2019 to 2022. His research interests include intelligent wireless communication systems, post-quantum cryptography, drug discovery and development, and graph theory. He was awarded the Bronze Prize in Samsung Best Paper Award Contest in 2018. He was also the recipient of the ETRI Journal Exemplary Reviewers from ETRI in 2013, of the Best Paper Award from the Journal of Korean Institute of Communications and Information Sciences (KICS) in 2015, of the Best Paper Award at the International Symposium on Information Theory and Its Applications (ISITA) in 2016, of the Best Paper Award at the KICS Fall Conference in 2022.

Seungnyun Kim received the B.S. in the Department of Electrical and Computer Engineering from Seoul National University, Seoul, Korea, in 2016. He is currently working toward the Ph.D. degree in Electrical and Computer Engineering with Seoul National University, Korea. His research interests include signal processing, optimization, and deep learning techniques for the 5G and 6G wireless communications.

Wonjae Shin received the B.S. and M.S. degrees from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2005 and 2007, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, Seoul National University (SNU), Seoul, South Korea, in 2017. From 2007 to 2014, he was a Member of the Technical Staff with the Samsung Advanced Institute of Technology and Samsung Electronics Company Ltd., Suwon, South Korea, where he contributed to next generation wireless communication networks, especially for 3GPP LTE/LTE-advanced standardizations. From 2016 to 2018, he was a Visiting Scholar and Postdoctoral Research Fellow with Princeton University, Princeton, NJ, USA. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, Ajou University, Suwon. Prior to joining Ajou University, he was a Faculty Member with Pusan National University, Busan, South Korea, from 2018 to 2021. His research interests include the design and analysis of future wireless communications, such as interference-limited networks and machine learning for wireless networks.

Dr. Shin was the recipient of the Fred W. Ellersick Prize and the Asia-Pacific Outstanding Young Researcher Award from the IEEE Communications Society in 2020, Best Ph.D. Dissertation Award from SNU in 2017, Gold Prize from the IEEE Student Paper Contest (Seoul Section) in 2014, and the Award of the Ministry of Science and ICT of Korea in IDIS-Electronic News ICT Paper Contest in 2017. He was a co-recipient of the SAIT Patent Award (2010), Samsung Journal of Innovative Technology (2010), Samsung Human Tech Paper Contest (2010), and Samsung CEO Award (2013). He was recognized as an Exemplary Reviewer by IEEE WIRELESS COMMUNICATIONS LETTERS in 2014 and IEEE TRANSACTIONS ON COMMUNICATIONS in 2019. He was also awarded several fellowships, including the Samsung Fellowship Program in 2014 and the SNU Long-Term Overseas Study Scholarship in 2016. He is an Editor for IEEE OPEN JOURNAL OF COMMUNICATIONS SOCIETY.

Byonghyo Shim received the B.S. and M.S. degree in Control and Instrumentation Engineering (currently Electrical Eng.) from Seoul National University (SNU), Seoul, Korea, in 1995 and 1997, respectively, and the M.S. degree in Mathematics and the Ph.D. degree in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign (UIUC), Urbana, in 2004 and 2005, respectively. From 1997 and 2000, he was with the Department of Electronics Engineering at the Korean Air Force Academy as an Officer (First Lieutenant) and an Academic Full-time instructor. He also had a short time research position in the LG Electronics, Texas Instruments, and Samsung Electronics in 1997 and 2004, 2019, respectively. From 2005 to 2007, he was with the Qualcomm Inc., San Diego, CA as a Staff Engineer working on CDMA systems. From 2007 to 2014, he was with the School of Information and Communication, Korea University, Seoul, Korea, as an Associate Professor. Since September 2014, he has been with the Dept. of Electrical and Computer Engineering, Seoul National University, where he is currently a Professor. His research interests include signal processing for wireless communications, statistical signal processing, machine learning, compressed sensing, and information theory. Dr. Shim was the recipient of the M. E. Van Valkenburg Research Award from the ECE Department of the University of Illinois (2005), the Hadong Young Engineer Award from IEIE (2010), the Irwin Jacobs Award from Qualcomm and KICS (2016), the Shinyang Research Award from the Engineering College of SNU (2017), the Okawa Foundation Research Award (2020), and the IEEE COMSOC AP Outstanding Paper Award (2021). He was a Technical Committee Member of Signal Processing for Communications and Networking (SPCOM), and currently serving as an Associate Editor of IEEE Transactions on Signal Processing (TSP), IEEE Transactions on Communications (TCOM), IEEE Transactions on Vehicular Technology (TVT), IEEE Wireless Communications Letters (WCL), Journal of Communications and Networks (JCN), and a Guest Editor of IEEE Journal of Selected Areas in Communications (location awareness for radios and networks).
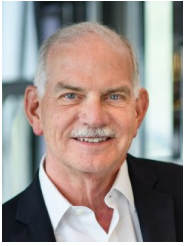
H. Vincent Poor received the Ph.D. degree in EECS from Princeton University in 1977. From 1977 until 1990, he was on the Faculty of the University of Illinois at Urbana-Champaign. Since 1990 he has been on the Faculty at Princeton, where he is currently the Michael Henry Strater University Professor. During 2006 to 2016, he served as the Dean of Princeton's School of Engineering and Applied Science. He has also held Visiting Appointments at several other universities, including most recently at Berkeley and Cambridge. His research interests are in the areas of information theory, machine learning and network science, and their applications in wireless networks, energy systems and related fields. Among his publications in these areas is the recent book Machine Learning and Wireless Communications (Cambridge University Press, 2022). Dr. Poor is a Member of the U.S. National Academy of Engineering and the U.S. National Academy of Sciences, an Honorary Member of the National Academy of Sciences, Republic of Korea, and a Foreign Member of the National Academy of Engineering of Korea. He received the IEEE Alexander Graham Bell Medal in 2017.