CONTRIBUTED PAPER

# 5G Key Technologies for Smart Railways

*This article presents a complete framework of 5G technologies for smart railways and points to exciting future research directions.*

By Bo Ai, *Senior Member IEEE*, Andreas F. Molisch, *Fellow IEEE*, Markus Rupp, *Fellow IEEE*, and Zhang-Dui Zhong, *Senior Member IEEE*

**ABSTRACT** | Railway communications has attracted significant attention from both academia and industries due to the booming development of railways, especially high-speed railways (HSRs). To be in line with the vision of future smart rail communications, the rail transport industry needs to develop innovative communication network architectures and key technologies that ensure high-quality transmissions for both passengers and railway operations and control systems. Under high mobility and with safety, eco-friendliness, comfort, transparency, predictability, and reliability. Fifth-generation (5G) technologies could be a promising solution to dealing with the design challenges on high reliability and high throughput for HSR communications. Based on our in-depth analysis of smart rail traffic services and communication scenarios, we propose a network slicing architecture for a 5G-based HSR system. With a ray tracing-based analysis of radio wave propagation characteristics and channel models for millimeter wave (mmWave) bands in railway scenarios, we draw important conclusions with regard to appropriate operating frequency bands for HSRs. Specifically, we have identified significant 5G-based key technologies for HSRs, such as spatial modulation, fast channel estimation, cell-free massive multiple-input–multiple-output (MIMO), mmWave, efficient beamforming, wireless backhaul, ultrareliable low latency communications, and enhanced handover strategies. Based on these technologies, we have developed a complete framework of 5G technologies for smart railways and pointed out exciting future research directions.

**KEYWORDS** | Artificial intelligence (AI); channel model; fifth-generation (5G)-R; Internet of Things (IoT); massive multiple-input–multiple-output (MIMO); millimeter wave (mmWave); smart railways; ultrareliable low latency communication (URLLC).

## NOMENCLATURE

| | |
|---|---|
| 5G | Fifth generation. |
| AI | Artificial intelligence. |
| AoA | Angle of arrival. |
| AoD | Angle of departure. |
| AP | Access point. |
| ARQ | Automatic repeat request. |
| ASA | Angular spread of arrival. |
| ASD | Angular spread of departure. |
| ASM | Adaptive spatial modulation. |
| AT | Attention. |
| ATO | Automatic train operation. |
| AWGN | Additive white Gaussian noise. |
| BBU | Baseband unit. |
| BEM | Basis expansion model. |
| BER | Bit error rate. |
| BF | Beamforming. |
| BPSK | Binary phase shift keying. |
| BS | Base station. |
| CCT | Cut and cover tunnels. |
| CIR | CAD-integrated radio. |
| CN | Core network. |
| C-RAN | Cloud radio access network. |
| CSI | Channel state information. |
| CTC | Centralized traffic control. |
| CUPS | $C/U$-plane splitting. |

Manuscript received November 28, 2019; revised February 13, 2020; accepted April 9, 2020. Date of current version May 28, 2020. This work was supported in part by the National Key Research and Development Program under Grant 2016YFE0200900; in part by NSFC under Grant 61725101, Grant 6196113039, and Grant U1834210; in part by the Royal Society Newton Advanced Fellowship under Grant NA191006; and in part by the State Key Laboratory of Rail Traffic Control and Safety under Grant RCS2020ZT010. The work of Andreas F. Molisch was supported by the U.S. National Science Foundation. *(Corresponding author: Bo Ai.)*

**Bo Ai** and **Zhang-Dui Zhong** are with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China (e-mail: boai@bjtu.edu.cn; zhdzhong@bjtu.edu.cn).

**Andreas F. Molisch** is with the Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089 USA (e-mail: molisch@usc.edu).

**Markus Rupp** is with the Institute of Telecommunications, Technische Universität Wien, 1040 Vienna, Austria (e-mail: mrupp@nt.tuwien.ac.at).

Digital Object Identifier 10.1109/JPROC.2020.2988595

| | |
|---|---|
| D2D | Device-to-device. |
| D-RAN | Distributed radio access network. |
| EE | Energy efficiency. |
| E2E | End-to-end. |
| eMBB | Enhanced mobile broadband. |
| eMLPP | Enhanced multilevel precedence and preemption. |
| ERAB | Evolved radio access bearer. |
| ESA | Elevation angular spread of arrival. |
| ESD | Elevation angular spread of departure. |
| ETCS | European train control system. |
| F-RAN | Fog-computing-based radio access network. |
| FBC | Finite block length codes. |
| FDD | Frequency division duplex. |
| GMM | Gauss–Markov model. |
| GPS | Global positioning system. |
| GRIS | Group packet radio service interface server. |
| GSM | Generalized spatial modulation. |
| GSM-R | Global system for mobile communications for railways. |
| HARQ | Hybrid automatic repeat request. |
| H-CRAN | Heterogeneous cloud radio access network. |
| HD | High definition. |
| HD-SDI | High definition serial digital interface. |
| HiBEM | Historical information-based basis expansion model. |
| HPC | High-performance computing. |
| HPN | High power nodes. |
| HSR | High-speed railways. |
| HST | High-speed train. |
| ICI | Interchannel interference. |
| IoT | Internet of things. |
| IoT-R | IoT for railways. |
| IRTS | Intelligent rail transport system. |
| ITS | Intelligent transport system. |
| ITU | International telecommunication union. |
| KPI | Key performance indicator. |
| LMMSE | Linear minimum mean square error. |
| LoS | Line of sight. |
| LS | Least square. |
| LTE | Long-term evolution. |
| LTE-R | Long-term evolution for railways. |
| MANO | Management and orchestration. |
| MBB | Mobile broadband. |
| MBB-R | Mobile broadband access for railways. |
| MBS | Macro base station. |
| MCBF | Multicell-coordinated BF. |
| MIMO | Multiple-input–multiple-output. |
| MO | Modulation order. |
| MR | Mobile relay. |
| MRN | MR node. |
| MTTM | Mean time to maintain. |
| MTTR | Mean time to restore. |
| MUT | Mean up time. |
| MWB | Mobile wireless backhaul. |
| NFV | Network function virtualization. |
| NGMN | Next generation mobile networks. |
| NLoS | Nonline-of-sight. |
| NOMA | Nonorthogonal multiple access. |
| NSA | Nonstandalone. |
| OFDM | Orthogonal frequency division multiplexing. |
| PLE | Path loss exponent. |
| QAM | Quadrature amplitude modulation. |
| QoE | Quality of experience. |
| QoS | Quality of service. |
| RAMS | Reliability, availability, maintenance, safety. |
| RAN | Radio access network. |
| RATG | Radio access technique group. |
| RF | Radio frequency. |
| rms | Root mean square. |
| RRH | Remote radio head. |
| RRU | Remote radio unit. |
| RT | Ray tracing. |
| Rx | Receiver. |
| SA | Standalone. |
| SAGE | Space-alternating generalized expectation. |
| SDN | Software-defined network. |
| SE | Spectrum efficiency. |
| SIP | Session initiation protocol. |
| SISO | Single-input–single-output. |
| SM | Spatial modulation. |
| SM-SVD | SM signal-vector-based detection. |
| SNR | Signal-to-noise ratio. |
| ST-ASM | Space-time-coded adaptive spatial modulation. |
| STD | Standard deviation. |
| TDCS | Train dispatch control system. |
| TDD | Time division duplex. |
| TDMA | Time-division multiple access. |
| THR | Tolerate hazard rate. |
| TTR | Time to restore. |
| TTRS | Time to return to safety. |
| UCN | User-centric network. |
| UDN | Ultradense network. |
| UEs | User equipments. |
| UPA | Uniform planar array. |
| URLLC | Ultrareliable low-latency communication. |
| V2V | Vehicle-to-vehicle. |
| V2X | Vehicle-to-everything. |
| VBS | Voice broadcast service. |
| VGCS | Voice group call service. |
| VIM | Virtualized infrastructure manager. |
| VNF | Virtual network function. |
| WB | Wireless backhaul. |
| WRC | World radio conference. |
| xMBB | Extreme mobile broadband. |
| ZF-BF | Zero-forcing beamforming. |

## I. INTRODUCTION

Rail traffic systems are widely acknowledged as an alternative green transportation for goods and people due to their higher mobility, higher EE, and/or substantially lower environmental impact than conventional transportation

systems. In the near future, people will see further advances in railway systems, such as the development of fully automatic train operation, real-time monitoring, prewarning for rolling rocks and infrastructure, highly accurate passenger information, and predictive maintenance planning. All these are the services in smart rail systems, whose target is not only to improve operation mobility, safety, and reliability, but also to improve QoS, eco-friendliness, comfort, and cost efficiency.

Industry players around the world have been interested in the future smart rail implementations. For one, the developers of the "Shift2Rail Plan" of the European Union have been working on the evolution of railways [1]. Infrastructure, trains, travelers, and goods are expected to be more interconnected in the future to provide consumers with more comfort and security. For the first time, the WRC has just published the railway train-to-ground communications in WRC-19. IEEE has also issued a series of standards to support the development of smart rail communications. For the Beijing 2022 Winter Olympics, the Chinese government has built a smart HSR line with a speed at 350 km/h to allow Olympic athletes, reporters, and audiences to move efficiently between Beijing and Zhangjiakou. Generally, smart railways are under construction all over the world.

As pointed out in [1], future smart rail implementations need to have a new generation of signaling and control systems to enable intelligent traffic management. Some of recent railway accidents in the United States, Europe, and Asia could have been prevented if technologies such as positive train control were implemented. Accordingly, this has reminded the public of deploying advanced wireless communications-based railway technologies. As pointed out in [2], the railway industry is now looking into wireless communications with great interest. In recent years, a number of IEEE journals have published special issues that focus on railway communications technology. Above all, both academia and industries have developed many initiatives that focus on communications technology to fulfill the goals for the future smart railways.

A critical component in smart rails is train operational control systems, which require a reliable communication link between a train and the ground. The most widely used railway communications system around the world is the GSM-R, which has been used for decades and is still in operation in many railway systems [3]. GSM-R is essentially the same as GSM, but with railway-specific functionalities. The uplink/downlink (UL/DL) peak data rate of GSM-R is 172 Kb/s. Typical railway-specific services offered by GSM-R include VGCS, VBS, eMLPP, shunting mode, functional addressing, and location-dependent addressing. In general, most of the above services do not require Mb/s-level data rate.

Currently, communication systems following the LTE-R standard are under consideration [3]. However, the maximum bandwidth of LTE-R systems is only 20 MHz, which can only support Mb/s-level data rate. For the future smart railways, services with higher data rate are required, such as real-time 4K/8K ultrahigh definition (UHD) video transmissions, security closed circuit television (CCTV) in the train cabin, remote maintenance of trains, and high data rate wireless offloading at railway stations. The above services of smart railways generally require Gb/s-level data rate and LTE-R cannot meet these requirements.

Moreover, smart rail traffic needs to have more thorough coverage of sensors; the train, goods, and passengers need to be more interconnected; and operators need more intelligent methods to deal with sensing data. All of these require a dedicated wireless communications system that can support a massive number of sensors, ms-level delay, ultrareliability, and/or Gb/s transmission rate at over 500-km/h train speed. Similar to the issues specified in [1], smart rail infrastructure, smart mobility management, smart rail services, and a new generation of trains ultimately form requirements for seamless high data rate communications for future rail developments.

However, not many innovative methodologies were available to enable designers to achieve these goals, and thus further push the smart rail development back until the 5G key technologies appeared. In particular, 5G key technologies offer much faster data rate and larger coverage area. Ericsson Mobility Report forecasts in its June 2019 edition that 35% of global mobile connections will be carried out on 5G networks and 5G users will reach 1.9 billion, which is up to 65% of the global population [4]. In 5G, a plethora of novel applications will be introduced by providing a diverse set of services for both machine-type and human-end users. Three important scenarios for 5G include eMBB, URLLCs, and massive machine type communications (mMTCs) [4]–[6]. In December 2017, the first global 5G standard was approved by 3GPP as an interim set of Release 15, which is the initial phase NSA deployments of eMBB. At the time of this writing, with Release 16, 3GPP is moving ahead to SA deployments focusing on URLLC and mMTC. Further 5G enhancements will be completed in Release 17, which starts in early 2020.

By now, 5G service has already become available in some countries. For example, Verizon has launched 5G fixed wireless broadband Internet and mobile services in some cities in the United States. In South Korea, the first 5G services were provided to general customers in April 2019, reaching a total of 2 million 5G subscribers by the end of August 2019. In June 2019, China handed out commercial licenses to operate 5G mobile networks. By the end of 2020, more than 100 million 5G subscribers are expected in China. However, a widespread coverage of 5G networks around the world is not available now.

There are many disruptive enabling technologies integrated into the realization of 5G services [7].

### A. Millimeter Wave (mmWave)

The need for higher data rate in 5G motivates the exploration of the mmWave bands from 30 to 300 GHz [8].

The most promising bands for 5G could be the $Ka$-band at 26.5–40 and 42–48 GHz, the license-free band at 60 GHz, and the $E$-band at 70–76, 81–86, and 92–95 GHz [9]. However, the cell size needs to be greatly reduced and new network architecture should be designed [10]. Another important factor that affects the deployment of mmWave transmissions is beam management, which is composed of beam sweeping, beam measurement, beam determination, and beam reporting [11]. It is well known that mmWave signals suffer from high isotropic free-space path loss and are prone to blockage [12]. In this regard, massive MIMO transmissions with beam management could be a good way to handle the former problem, while other approaches such as beam management and distributed antennas could help with the latter problem.

### B. Massive MIMO

Another important research direction to improve the data rate is to utilize massive MIMO, which played a major role in early 5G trials and becomes more central as 5G matures [13]. Although massive MIMO has been used for 5G networks operating in conventional sub-6-GHz bands, it still presents many research challenges [14]. For example, how will massive MIMO and mmWave work in cooperation to ensure a successful 5G network?

### C. Network Slicing

Different KPIs of 5G networks can support a number of vertical industries with different requirements, such as critical communications, massive IoT, and V2X communications [15], [16]. Instead of "one size fits all" pursuing design, network slicing attempts to support such optional and functional diversity. However, the next phase 5G network should address several open issues, including handover for multislice connected UEs.

### D. Ultrareliable Low Latency Communication

To meet the ultrareliable and low latency requirements in industrial IoT and V2X scenarios, 3GPP has incorporated URLLC into the 5G standards. In a typical use case of factory automation, the cycle time should be between 0.25 and 50 ms, and the packet error rate is lower than $10^{-9}$ [17]. The technical challenges for URLLC include synchronized operations in the licensed and unlicensed band, over-the-air intersystem coordination, and traffic prioritization.

### E. Nonorthogonal Multiple Access

Current orthogonal multiple access technologies pose important challenges to 5G networks with a large number of UEs. The underlying idea for NOMA is to serve more than one UE by using the same time, code, space, and frequency resource [18]. UEs strong processing capability for interference cancellation is the key factor for the implementation of NOMA in 5G networks [19], [20]. Furthermore, NOMA may provide the solution to grant free access for a large number of IoT devices. More comprehensive research is required to better understand the effect of practical conditions on NOMA performance.

European 5G organizations, such as NGMNs, and the 5G promotion group IMT-2020 in China have identified HST as one of the 5G typical scenarios that represent $\geq$500-km/h mobility [21], [22]. However, railway communications are significantly different from other public mobile communications in terms of mobility, scenarios, and KPIs such as dependability and data integrity [23]. Therefore, developers need to find viable innovative network architectures and key technologies for the future development of smart rail. Since 5G researchers have given little attention to railway communications in the past, only a few articles deal with 5G technologies in the rail systems. To fill this gap, we present a comprehensive analysis on 5G key technologies for smart railways. This includes both a review of the state-of-the-art as well as the original research.

Based on the aforementioned discussions, we present our major contributions in this article as follows.

1) We summarize new applications and KPIs for 5G-R based services that we develop in Section II.
2) We introduce state-of-the-art, drawbacks, and design challenges of typical 5G application scenarios, including geographical and communication scenarios, for smart rail systems in Section III.
3) With a comprehensive consideration of propagation and channel characteristics under the appropriate deployment frequency bands, we propose a 5G network slicing architecture suitable for smart rail systems in Section IV.
4) We analyze the propagation characteristics and channel models for massive MIMO in mmWave bands for smart rail systems. Based on this investigation, we offer our recommended frequency bands appropriate for smart rail communication networks in Section V.
5) With a comprehensive consideration of smart rail 5G network architecture, propagation, and channel characteristics under appropriate deployment frequency bands, we propose several 5G key technologies suitable for smart railways in Section VI.
6) Future design issues and research directions are given in Section VII.

## II. SERVICES AND KEY PERFORMANCE INDICATORS

In this section, we detail the list of services for smart railways. The fundamental purpose of smart railways is to have a traffic system that is highly efficient, comfortable, convenient, transparent, predictable, and environmentally friendly.

### A. Services Provided for Smart Rail

We developed a basic architecture for smart rail. As shown in Fig. 1, the intelligent rail infrastructure and

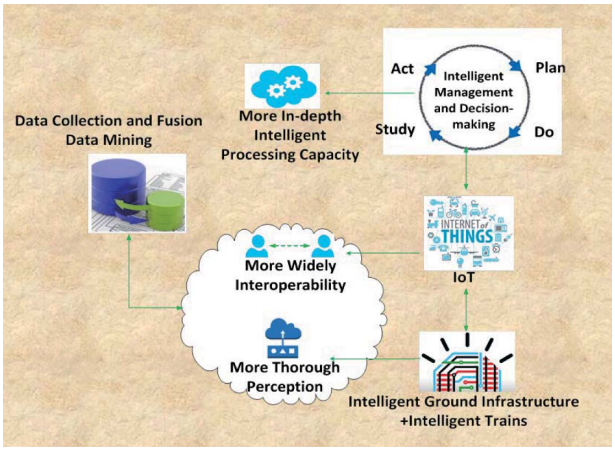


**Fig. 1.** ***Brief architecture for smart rail.***

intelligent train act as the physical units to provide a more thorough (i.e., a faster, more sensitive, reliable, efficient, and comprehensive) perception. The IoT acts as a platform for deep and intelligent perception. With more thorough perception and wider interoperability, we need a more in-depth intelligent management and decision-making system. As such, big data and AI technologies are needed to be able to collect, fuse, and mine data and to promote an intelligent, reliable, and highly efficient management and decision-making system.

The UIC divides railway requirements mainly into two categories: railway operation communications and railway support communications. Meanwhile, there should be IT solutions that would make railway services more attractive to passengers in smart rail. For example, such IT solutions could be geared toward helping the passengers to share high data rate services from the Internet while onboard the train. In addition to onboard real-time high data rate services for passengers, it also helps to provide IoT for railways. According to the business attributes, the services and applications for smart railways can be categorized into four categories: railway safety-critical service, railway non-safety service, passenger-oriented, and IoT for railways.

### B. Key Performance Indicators for 5G-R

To establish KPIs for smart rail communication systems, we first need to distinguish between the two main categories of railway communication services. The first category is developed for the passengers and is almost similar to the KPIs of the public land mobile communication network. The 5G KPIs are mainly about user-experienced data rate, contention density, E2E latency, traffic volume density, mobility, peak data rate, and so on. This KPI category mainly focuses on the stringent requirements in terms of throughput and QoS. On the other hand, the second category applies to train operation control and safety. This category is closely related to the characteristics of RAMS, resilience, integrity, mission criticality, safety criticality, and dependability.

Researchers and developers need to evaluate the 5G-R system to better suit the requirements of use case categories because use cases have very diverse and, sometimes, conflicting requirements. For example, sensor applications generally require low data rates and can sometimes tolerate high latencies. Meanwhile, evolving applications, such as remote control, require high data rates and low latency.

We estimate the spectrum needs of the train radio applications for future smart rail communications based on the spectrum calculation methodology introduced by Recommendation ITU-R M.1768-1 [24]. We show that the modified flow for the spectrum needs calculation of future smart rail communication as follows [25].

*Step 1:* We present the different definitions used in the methodology, including the geographical area, service category, and service environment.
*Step 2:* We list the SE consideration for the different SEs.
*Step 3:* We then assign the area SE matrix (SE consideration) for one RATG.
*Step 4:* We obtain the values for the calculation (session arrival rate per user, mean service bit rate, and mean session duration) from the calling record files of railway operators.
*Step 5:* We then calculate the aggregated spectrum need.

When considering latency requirements, the requirements on minimum latency are expressed in terms of E2E latency, which is the latency perceived by the end user. Note that this assumes that the delay arising from the application layer processing due to transport and switching is negligible.

The reliability is also a crucial KPI in smart rail communication systems, which can be characterized by the reliability rate defined as follows: the amount of sent packets successfully delivered to the destination within the time constraint required by the targeted service, divided by the total number of sent packets.

## III. OVERVIEW OF RELATED WORKS

Most railway communications today use GSM-R, a narrowband and low-rate communication system that cannot fulfill the requirements of smart rail services such as railway multimedia dispatching communications in the future. Although there have been some studies on smart rail communications since IBM proposed the idea of smart rail in 2009 [26], there are still significant gaps. In this section, we briefly introduce major related works on this topic.

Ning *et al.* [27] presented the concept of IRTSs, which includes an intelligent client service system, intelligent train control, and dispatching system, as well as an intelligent emergency rescue and supervision system. All of these bring together a cornucopia of technologies (e.g., sensing, communication, computing, and intelligent control) to address various aspects of IRTS. One of the benefits of IRTS is that it is capable of increasing train operation speed and track capacity while maintaining high safety

level. The crucial technological ingredients in IRTS include low-cost sensors, clever software for analytics and visualization, and high computing power. Niculescu *et al.* [28] presented a smart infrastructure, which refers to all communication and information equipment that supports the development and the implementation of IRTS services. However, it mainly focused on smart rail infrastructure, with little consideration of smart rail communications. Above all, the discussions on smart rail systems in the existing literature have presented these systems as just some concepts the studies do not provide specific architectures or advanced technologies that could help in developing these systems.

### A. Network Architecture

We first review the general network architectures that have been proposed for 5G. Note that C-RAN has been designed and tested since 3G [29], and D-RAN have been applied to 4G as well [30]. Moreover, CUPS was already implemented in the 4G CN. However, due to the high demand for various services and due to the variety of communication scenarios, the performance requirements for wireless networks have become more stringent in recent years. To handle such requirements, 5G researchers have developed some novel RAN architectures, such as fog-computing-based radio access network (F-RAN), H-CRAN, and UDN. Based on SDN and NFV techniques, both RAN and CN can be made programmable, flexible, and elastic [15].

The HPNs (e.g., macro or micro BSs) existing in HetNets take charge of the seamless coverage and all of the control signaling interaction. Both cloud computing and deployment heterogeneity may ensure the backward compatibility to the current wireless system and for it to evolve smoothly to the next-generation networks. HPNs can be introduced into C-RAN systems to create H-CRAN, taking full advantages of both HetNets and C-RAN [31]. In contrast to the general C-RAN architecture, the H-CRAN system also supports the separation of the control plane ($C$-plane) and the user plane ($U$-plane). The paradigm HPNs deliver the overall control signaling throughout the system, including the signaling interaction with the BBU pool, such that the centralized control-cloud functions detached from the BBU pool can efficiently improve the round-trip latency. However, the constraints caused by the fronthaul links, as compared to those by ideal links, limit the overall SE and EE of the H-CRAN architecture.

To mitigate the shortcomings of C-RAN and H-CRAN, Peng *et al.* [32] proposed the F-RAN architecture. Incorporating fog computing into the edge cloud, which is a traditional RRU equipped with cooperative network functions (including web caching, signal processing, and radio resource management), can evolve into an innovative network element termed as fog-AP. Adjacent "smart" UEs (denoted as fog-UEs) can communicate with each other via D2D or relay modes to improve SE. The cross-tier interferences between the fog-AP and HPN can be suppressed in the BBU pool by using coordinated multipoint techniques. Four candidate transmission modes can be selected such that UEs can access F-RAN adaptively in accordance with the UE mobility characteristics: D2D and relay mode, local distributed coordination mode, global C-RAN mode, and HPN mode.

**Table 1** Merits and Challenges of 5G Network Architectures

| Network architecture | Merits | Challenges |
|---|---|---|
| H-CRAN | Backward compatibility, less handoff, front-haul efficiency | Serious interference in cross-tier and inter-tier |
| F-RAN | Front-haul efficiency, high capacity in hot spots | Complex and mass node reconstruction |
| UDN | Improve spectral and area throughput | Complex interference management |
| UCN | Reduce inter-user interference | Global CSI and huge fronthaul signaling |

In contrast to previous cellular generations, 5G network infrastructure densification introduces the UDN paradigm to obtain a large system performance gain [33]. This is due to each user having one or more BSs exclusive service. Thus, proximal communications and improved spatial degree-of-freedom can be leveraged. However, UDN deployment also introduces highly challenging interference management. In order to tackle this challenge, efficient and realistic network deployment strategies are required by taking backhaul overhead, cost constraint, and computational complexity into account.

Another promising 5G network architecture is a UCN, which facilitates the centralized signal processing, low hardware cost, interference-free connection between users, and many more [34]. It takes advantages of the rapid development of cloud and edge computing techniques. However, UCN requires global CSI knowledge for efficient cooperation between APs. Furthermore, a large amount of fronthaul capacity is required for the UCN. In the future research, it is promising to investigate the dynamic AP selection, fronthaul compression, pilot assignment, power control, and mobility management.

Table 1 compares the merits and challenges of 5G network architectures.

With a huge available bandwidth in the mmWave band, such as the 60-GHz band and $E$-band (71–76 and 81–86 GHz), mmWave wireless communications can provide multigigabit transmission rates and support a lot

of high-speed data services [35]. Since small cells and macrocells are in different bands, there is no interference between small cells and macrocells. Combining the coverage and reliability from macrocells and high bandwidth from small cells, we expect small cells in high frequency bands underlying the macrocells to be able to deliver a good performance [35]. To ensure reliability and coverage, an intuitive method is to offload the control signaling to the overlaid MBS, providing continuous coverage and reliable connectivity at lower frequency bands, and the small cell provides high-speed data transmission at higher frequency bands within small coverage [35].

Compared with other electromagnetic waves at lower frequencies, mmWave communication has two main characteristics: high propagation loss and vulnerability to obstacles. MmWave directional BF is adopted to combat high propagation loss [36]. There have been several works on efficient BF for mmWave communications. By exploiting useful information such as location and mobility pattern, the overhead for BF is hugely reduced. There are several works on exploiting mmWave communications for HSR communication [37]. For example, prior works on mmWave channel modeling and mmWave BF design have validated the efficacy of HSR communications [38]. Furthermore, to combat the random blockage in practical scenarios, a variety of solutions have been proposed, such as relay-assistance [39], reflection exploitation [40], and large intelligent surfaces [41].

We now turn to the networks for railways. Considering the train safety-related services, mission-critical data information should be delivered with high reliability at lower frequency bands by the MBS, and thus only the passenger-oriented data information is offloaded to small cell BSs at high frequency bands [42]. The CUPS architecture is based on a HetNet deployment, in which small cells are overlaid on the coverage of macrocells. In general, macrocells working at low-frequency bands provide universal coverage and mobility support whereas small cells use high-frequency bands to provide high data rate transmission. Because frequency bands are scarce, passenger-oriented services requiring huge capacities depend on small cells for the $U$-plane. The corresponding $C$-plane bearing the control signaling remains in the macrocell and thus avoid frequent handover and severe intercell interference. By separating the user layer and the control layer resources, the operator can provide flexible services. The user layer can be distributed to different areas in order to get closer to the UEs, thereby reducing delay and bandwidth requirements. The control layer then reduces the complexity of operation and maintenance through centralized management [43].

Besides the frequent handover issue, another problem caused by handover is the group handover issue, as a number of active information users may desire to connect with Internet, enjoying multimedia services and online games. In this case, all active users have to launch their handover requests at almost the same time and all handover requests are required to be completed within tens of milliseconds, which is challenging especially when the number of active users is large. This may put a heavy burden on both the train-ground communication and the handover computing processing. To alleviate such a problem, the two-hop architecture proposed in [44] may be an efficient solution. With the two-hop architecture, all information users in the train access the Internet via the train relay station (TRS) rather than directly linking the ground BS. In this way, the handover requests are aggregated and merged into a single one, so the UEs can be treated as a virtual big UE, and the handover processing burden is greatly reduced.

To guarantee the special needs of the required QoS, several studies have proposed to apply a network-slicing technology. This technology is a type of a virtual network architecture that allows multiple virtual networks operating on top of a common shared physical infrastructure to be separated in order to meet the specific needs of applications, services, devices, customers, or operators in terms of customized requirements such as latency, bandwidth, reliability, and security [45]. It is service-driven and aims at addressing the need of different use cases with highly diverse requirements. Network slicing further allows developers to set up new services and applications or modify existing ones. It can also support the communication services of a particular connection type with a specific way of handling $C$-plane and $U$-plane for these services.

Network slicing is created by E2E logical networks. Each network is flexible enough to provide one or more network services that are in accordance with the needs of the slice requirements such as vertical industry users, virtual operators, and business users [46]. Each slice can serve a specific vertical application to provide network services efficiently, to support the concurrent work of multiple use cases, and to provide flexibility to the network [47]. In other words, developers can take advantage of relevant technologies and key functions with network slicing; they can create specialized, dedicated, and logically independent virtual networks based on generic, programmable physical infrastructures. This technology also supports operators since it provides customized virtual networks for specific needs and scenarios and meets the KPIs. However, as far as we know, there is limited work on network slicing for railway communications.

An mmWave communication network architecture for railways has been presented in [48], emphasizing its capability to trigger handover in advance and thus enhance the probability of handover success. Yan and Fang [49] developed a network architecture that integrates mmWave communication and radar detection for railways. The proposed integrated network was deployed based on C-RAN. However, due to the difficult propagation characteristics of mmWave and due to the particularity of railway scenarios, there are still many challenges ahead. Moreover, these architectures are based on LTE but not on 5G.

A specific smart collaborative network architecture for railways was presented in [50], aiming to promote efficient and reliable communications under railway scenarios.

However, this architecture is not suitable for chain-shaped communication networks for the railways. It also lacks backward compatibility with redesigned network components and smart hierarchies.

### B. Channel Models

Channel models for railway communications are different from those for traditional cellular communications, which necessitates both new measurement campaigns and modeling approaches for the former. Wang *et al.* [51] provided a comprehensive review of the channel measurement campaigns conducted in different railway scenarios; another survey is provided in [52].

It is generally useful in channel modeling to distinguish between frequencies below 6 GHz and above 20 GHz. At below 6 GHz, most of the existing investigations have focused on narrowband measurements, characterizing path loss, small-scale fading, and shadowing in a variety of environments such as open area, cuttings, viaducts, and train stations. For example, Liu *et al.* [53] characterized narrowband railway scenarios at 2.35 GHz and then established a statistical position-based channel model. Based on the measurements performed along the Zheng-Xi railway line, a stochastic channel model for cuttings, viaducts, crossing bridges, and train stations at 930 MHz was provided by He *et al.* [54].

Many researchers have also used a variety of channel modeling approaches to characterize railway channels. Generally, geometrical approaches are preferred for characterizing the nonstationarities of the channels and to implicitly provide all the parameters required for multi-antenna characteristics. Ghazal *et al.* [55] developed a generic nonstationary wideband geometry-based stochastic model for MIMO systems in railway scenarios, and good agreement was obtained between the statistical properties of the proposed generic model and measurement data. Lin *et al.* [56] developed a finite-state Markov chain channel model while considering the impact of movement speeds on the temporal statistical channel characteristic. A geometry-based random-cluster model revealed the behavior of the clusters and their temporal changes for HSR [57]. Unlike in 3GPP-type models, these models characterize the time evolution of the clusters with high accuracy. Moreover, Zhou *et al.* [58] modeled the cluster evolution as well as the impact of railway-specific structures such as power poles based on extensive measurements. In summary, the channel characteristics and channel models under HSR scenarios below 6 GHz have been subject to a number of investigations but are far from completion and thus should be further explored.

In terms of mmWave bands, to the best of our knowledge, only measurements within a train car have been performed. Although no measurements between moving trains and infrastructure have been conducted, a few researchers have studied these channels through RT. Some of the environments that have been investigated include a typical straight subway tunnel [59] as well as an arched tunnel [60]. The results have shown that systems with high gain directional antennas can support communication links longer than 1 km while inside a tunnel. The roughness of the wall inside the tunnel, antenna radiation pattern, and orientation in railway tunnels at mmWave bands have been studied in [61]. Meanwhile, Chang *et al.* [62] presented a channel model based on RT at 60 GHz for an HSR open area.

A 5G massive MIMO channel modeling method based on a theoretical nonstationary 3-D wideband twin-cluster channel model was presented in [63]. Moreover, the performance of 5G massive MIMO systems has been analyzed in [64] and [65], which reveals that the achievable SE can be assessed only through appropriate massive MIMO channel models. An important finding in [66]–[70] is that massive MIMO systems can be employed with low-cost hardware to achieve remarkable SE and EE. Guan *et al.* [71] defined and developed six scenario modules for mmWave and THz train-to-infrastructure channels for the "smart rail mobility."

Overall, many articles have dealt with mmWave channel measurements and channel models; however, few have studied channel models at mmWave bands in rail traffic scenarios.

### C. Key Technologies

In railways, the requirement for seamless Gb/s-level transmission motivates researchers and developers to explore mmWave communication and BF technologies. However, an open and important question is whether mmWave and BF can really work in HSR. In order to answer this question, the researchers in Japan and South Korea have conducted field trials of mmWave communications for HSR to obtain high data rate [72]. Moreover, standardization efforts have been conducted on the mmWave-based HSR communications in 3GPP 5G NR specification from the aspects of network architecture, channel model and estimation, frame structure, Doppler compensation, efficient handover, and so on [73]. Therefore, mmWave is a promising and practical technique for HSR communications.

To extend the coverage in HSR, it is meaningful to use high directivity and narrow BF to combat the large isotropic path loss in the mmWave band. Several studies have dealt with massive MIMO for railways. Cui and Fang [74] designed a multistream BF scheme and used an adaptive beam-selection algorithm to exploit the train location information. However, the performance severely degrades when the interbeam interference increased.

Massive MIMO has been studied extensively as one of the core technologies in 5G to improve the system capacity [75]. SM is a candidate scheme of massive MIMO, effectively reducing energy consumption and complexity, making it a potential technique in beyond 5G. In particular, Cui and Fang [76] investigated the performance

of massive SM MIMO over a spatial-temporal correlated Rician fading channel for railway scenarios. The study theoretically determined that higher velocity makes temporal correlation more dominant when the train moves through a spatially variant field. Also, a hybrid analog-digital SM-BF scheme operating at mmWave frequencies for future railways was proposed in [77]. This scheme, however, performs poorly in the low SNR regime with velocities larger than 360 km/h. As such, Chen and Fan [78] suggested a low-complexity BF method based on the train location information, with a suboptimal solution for eliminating the interbeam interference and maximizing BS service ability. An mmWave BF scheme for railway disaster detection has been presented in [79]; in the study, the concerned area around the railways was divided into different detection areas with their corresponding danger sensitivity levels. In addition, the channel coherence time is small in mmWave communications with wide beams. Therefore, massive MIMO with BF is challenging in railway scenarios.

Channel estimation is of vital importance for wireless communication systems, especially in the case of high mobility. Fading channels for HSRs are fast time-varying due to train speeds of 360 km/h and more. Therefore, channel estimation for 5G on HSRs is a big challenge. Furthermore, channel estimation is a new research topic when combining with other 5G key technologies, such as massive MIMO and mmWave.

Currently, most of the existing articles that estimate the channel parameters for massive MIMO, both in TDD and FDD modes, assume that wireless channels are static or quasi-static. Besides, existing studies for channel estimation utilize training symbols and traditional estimators such as LS and LMMSE. However, as pointed out in [80], another two types of information can aid the channel estimation for HSRs: the wireless data of the past trains on the same track and the scenarios of the trains such as tunnels, viaducts, and cuttings. Accordingly, data-motivated and environment-sensing channel estimators, different from traditionally training-aided channel estimators, can be applied in 5G systems on HSRs to enhance the estimation accuracy.

As the Doppler spread increases in HSR, the resulting large intercarrier interference may degrade the performance of OFDM; thus, countermeasures or alternative modulation methods need to be developed. The orthogonal time-frequency space modulation, which performs modulation in the delay Doppler domain, has been proposed in [81] to address this issue. This method spreads each modulation symbols over the entire time-frequency domain to increase diversity.

As for the body of research that deals with reliable train-to-ground communication technologies, the most popular technique avoiding large penetration loss into the railway cars is the MR-based communication. UEs onboard the train communicate with the BS aided by antennas mounted on top of the train as gateways [82], whereas distributed antennas constituting the other link end should be placed along the railway line. The key issue in the latter case lies on selecting an optimal cooperative remote antenna unit (RAU) of a linear distributed antenna for railway scenarios. Accordingly, Liu and Fan [44] presented a low-complexity approach based on the theory of convex optimization to select a transmit set of RAUs for capacity maximization. Likewise, Cheng *et al.* [83] proposed an adaptive antenna activation-based BF scheme to mitigate interbeam ambiguity. The proposed scheme can achieve SE gain over nonadaptive schemes and shows more robustness against direction-of-arrival estimation errors.

Moreover, Fan *et al.* [84] proposed a novel Doppler shift estimation algorithm for HST based on the reconstructed received signal via a discrete Fourier transform with massive linear receiving antennas. A filter bank multicarrier-based technique was presented in [85] to estimate the effect of time-varying channels on multicarrier signals, proving to be accurate for WiMAX and LTE at high speeds. High mobility also affects the media access control (MAC) and networking layer; thus, You *et al.* [86] presented a simple but effective distributed load-balancing algorithm to relieve service interruption due to frequent handovers in high-mobility scenarios.

As for the research on URLLC for railway scenarios, the first issue that needs to be considered is the design of the frame structure that would meet the requirements of ms-level latency. For example, the larger the block length is, the higher the coding efficiency would be [87]. However, this does not consider implementation complexity. For uncontrollable external interference signals in URLLC systems, Lee *et al.* [88] presented a "virtual pilot" frame structure that uses the data field to estimate the interference covariance matrix such that the transmission delay decreases. This frame structure, however, is suitable only for low-rate transmission scenarios; otherwise, the covariance matrix estimation error may cause an error diffusion problem, and consequently affect transmission reliability. As such, 3GPP has proposed the concept of mini-slots to meet the demands of URLLC [89] as it has been suggested that the subcarrier spacing would be larger under this condition. Variable length coding and feedback mechanisms can greatly improve the maximum achievable rate of the system [90].

Using time diversity via retransmissions can reduce the effect of deep fading. Accordingly, ARQ is the traditional retransmission technology, albeit reducing system transmission throughput. Although HARQ reduces the throughput loss, a major challenge in taking advantage of time diversity is how to characterize effectively the system interrupt probability after a certain number of transmissions [91]. If the code length is adaptively changed, then the system throughput can improve even if the HARQ feedback delay is larger. However, the delay introduced by the signaling overhead of ARQ mechanisms cannot be ignored. Therefore, some scholars have proposed

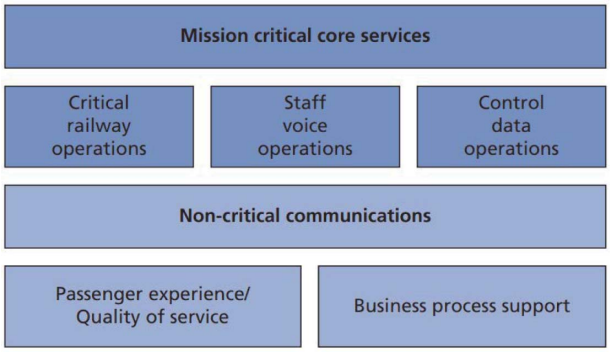


**Fig. 2.** ***Different categories of railway group services.***

a method to avoid the retransmission of information confirmation.

The existing research on reliability lacks the design and optimization of the frame structure, retransmission, and feedback mechanism of URLLC for multiple users in railway scenarios. Thus, this subject needs to be further studied and explored. Most of the current frame structure designs are based on heuristics and can be evaluated only through numerical simulation. Therefore, researchers need to establish a novel optimization framework for designing URLLC frame structures while considering both signaling information and data payload. Furthermore, only a few studies exist with regard to the ARQ/HARQ mechanisms of short packet transmission in URLLC systems. The retransmission mechanism based on information acknowledgment makes the data transmission delay larger in URLLC systems, especially in cases of time-varying channel conditions.

User association was implemented using cell range extension by maximizing a user utility function by Hattab and Cabric [92]. Meanwhile, a scheme jointly considering user association and scheduling for load balancing in heterogeneous networks was investigated in [93]. Also, Su *et al.* [94] proposed distributed algorithms to solve the user association problem according to the backhaul capacity constraint and BF service in a two-tier heterogeneous networks. In addition, Kalantari *et al.* [95] presented an algorithm to find a suboptimal solution of user association to maximize the sum logarithmic rate of the users in a drone scenario. However, these schemes do not consider how to quickly switch associated BSs in a fast-moving railway network. Thus, it is a challenging task on how to deploy APs on the train or along the track to make a huge seamlessly connected heterogeneous network and establish a user association model for service provisioning in a heterogeneous wireless railway network with velocities larger than 360 km/h.

In railway environments, mission critical services, operation assistance services, and passenger services should be supported by reliable mobile communication systems, as depicted in Fig. 2. Mission critical services generally include critical railway communications, train operational voice services, and operational data applications. Expected future railway mission critical systems should support intelligent transportation and control services, onboard and wayside HD video surveillance, distributed emergency communication, remote monitoring and diagnosis system, and so on. Due to the nature of the railway environment, future railway radio systems should fulfill the specific RAMS and QoS requirements demanded by railway services. In addition, mission critical services demand stronger delay, reliability, availability, and safety assurances, while additional services requirements are mainly based on the bandwidth capacity.

In general, there is limited work on network architecture, KPIs, channel models, and appropriate 5G key technologies for smart rail communication systems. However, these aspects are highly important to advance smart rail traffic systems.

## IV. NETWORK ARCHITECTURE FOR 5G-R

The emerging services for smart rail may pose special requirements for the network architecture. Based on cloud and fog computing, we can expect that open network architectures will be converted from the seamless evolution of the current networks into SDN and NFV in order to facilitate CN slicing. The software-defined air interface, together with the anticipated advanced access technologies and physical layer technologies, can enable seamless radio access, optimize business management, and deliver diversified customer services and better QoE. However, the railway network architectures discussed in the existing literature focus only on TDMA and LTE-R that are not consistent with a 5G-R network architecture.

### A. Proposed Network Architecture for Smart Railways

The requirements of the network architecture for railways are different from those of the public mobile communication network in many aspects. First, the network architecture for railways should be a dedicated communication transmission network that connects the dispatching center at all levels. This ensures that all elements interact with each other through different modes of information, such as voice calls between train driver and dispatching center, train operation control data, and video monitoring data for smart rail infrastructure. In addition, the rail communication network strives to establish an integrated and reliable emergency command system that would take real-time scene information as the decision-making base during emergencies (e.g., natural disasters or traffic accidents). Furthermore, the specific performance requirements of the train operation and control are embodied in reliable business, specialized equipment, timely transmission, and other aspects.

The biggest challenge for 5G-R is how to ensure the operational safety and reliability of the train and

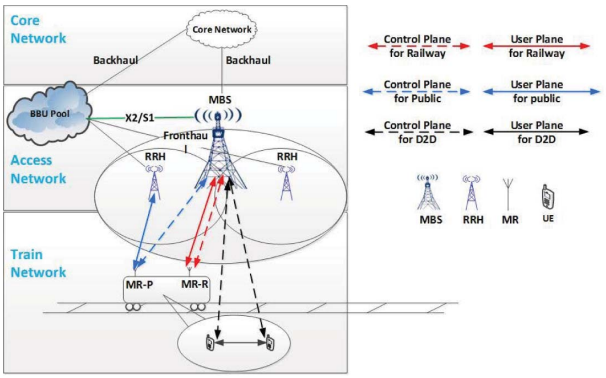


**Fig. 3.** ***Heterogeneous network architecture for smart railways.***

the reliability of MBB communication for passengers at high-moving speeds and complex scenarios. In addition, the compatibility of legacy operational management systems should not be ignored. Accordingly, to protect the current infrastructure investments and to ensure normal operation during evolution, there should be no doubt about the long-term coexistence of multiple rail-oriented transportation systems. In other words, 5G-R will certainly evolve with the legacy networks, as it maintains backward compatibility and smooth transitions.

The design of a 5G-R network architecture mainly involves three parts: access network, CN, and air interface. Unlike C-RAN, D-RAN, F-RAN, and H-CRAN, the network architecture should fully take into account real-time and high-reliability access and always-online transmission. It should be capable of executing hybrid networking using both low (below 6 GHz) and high (mmWave) frequency bands. The CN design should likewise account for 5G-R services and applications, in which the network slicing and SDN/NFV networking architecture are used. Also, the 5G-R network architecture should support both ultrareliable train operation control signal transmission and high-quality user experience that would satisfy the diverse needs of diversified services. New CN elements should be added to provide different levels of QoS. For example, a train group call register needs to be added to the CN, and this should be different from the public mobile communication network.

There are two categories of MR: one that targets the passenger services (MR for passengers, MR-P), whereas the other focuses on train-related services (MR for railways, MR-R). In the heterogeneous network architecture for railways, as shown in Fig. 3, the MBS offers seamless handover and signaling interaction, the RRH provides data transmission, and the BBU pool performs signal processing and resource allocation. The MBS guarantees the reliability of critical missions, the compatibility with current BSs, and the wide-range seamless coverage.

Basically, there are three communication models for data transmission: one that the RRH delivers, the MBS, and another one that will be supported by D2D communications. Unlike the situation of conventional BS deployment, a large number of the RRHs can be geographically distributed in a cost-effective and energy-efficient manner; thus, they are more suitable for remote deployment along the railway line. The BBU pool provides large-scale collaborative processing and enables tremendous amount of real-time connectivity. RRHs should be connected with the BBU through a high-capacity optical fiber fronthaul.

Fig. 3 shows three network domains that are based on the system resource layer, which connect the onboard access unit and the private network or the Internet. In the resource layer, the available physical resources can be virtualized into virtual resources. The radio access domain that is compatible with the current BSs integrates multiple access technologies, such as distributed heterogeneous hierarchical network, D2D connection, C-RAN connection, and satellite communications. We can expect to enable multitype and full-connectivity communications, allocate the system resource dynamically, and manage the network capability efficiently in railway scenarios.

The control domain based on the SDN and NFV establishes a componentized network. The network then decouples the $C$-plane and $U$-plane completely and flexibly reconstructs the network functions such as system control, network orchestration, and capacity openness. Accordingly, network slicing for specific use cases, reliability of train critical missions, and high QoE can all be achieved through the network function orchestration on the NFV. Meanwhile, the SDN enables the system to forward and control element separation completely. The forwarding domain concentrates on topology-aware routing and distributed forwarding of user data whereas the control domain facilitates the centralized control of the network. To meet the different demands of train-related and passenger-oriented services, the simplified design and flat distribution of the gateway equipment focus on traffic transmission and bandwidth expanding, enhancing both computation and storage capacity at the edge of network, and meeting the strict latency of time-sensitive service; therefore, they promote the full-capacity extension of the 5G-R network architecture.

In the logic view, CUPS has been clarified in railway scenarios. Subsequently, the specific network functions will be chained-up to deliver customized services, namely, the network slicing as a service. Fig. 4 depicts that the functional view consists of five layers and critical components. The modularized function layer is built on the infrastructure layer and plays a pivotal role. The service performance requirements lead to the demand for critical features, thereby providing network slicing as a service support for the upper applications. The MANO layer schedules and manages the system resources flexibly and collaboratively.

The infrastructure layer derives virtualized resources, accommodates the basic infrastructure, and facilitates the overall network service applications. Among these,

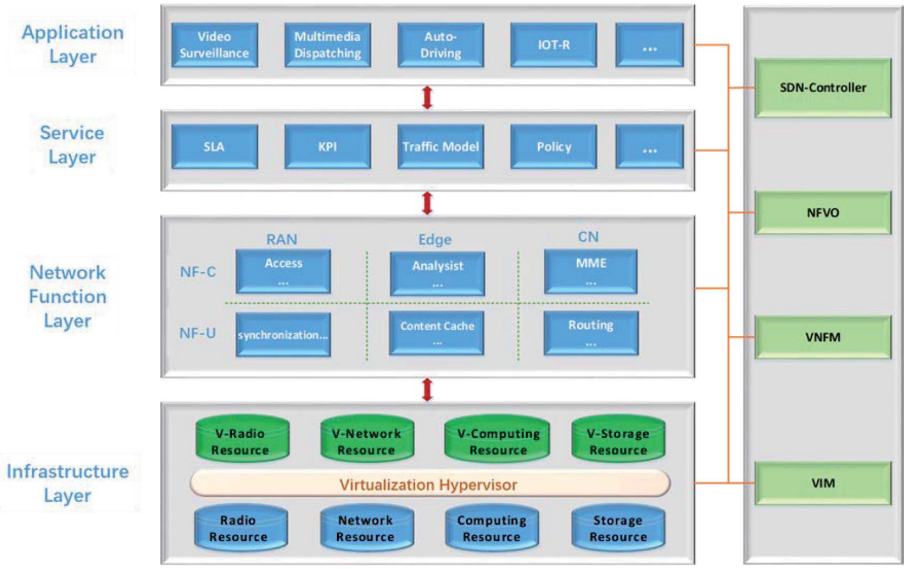


**Fig. 4.** ***Functional view of the 5G network architecture for railways.***

the hypervisor virtualizes and manages physical resource allocation.

The function layer separates the $C$-plane and the $U$-plane, decouples the network functions from the hardware, constructs the modularized functions management system, and provides an E2E network-slicing framework. The network-slicing architecture encompasses the RAN, the transport network, and the CN resources or functions to exploit the multidimensional reorganization and association of function elements, thereby meeting the specific requirements of a certain use case.

The application layer provides diversified service applications to train access units and passenger terminals, consolidated application support and performance guarantee, and extra network capacity expansion and opening based on the customized individual network slices.

The service layer defines the features and requirements of diversified services, assuring the availability and reliability for all types of train-related subscribers. As such, this should be contracted between service providers and their customers. Specifically, the demanding description pattern comprises traffic characteristics (arrival rate, average packet size, and flow type); supplementary services (firewall service and open service chains); service-level agreement (customer technical support and traffic priority); and KPIs (RAMS, latency, and robustness).

The MANO layer consists of four function entities, namely, the SDN controller, the VIM, the NFV management entity, and the NFV orchestrator element. Among them, the SDN controller manages the network elements and controls the traffic processing. VIM deals with both physical and virtualized resources in the infrastructure layer. The VNF management entity supports VNF configuration and life cycle management [96]. Finally, the NFV orchestrator element arranges both the functions and resources.

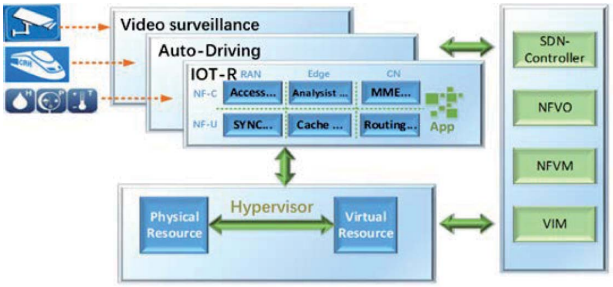


**Fig. 5.** ***Platform view of the 5G-R network architecture.***

Four functional entities guarantee that the MANO layer fulfills the life cycle management and arrangement, thus implementing the E2E network slicing for railway wireless communications [97].

Fig. 5 shows the service and application platform for railway communications. E2E railway transportation network slices stretch across the business layer; the network and the infrastructure layer function through open interfaces by mainly adopting NFV/SDN for 5G-R to provide specialized business functions and customized protocol stack and orchestration. Context awareness allows the network to adapt to the needs of applications within the framework of network constraints and operator policy. In addition, the ability of both the network and device to use context awareness can help to further enhance user experience. This ability also enables the concept of the Internet to come to the user and provide the user with the most relevant and timely information, rather than the user having to go to the Internet to retrieve information and then filter out the irrelevant pieces of information. Context awareness includes awareness of the following.

1) Network analytics, including alternative radio access techniques, network layers (macrocell, mmWave, small cell, WiFi), and the corresponding congestion levels, capabilities, and performance characteristics.
2) Subscriber analytics, including subscription attributes, wireless activity level, loyalty management status, experience analytics, historical subscriber activity, location history, current location, subscriber contacts, and application usage.
3) Device attributes and capabilities, including information on single function versus multifunction devices, device support for specialized applications, machine-type communication versus subscriber devices, and radio and network optimization capabilities;
4) Application requirements, including QoS requirements, connection reliability, access price, power consumption, and security level.
5) Subscriber preferences, including preferred access options, power savings versus performance and access cost.
6) Operator policies and subscription context, including allowed services, service, attributes, and QoS.

Context information may be gathered from the device, network monitors, network elements, network databases, and analytics platforms. The network processes the context information when a device attaches to it or when an application is invoked. This determines the service attributes that govern how the network will treat the device and application. The service attributes for access may, for example, include cost, reliability, power consumption, security level, QoS, and mobility. The service attributes for access can be mapped to configurable 5G features, which the network assigns. For example, context information may determine that low-cost access, with no support for active mobility and long battery life, is best for providing service to a nomadic sensor device that attaches to the network. As a result, the network configures connectionless access with low priority, simple IP networking with no tunneling, and an idle-mode wake-up period of one day.

## B. Network Slicing for 5G-R

Network slicing is the key technology allowing to apply NFV to 5G. We can anticipate that a 5G-R network will be able to support a diverse set of emerging smart rail services and applications. Therefore, network slicing can support the multiple QoS needs of railways. Based on common infrastructures and on a range of techniques, network slicing can offer isolated and dedicated logical networks to certain users. However, the network slice types need to be further excavated and enriched.

In contrast, to our knowledge, network slicing for railway communications has yet to be investigated. In this section, we investigate how network slice types can be identified in terms of their service and use case characteristics. The proposed design process of the smart rail network slicing has the following procedures: network slice formation, network slice selection, network slice switching, user state maintenance, and new function identification.

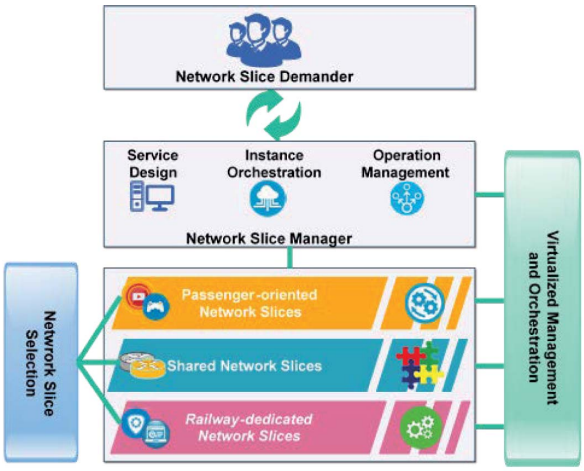


**Fig. 6.** *Network slicing based on the architecture of 5G-R.*

As shown in Fig. 6, the network-slicing architecture for 5G-R contains two basic components: the slicing management and the slicing selection. The network-slicing management module is the core technology of the network-slicing operation maintenance. First, the user requirements at the service level are transmitted to the NFV/SDN-based cross-domain deployment descriptor and service profile, and then the whole life cycle of the network slice is managed to ensure the QoS requirements.

The network-slicing management function could provide an isolated, secure, and highly automatic E2E logical network to different kinds of service demanders, such as vertical industry users, virtual operators, and enterprise users. The process has three stages as shown in Fig. 6.

1) *Business design stage:* In accordance with the network slice templates and design tools, service demanders set the related parameters of the network slicing, including network topology, function component, interactive protocol, performance indicators, and hardware requirements.
2) *Instance orchestration stage:* A network slice management function sends the slice description file to the NFV MANO function module for the network slice instantiation. Furthermore, it will hand out the network element function configuration information via the interslice interface. Finally, it starts the connectivity test through migrating to active state.
   a) *Implementation MANO:* An example of a management and choreography entity belonging to NFV; NFV MANO is broken up into three functional blocks.
      i) *NFV orchestrator:* Responsible for onboarding new network services and VNF packages, network service life cycle management, global resource management, and validation and

authorization of NFV infrastructure resource requests.

ii) *VNF manager:* Oversees the life cycle management of VNF instances and the coordination and adaptation role for configuration and event reporting between NFV and E/NMS.

iii) *VIM:* Controls and manages the NFV infrastructure compute, storage, and network resources.

3) *Operation management stage:* While in a running state, the slice owner can perform real-time monitoring and dynamic maintenance of the slice with relevant slicing management functionsincluding dynamic allocation of resources, design, creation, deletion, and update of the network slicesas well as the fault-alarming module.

The network-slicing selection module maps the user terminal to the appropriate network slice. Considering services subscription and functional properties and other various factors, it is capable of provisioning proper network-slicing for different user requirements. Shared or separate network slices can be accessed by different users. Also, the slicing selection contains RAN slicing (LTE or wired networks) and CN slicing (various combinations of $C$-plane and $U$-planes from different CNs). Industrial network slice instances include intelligent control system for railways (in need of low latency and high reliability). This control system is required for dedicated RAN and specific CN and for broadband communications, which is required for proper public mobile network in cooperation with trackside private network for railways.

E2E network slices connect the UE, access slices, CN slices, and the mapping relationship between the slices into a complete logical network. The slicing selection function routes data transmission from RAN slices to appropriate CN slices. The CN slices compose a set of supporting service functions that can either be shared or be exclusive to certain E2E network slices. The virtual LAN technology of virtual-switch can enable service isolation inside a network slice whereas the virtual-switch technology of the virtual-router can offer the business isolation between network slices.

*Service slicing category:* The nine types of network slicing are summarized as follows.

1) *Auto-driving slicing:* Shows strict requirements for high safety, high reliability, and low latency.
2) *D2D slicing:* Establishes direct communication links between the interactive devices and enables low-latency and large-capacity data transmission.
3) *MBB slicing:* Provides travelers with access to MBB communication networks and supports interactions inside the coaches.
4) *Ultrahigh reliable slicing:* Attends to those services that need ultrahigh reliability.
5) *Ultrahigh available slicing:* Fits with applications that call for ultrahigh availability.
6) *Ultra-HD slicing:* Capable of ultra-HD video streams.

Meanwhile, the three types of IoT slicing are as follows.

1) *Massive IoT:* Monitors the status of the massive number of static sensors of the fundamental fixed infrastructure and of the train equipment and requires high reliability.
2) *Scheduling IoT:* Delivers the information of train scheduling and crew arrangement with high flexibility.
3) *Asset IoT:* Responsible for passengers luggage, cargos, and containers and also supports railway ticketing.

## V. CHANNEL MODELS AND OPERATING FREQUENCY BANDS FOR SMART RAIL

The characteristics and models of wireless channels for smart rail are critical for evaluating 5G-R key technologies. Among other decisions, the operating frequency bands suitable for rail communication systems are determined based on the channel models. In this section, we investigate the channel models and operating frequency bands for smart rail.

### A. 5G Physical Communication Scenarios for Smart Rail

Due to the variety of applications and deployment scenarios, 5G networks need to support extremely varied requirements and propagation channels with different properties and attributes. For example, the path loss and multipath effects are different in different propagation environments. Thus, radio wave propagation scenario partitioning is a fundamental step for wireless channel modeling. Although the International Mobile Telecom System-2000 [98], the Universal Mobile Telecommunications System developed by 3GPP [99], and the WINNER project group [100] all define typical scenarios, none of them specifically relate to railways. On the other hand, railway scenarios (e.g., cuttings, viaducts, tunnels, and crossing bridges) significantly affect the propagation characteristics. Therefore, developers need to establish a detailed scene-partitioning scheme for railways to improve the quality of wireless network planning and optimization. Based on practical measurements of typical railway lines and railway stations, and considering some other factors, such as physical and user attributes, we have defined 16 scenarios for railways [101].

Another important aspect in evaluating 5G-R key technologies are the communications scenarios, which mainly include train-to-infrastructure, intercarriage, intracarriage, inside station, infrastructure-to-infrastructure, and train-to-train.

1) *Train-to-infrastructure:* The links are between the AP/transceivers of the train and the infrastructure nodes of the fixed networks. The links deliver bidirectional streams with high data rates and low latencies and provide robust communication links with

latencies lower than 10 ms, together with an availability of 99.9%–99.999%, while moving at speeds of up to 500 km/h.

2) *Intercarriage:* A wireless network operates in-between carriages since it is very costly to wire a train for network access. Likewise, rewiring when a train needs to be reconfigured is inconvenient. This scenario of a wireless network requires a high data rate and low latency because the APs are arranged in each carriage such that each AP serves as a client station for the AP in the previous car while also serving as an AP for all UEs within its car.
3) *Intracarriage:* The links provide wireless access between the APs in the carriage and the passengers or the sensors of the equipment inside the carriage. In this scenario, real-time HD videos need to be accessed with low latencies.
4) *Inside station:* The links provide wireless access between the APs and the UEs in railway stations. Passengers can have access to MBB communication services. Furthermore, stations provide communication infrastructure to support commercial (e.g., cash desks) and operational services (e.g., automatic doors, surveillance, and fire protection).
5) *Infrastructure-to-infrastructure:* HD video and other information is transmitted in real time between multiple HDTV IP/HD-SDI cameras. The APs are deployed at station platforms and at the wayside along rail tracks as a high date rate WB or as IoT. The infrastructure nodes are connected in real time and interactive; they are supported by bidirectional data streams with very high data rate and low latencies.

## B. Channel Characteristics and Models Below 6 GHz in Smart-Rail Scenarios

In recent years, researchers and developers have conducted extensive sub-6-GHz channel measurements in various railway environments. Several studies on railway propagation channels focus on narrowband fading and the behavior of composite channel parameters. For example, the analytical model in [61] provides a fit for path loss, shadowing, $K$-factor, and other statistical channel parameters. Meanwhile, Guan *et al.* [102] and He *et al.* [103] studied the impact of viaducts, cuttings, stations, and cross-bridges on path loss and fading. Based on the extensive channel measurements conducted along the Zhengxi railway line, the PLEs are estimated, for example, PLE of the viaduct scenarios is about 3.5, and the PLE of the cutting scenarios is around 4.3. Meanwhile, He *et al.* [104] proposed a path loss model at 930 MHz based on Hata's formula that is modified with two correction factors to ensure a sufficient fit in different environments.

For the amplitude distribution of small-scale fading in railway scenarios, in most cases, the Rician distribution provides the best fit [67]. Therefore, the Rician $K$-factor can be used as a useful parameter in assessing the severity of fading and the resulting impact on system reliability. Measurements show that one can model the dB-scale $K$-factor as a Gaussian random variable that has mean $m_k$ and STD $\sigma_s$ that are dependent on the location within the cell. The parameters of the distributions are summarized in [103].

## C. Channel Characteristics and Models at mmWave Bands in Smart Rail Scenarios

Recently, researchers have gradually shifted their focus from railway channel measurements to ultrahigh frequency (mmWave) range, emphasizing multidimension channel characteristics, for example, delay, Doppler, and spatial domain.

Prototype experiments on vehicular environments have confirmed the viability of highly mobile mmWave communications. For example, Samsung achieved in 2014 a rate of 1.2 Gb/s on a vehicle-to-BS communication link at 110 km/h in the 28-GHz band [105]. In 2015, Korea Electronics and the Telecommunications Research Institute achieved 1-Gb/s transmission rate in the 30-GHz band in the Seoul subway [106]. However, due to the lack of channel measurement data, simulation technology, and channel models, neither academia nor industry can fully evaluate and optimize their system performance under various HSR communication scenarios at higher speeds. The challenges include the exploration of static or dynamic radio propagation mechanism and multipath birth and death mechanisms, mmWave large-scale antenna channel characteristics, and a modeling theory for HSR scenarios.

Yang *et al.* [107] and He *et al.* [108] measured the mmWave propagation in railway scenarios, and accordingly studied the path loss characteristics. However, the measurement data do not provide a full picture. Moreover, the measurement data validate only the path loss model while the other channel characteristics are not verified since relevant data were lacking.

Other effects that have been mostly ignored up to now and thus will require particular attention are as follows: 1) the elevation characteristics of the propagation (since radio waves propagate in three dimensions and reflection/scattering disperses radiation in azimuth and elevation) and polarization and 2) 3-D channel measurements are necessary, especially in designing and evaluating smart antenna and massive MIMO. Similarly, polarization needs to be taken into account to assess its impact on the possible diversity and multiplexing gain.

Existing mmWave directional channel measurements were not performed under dynamic environments due to a lack of measurement equipment. A few exceptions exist, but none of them were done in railway environments. There are no measurement results under such circumstances since it is difficult to build suitable channel sounders for railway mmWave applications. Hence, in railway scenarios, comprehensive mmWave channel measurements are needed to establish and/or validate suitable channel models.

**Table 2** Definition of Six Modules for 5G mmWave Railway Channels

| Module index | Module 1 | Module 2 | Module 3 | Module 4 | Module 5 | Module 6 |
|---|---|---|---|---|---|---|
| Definition | Tunnel entrance on steep wall connecting the cutting to the crossing bridges | Viaduct with open train station | Urban with semi-closed train station | Rural with CCT tunnel | Rural connecting double-track tunnel | Single-track viaduct |
| Special objects | Steep wall, cutting walls, crossing bridges | Open train station, indicators | Buildings, semi-closed station | Vegetation, CCTs | Dual-track tunnel, barriers | Single-track viaduct, barriers |
| Common objects | Trains (metal), tracks (concrete), pylons (metal), traffic signs (metal), billboards (metal/LED/concrete) | | | | | |

To verify new communication regimes in railway environments, we first need to define 5G mmWave railway scenario modules together with their distinct propagation features. Six modules, representing various typical railway scenarios, were constructed in [109]. Note that the coverage ranges for which these modules are valid range from 100 to 200 m, which is much shorter than the macrocell or microcell range with carrier frequencies below 6 GHz. The high isotropic path loss of mmWave bands, which is partially compensated by BF gain, considerably shortens the link length, thereby providing us with the chance to concentrate on the six modules in order to compose a comprehensive smart rail environment. As summarized in Table 2, apart from common objects (e.g., trains, tracks, pylons, traffic signs, and billboards), each module includes its special objects, that is, steep walls, cutting walls, crossing bridges, train stations, indicators, barriers, vegetation, CCT, dual-track tunnel, and single-track viaduct. The 3-D models of both the comprehensive scenario and the six modules are publicly available and freely downloadable (see http://raytracer.cloud).

*Module 1:* Tunnel entrance on steep wall connecting cutting with crossing bridges: In a railway environment, a tunnel is an artificial underground passage that is specifically built through a mountain. Thus, when the mountain is high, rail builders usually construct steep walls to protect the tunnel entrance. Such steep walls are huge reflectors for wave propagation. Cutting is another common scenario in railways. A rail cutting is a man-made valley that carries the track as its base; it is used to pass through the hill that is not particularly high. Cuttings are made of concrete and stone with some vegetation on the surface. The depth of the cutting is usually 3–10 m. The slope height is about 13–14 m, and the inclination angle is 35°–40°. The distance between the two bottoms of the slopes is 12–13 m. The most common semi-closed obstacles in cutting walls are crossing bridges. They can block the LoS connection and generate strong multipath propagation.

*Module 2:* Viaduct with open train station: A viaduct is a long bridge across the uneven ground found in rural or urban settings. It is a common scenario in railways; for example, viaducts make up almost 70% of the Beijing–Shanghai railway line. We assume that we can obtain a LoS propagation condition most of the time. This module is composed of an open train station (awnings that only cover the platforms and not the rails) and a viaduct in a rural environment. We note that there are three types of modern train stations: closed type, semi-closed type, and open type. The first type is similar to the typical indoor environment whereas the latter two types are more particular. Usually, smaller, open stations appear in rural scenarios whereas larger, semi-closed stations are constructed in urban environments. Thus, this module is composed of an open train station and a viaduct in rural areas to represent the features of these two structures.

*Module 3:* Urban with semi-closed train station: The urban scenario represents the propagation in urban areas when the train is entering, leaving, or passing the city. Most of the buildings are higher than 10 m and are at least 10 m away from the barriers. In the semi-closed scenario, the huge awnings are usually designed to keep the rain from the passengers and trains. Thus, the channel will exhibit strong multipath because the semi-closed station, barriers, and buildings are present in this module.

*Module 4:* Rural with CCT tunnel: The rural scenario represents the environment where there is a large range of open areas, very few buildings, and where certain vegetation grows adjacent to the track. A CCT tunnel is a method for building shallow or short tunnels, in which the builders excavate a trench and then roof it over with an overhead support system. The overhead support system should be strong enough to carry the load of what needs to be built above the tunnel. CCTs are built in trains for two reasons: 1) to prevent potential landslides and 2) to cover the rail as a defense against wind and other unfavorable climate conditions. In fact, a CCT is not a complex construction method; however, it is challenging for wireless

planners and system designers of railway communications. Thus, in this module, the CCT, barriers, and the dense vegetation adjacent to the barriers are the main structural and influential factors that need to be evaluated.

*Module 5:* Rural environment connecting double-track tunnels: The cross section of double-track railway tunnels is usually vaulted or built in a semicircle fashion, with a height of 5–10 m and a width of 10–20 m. Two trains have the chance to pass by each other inside the tunnels, and the effect of this situation needs to be further evaluated. Thus, we design this module such that we can study the channels that connect a double-track tunnel in rural settings.

*Module 6:* Single-track viaduct: The difference between two single-track viaducts and one double-track viaduct is the presence of barriers between the two tracks. In this module, we can further determine the influences of barriers and railway crossing. However, the barriers and tall objects (e.g., tall trees or billboards) on both sides of the viaduct still affect propagation. The train station is another kind of important structure along the railways. The huge awnings, LED indicators, steel frames beside and above the track, and metallic pylons can block parts of the LoS, and consequently influence the channel. Note that if the target cell includes only a part of the scenarios, the users can further divide this module into more submodules, such as pure viaduct and pure open train station.

MmWaves are highly sensitive to the propagation environment and mobility. As such, multiple antenna technologies adopted in the mmWave band, such as BF, have very strict requirements for the spatial resolution of the channel. As discussed above, it is difficult, expensive, and complex to perform dynamic channel measurement at mmWave bands. This makes it impossible for us to rely only on measurements in order to obtain comprehensive and elaborate channel properties (e.g., time, delay, and frequency domains). Accordingly, the design of BF algorithms and other system design questions can be investigated based on detailed simulations of electromagnetic propagation.

The most suitable way to simulate radio wave propagation at high-frequency bands is through RT, which approximates electromagnetic waves as quasioptical rays. Propagation at high frequencies is similar to optical rays; thus, an RT approach is inherently suited to obtain highly accurate channel simulation and modeling at mmWave or even THz frequency bands. Different rays that correspond to different multipath components (MPCs) are traced from the Tx to the Rx. Each MPC is characterized by multiple output parameters: time delay; real part of field intensity; imaginary part of field intensity; AoD azimuth; AoD elevation; AoA azimuth; AoA elevation; type (transmission, reflection, scattering, and diffraction); and number of reflections or other interactions.

Although RT is well suited for high frequencies, both the calculation complexity and the simulation time exponentially increase as the number of objects and faces in a 3-D-scenario increase. Yet the required resolution

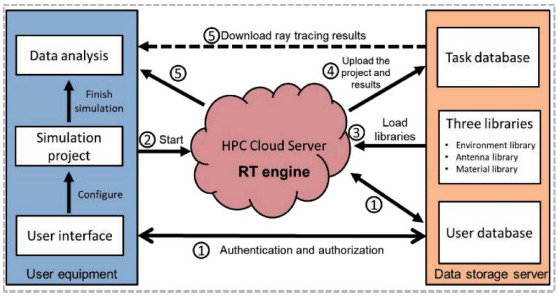


**Fig. 7.** ***RT simulator based on HPC.***

(number of faces representing complex objects) increases with frequency. This makes it much more challenging to achieve accurate and efficient RT simulations for railway communications at mmWave and THz bands than when simulating at traditional frequencies below 6 GHz.

One promising solution to address this challenge is to transfer an RT simulator from a personal computer to HPC platforms. This can make RT much more efficient without sacrificing accuracy. Thus, we develop a high-performance RT supercomputing platform. All the components are connected via the Internet such that the components can exchange commands and data. As shown in Fig. 7, the platform is composed of five layers: data input layer, data transmission layer, data storage layer, data analysis layer, and application layer. Developing the smart rail RT includes building the RT database and material parameters, RT correction, antenna modeling, RT calibration, and channel modeling based on the RT simulation.

To build the RT database, we first need to parametrize the target smart rail scenarios. We then get the typical objects and materials from the scenario model. According to the parameters (electromagnetic parameters, roughness, etc.) of the individual models of radio propagation mechanisms, (i.e., LoS, reflection, scattering, transmission, and diffraction), we need to construct the material library for smart rail systems (frequency band up to 100 GHz). For different antenna types (single antenna, multiple antennas, and large-scale multiantenna) and different array types (linear array, planar array, cylindrical array, etc.) that work in different frequency bands and operation mode, we need to study the antenna modeling approach and then establish the antenna library. The material library can be calibrated by comparing the RT results and sample measurements; the latter do not need to be performed at full train speed but can be performed quasistatically. We then use the SAGE algorithm to extract MPC parameters from the measurements and then calibrate the RT toolbox based on the reflection mechanism [109].

The data input layer mainly collects various data including environment information, antenna patterns, Tx and Rx deployments, frequency, and channel measurements, and accordingly transforms them into predefined structures. The data are then sent to the data storage layer (cloud),

which contains scenario library, material library, antenna library, user database, task database, and measurement database. On this basis, the data analysis layer can perform channel simulation and modeling for the specified site using acceleration algorithms (e.g., parallel computing and space partitioning), thereby making the RT simulation more accurate and efficient. The output of the data analysis layer can then be applied to the system-level simulation, network planning, and antenna design. Thus, with the aid of our CloudRT platform, designers can accurately study the channel modeling in railway scenarios at both mmWave and THz frequency bands.

The electromagnetic calculation is composed of the different radio propagation mechanisms (i.e., LoS, reflection, scattering, transmission, and diffraction), which highly depend on the material of the object. Therefore, the material parameters are important to the accuracy of the RT simulation. The ITU provides dielectric parameters of typical building materials at 6–100 GHz [110]. Due to the complex composition of the dynamic scenario, it is unrealistic to measure the different propagation mechanisms of all materials. Therefore, the existing literature and works cannot cover all the material parameters that affect the propagation mechanism in all 5G-R scenarios. Given such a problem, Zhang *et al.* [111] proposed a method for calibrating the RT simulator based on channel measurements using a simulated annealing algorithm. The parameters of individual simulated MPCs are compared with those of the corresponding measured MPCs, and the ultrawideband electromagnetic parameters of typical materials in the target scenarios are obtained at the end. Based on the path loss measurement at 90 GHz in a railway viaduct scenario, similar calibration works are also performed in [112]. The difference between the fit path loss coefficients and the calibrated RT and the measurement is 0.01. With this, the system performance can be analyzed further for various configurations with higher accuracy. However, the accuracy of RT depends not only on the accurate material characteristics but also on the accurate geometry model of the propagation scenario and accurate superposition of multiple propagation mechanisms. The existing calibration methods assume that only a part of the measured propagation mechanisms exists and that the geometry of the environment model is built correctly. Besides, the referred measured MPCs mainly include the delay and power domain. In the future, researchers should consider all sorts of propagation effects and accordingly calibrate RT under different domains with the flexibility of adjusting the preconstructed geometry model. Owing to the development of high-resolution channel estimation algorithms, a sufficiently fine resolution of the measurement results is possible, thus allowing for a more accurate analysis of the impact of geometric features of the scenario and the material characteristics.

In terms of channel modeling, CloudRT supports the dual-mobility characteristics of a mobile scenario (i.e., dynamic scatterers and dynamic Tx/Rx). The high computational complexity of RT has made research on its acceleration method a hot topic in the computer field. The typical approaches used include dimensionality reduction [113], space partitioning methods [114], ray-launching methods [115], and hardware- acceleration methods [116]. The basic principle of the dimensionality reduction method is to simplify the 3-D model in the elevation domain and then change the model to 2-D or 2.5-D. Although such method makes it much faster to compute multipath tracking, it makes the RT simulations less accurate. The idea in space partitioning is to divide the radio propagation scenario into several small regions such that researchers can avoid traversing all the objects in the scenario model for detecting MPCs. The space partitioning methods have no effect on the accuracy of the RT simulation; however, these methods mostly apply to a communication scenario where the scatterers are static. However, for 5G mobile communication scenarios for which the "dual mobility" feature exists, the appearance and disappearance of scatterers dynamically change the spatial distribution and the visibility relationship between the geometry planes, which makes performing RT more difficult. This also applies to ray launching methods. Accordingly, Nuckelt *et al.* [117] proposed a method for obtaining densely sampled channels by interpolation, which will lessen the RT simulation task and ensure that the simulation results are accurate. However, the interpolation algorithm in [117] focuses only on the amplitude domain and does not include the angular domain and polarization domain. Also, there is no theoretical analysis of reasonable sampling intervals. As the advantage of the hardware acceleration has already been proven in the aforementioned part, we need to combine new acceleration algorithms and hardware acceleration to make the RT calculation less complex.

Specifically, we propose to characterize the sparsity of dynamic scatterers and to determine the quasistationary interval in the specific scenarios. With regard to the sparsity characterization of dynamic scatterers, the global scattering center of the complex objects in the real mobile communication scenario can be extracted based on the full-wave analysis and image processing algorithms by setting a reasonable threshold. This method can be used to reveal the variation of the sparseness of the scatterer with frequency. Moreover, this method changes the modeling of complex objects (e.g., trains) from modeling deterministic polygons to multiparticle swarms that include explicit field strength and phase information in order to simulate RT accurately and efficiently under a complex dynamic scenario.

To determine the quasi-stationary interval in the spatial domain based on the geometric parameter transformation and spatial segmentation method, the CloudRT extracts the geometric features of a large number of similar scenarios from the scenario library. The relationship between the spatial geometric characteristics and the different propagation mechanisms is mapped through deep learning

methods. We can determine the quasi-stationary interval in the spatial domain using a predefined threshold and attain mathematical interpolation for different propagation mechanisms within a spatial quasi-stationary interval. Through this method, the sampling interval of dynamic RT simulations significantly lessens while maintaining accuracy. The railway channel characteristics at the 60-GHz band with 8-GHz bandwidth in the six scenario modules that use the two antenna height setups can be found in [118] and [119].

### D. Massive MIMO Channel Modeling in Smart Rail Scenarios

Many research institutes and organizations have conducted channel measurements on massive MIMO under urban or suburban LoS and NLoS scenarios. The moving velocity of the train (Tx or Rx) is high in HSR scenarios. The virtual array technique for (massive) MIMO channel measurements cannot be applied to such time-varying environments (with the exception of using the train movement itself to create a massive virtual linear array at the train side). On the other hand, although one can apply a real array with demodulator chains to each receive antenna element, this would increase cost and would make it more difficult to implement. To this end, a switched array seems to be the best compromise between measurement time and hardware effort. In such a switched array, different antenna elements are connected to a demodulator chain (conventional channel sounder) via a fast RF switch. This would enable developers to measure dynamic massive MIMO channels. However, massive MIMO channel measurements in railway scenarios have the following five key challenges.

1) *Signal power or maintaining adequate SNRs and coverage:* The use of switched antenna arrays implies a lack of BF gain. The measurement SNR might thus be low, particularly in the cell edge of railway scenarios. A promising solution to address this problem is to use beam switching [120]. In this case, multiple antennas are activated simultaneously. If each antenna is equipped with a power amplifier, the effective conducted transmit power can be increased while realizing BF gain.
2) *Clock synchronization:* Clock synchronization is important in massive MIMO measurements to ensure that the frequency and/or signal phase of each channel is consistent with each other. In the dynamic railway scenarios, it is not possible to share the clock with the cable in the same way as that done in indoor measurements. Therefore, a standard GPS clock (to receive GPS pps signals from satellites) plus atomic clock, such as a rubidium clock, needs to be used in both Tx and Rx to keep precise synchronization.
3) *Sample rate:* In measuring dynamic mmWave channels under smart rail scenarios, the Doppler effect is expected to be significant due to high carrier frequency and high movement speed. Thus, a high sample rate at the Rx is required to successfully measure the fast time-variant mmWave channel and acquire its Doppler characteristic. The Nyquist theorem in the time domain must be fulfilled. The sample for each subchannel of the mmWave massive MIMO system needs to be taken frequently to track changes in the dynamic channel. For a regular switching structure (each antenna measured in a regular sequence), this means that the switching cycle needs to be finished within the coherence time of the channel. This is challenging, particularly for mmWave massive MIMO channel measurements because the BS can have hundreds of antenna elements and more time is needed to measure all subchannels. Quasirandom switching sequences show promise to alleviate these problems; however, they have not yet been used in railway channel sounding applications. Moreover, the bandwidth of the excitation signal in the mmWave massive MIMO techniques always exceeds several hundreds of Megahertz. This large bandwidth also requires a high sample rate of the Rx.
4) *Fast switching:* As mentioned before, the measurements of the whole massive MIMO subchannels need to be finished within a limited time to ensure that all the measured data are within the same coherence time. Therefore, each subchannel needs to be quickly measured and fast switching is highly required. The measurement duration at one antenna element is determined by the actual subchannel measurement time and the time it takes to switch from one antenna element to the next. Obviously, the latter should be as short as possible, particularly in massive MIMO applications with hundreds of antenna elements. Thus, a switching system with a large number of ports, low insertion loss, and a short switching time is required. Note that since a single switch has limited RF ports, a cascade connection scheme of switches is required to build the above switching system. However, such scheme results in more challenging issues such as system calibration and synchronization. Moreover, Gb/s real-time data flow storage is needed in the measurement system to record data in real-time from the switching system.
5) *Dynamic BF tracking:* For mmWave massive MIMO measurements, BF technique is preferred to increase the gain. For the conventional static mmWave measurements with limited angular scanning range, horn antennas with large gains can be used for power concentrating. The transmitted signal beams are manually aimed at the Rx. In the dynamic scenarios of railways, how to accurately track the beam is significantly difficult. The channel sounding system should establish a feedback mechanism that can control the transmitter to adjust the beam angle according to the varying location of the Rx.

Based on the RT simulation platform, we mainly focus on analyzing the influence of massive MIMO configurations at the mmWave band inside the railway viaduct. The configuration is suggested by 3GPP and the propagation parameters are calibrated in our previous work [121].

### E. 5G Operating Frequency Bands for Smart Railways

Before we discuss potential frequency bands for the smart rail, we first need to reference the suitability of spectrum bands for the 5G generic services as defined by the European Union METIS: xMBB, mMTC, and ultrareliable MTC (uMTC).

1) xMBB requires a combination of frequencies, comprising lower bands below 6 GHz for both coverage and data traffic purposes and higher bands above 6 GHz with a large contiguous bandwidth. These requirements are needed to cope with the ever-increasing traffic demand, including WB solutions. An exclusive licensed spectrum essentially guarantees coverage and QoS, and this can be supplemented by a spectrum that other licensing regimes authorize to increase the overall spectrum availability. Such licensing regimes include licensed shared access or unlicensed access (e.g., WiFi offload) or new enhanced unlicensed access schemes (e.g., license-assisted access).
2) Some mMTC applications are most suited for frequency spectra below 6 GHz, particularly spectrum below 1 GHz, in cases when large coverage areas and good penetration are needed. Although an exclusive license is preferred, other licensing regimes may have to be considered depending on the specific application requirements.
3) Licensed spectrum is considered most appropriate for uMTC. For safety and for V2V and V2X communication, an option could be the frequency band 5875–5925 MHz harmonized for ITS. Another option could be the sub-1-GHz spectrum, which is particularly well suited for high-speed applications and rural environments.

NGMN requirements for 5G include the capacity to support a wide range of applications that may have differing requirements for the underlying mobile connectivity. This will consequently require access to a range of spectrum bands with differing characteristics in order to address a wide range of requirements for coverage, throughputs, and latency in the most cost-efficient manner and to enable the spectrum to be used effectively. Spectrum bands already licensed to mobile network operators will form an essential foundation for 5G mobile services. As such, it is important to allow operators to "refarm" existing spectrum bands to 5G technology according to their deployment strategy, which helps to improve SE and to introduce new capabilities. It will also help in planning the necessary long-term investments.

Further spectrum of MBB for both coverage and capacity may need to be organized in the future. A spectrum below 1 GHz is particularly useful for coverage, especially in indoor settings and rural areas. A spectrum above 6 GHz is particularly useful for supporting very high data rates and short-range connectivity.

In the view of the future demand of smart rail, the single bandwidth resource at 450 MHz will hardly be sufficient for the diverse potential services envisioned in evolved railway communication system. The available spectrum resources in TDD or FDD modes (at 20 or 25 MHz) can meet the concurrent traffic requirements of train-related services in all scenarios under the present communication system. In special scenarios (e.g., grand passenger train stations and marshaling stations), the frequency planning that supports massive-bandwidth requirement services adopts a spot-line combined approach. As such, this allows operators to attend to the application for shared licensed spectrum at 1.8 GHz for vertical customers or to utilize unlicensed spectrum at 2.4 or 5.8 GHz for complementation straightforwardly.

Researchers need to evaluate the possible candidate bands in higher frequencies to address the requirements for new spectrum for UDNs in 2020 and beyond. The industry needs such frequencies such that very wide bandwidth channels can support very high data rates and short-range mobile connectivity (e.g., 500–1000 MHz of contiguous spectrum per network to support the multitude of services). The total spectrum requirements also have to consider the potential need to accommodate multiple networks. Therefore, it is necessary to further develop the technical feasibility of the ranges between 6 and 100 GHz, particularly those ranges where primary allocations to mobile, as stated in the ITU radio regulations, already exist. Therefore, the lower limit for the band range (above 6 GHz) should be further assessed.

According to our simulation analysis, the low frequency band of 3.5 GHz and the mmWave band of 25–30 GHz can cover a 1-m railway line using directional antennas and reasonable transmitting power. The mmWave band can support higher transmission rate and larger bandwidth required by the dedicated communication for the smart rail scenarios. Considering spectrum utilization, smart rail services, propagation characteristics, BS deployment, frequent handover, and other factors, a reference frequency band 26–38 GHz can be adopted for smart rail communications.

## VI. 5G KEY TECHNOLOGIES FOR SMART RAIL

One of the key challenges for smart rail is achieving reliable high data rate transmissions at high speed under various railway scenarios. Providing transmission rates of up to Gb/s, and/or reliability of up to 99.9999% in high mobility is a very significant technical challenge, which is compounded in railway communications where speeds of up to 500 km/h occur. The physical-layer techniques

needed to achieve those goals are parallel to those that are used in 5G, in particular MIMO (including massive MIMO and distributed MIMO), as well as relaying and design of suitable codes and link margins for enhancing reliability. However, the special aspects of high mobility need to be considered. Sections VI-A–VI-H discuss all these aspects.

### A. ST-ASM in Railway Scenarios

SM is a digital modulation concept for multiantenna wireless systems, which has been introduced to increase SE while allowing for a low-complexity implementation. SM achieves multiplexing gain by mapping a block of information bits onto two information-carrying units: one symbol is chosen from a conventional constellation diagram and another unique transmit antenna index is chosen from the so-called spatial constellation diagram. At last, the point of the signal constellation diagram is transmitted through a single active antenna belonging to the spatial constellation [122]. At the Rx, the optimum maximum likelihood (ML) detector can simultaneously estimate the active Tx antenna index and the selected modulation symbol index. As a result, the overall SE increases (compared to transmission from a single antenna) by exploiting the transmit antenna index. Using SM in smart railway can ensure system robustness and reliable services.

It follows from the above description that in SM, the data transmission rate is proportional to $\log_2(N_t M)$ bits/s/Hz, where $N_t$ is the number of the transmit antennas, and $M$ denotes the MO. We can thus improve SE by increasing the number of transmit antennas. SM is different from a conventional MIMO transmission; in the latter, we can improve reliability through diversity and increase channel capacity by spatial multiplexing. In contrast, the former maps information to the Tx antenna index to obtain additional multiplexing gain.

When using SM, there is only one or a few active transmit antennas. Thus, only a single-RF chain is required; this reduces complexity and energy consumption, which make the commercial communication system more economical than conventional MIMO methods. This single RF scheme can also completely remove the intersystem interference at the Rx because only one Tx antenna is activated for any symbol duration. On the other hand, SM with multiple Rx antennas shows improved robustness because it offers excess degrees of freedom that can be used to cancel signals from man-made interference including intentional jamming.

The transmission of an $N_t \times N_r$ MIMO system can be represented as

$$y = \sqrt{\rho} H x_q + w \tag{1}$$

where $\rho$ is the average SNR at the transmit antenna and $H$ represents the $N_r \times N_t$ channel matrix, in which $h_l$ is the $j$th column vector of the channel matrix. Finally, $w$ denotes the $N_r$-dimension AWGN with i.i.d. elements

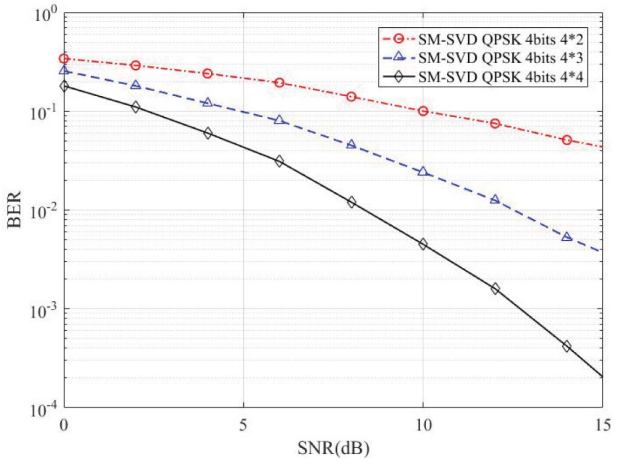


**Fig. 8.** ***BER performance versus SNR evaluated using SM-SVD approach for the case of a 4-b/s/Hz transmission with different receive antennas. The number of the transmit antenna is 4 and the number of the receive antenna is 2, 3, and 4, respectively. All of the three schemes employ a QPSK modulation alphabet.***

and $x_q$ denotes the transmitted symbol. The decision of the ML detection is given by

$$\langle \widehat{j}, \widehat{q} \rangle = \arg\min_{j,q} \|y - h_j x_q\|_F^2 . \tag{2}$$

Considering the large additional gains of SM with more Tx antennas, railway communications might be an especially beneficial application of SM. The roof of the train carriages can provide enough space; to deploy a few hundred antennas. Li *et al.* [123] proposed a different detection method called the SM-SVD. This method has a low complexity and is suitable for a massive MIMO system. The detection method includes two independent estimation steps and the estimation of the antenna index can be written as

$$\widehat{j} = \arg\min_{j} \arccos \frac{|\langle h_j, y \rangle|}{\|h_j\|_F \|y\|_F} \tag{3}$$

where $\langle \bullet \rangle$ denotes the inner product in the Hilbert space and $\| \bullet \|_F$ represents the Euclidean norm. Once the index of the transmit antenna is estimated, the transmitted symbol can be estimated in the traditional way.

Let us assume the channels are independent Rayleigh flat-fading and the noise is additive Gaussian. We then evaluate the SM using different receive antennas for a 4-b/s/Hz transmission. Fig. 8 shows the numerical results when we employ the SM-SVD. We can see that the SM scheme with four Tx antennas performs the best, which demonstrates that the extra Rx antennas can compensate for the loss of SNR, and thus bring additional gains [123].

Most of the recent works on SM have focused on generalized SM and ASM. Unlike SM, in which only one antenna is used in each time slot, two antennas are activated

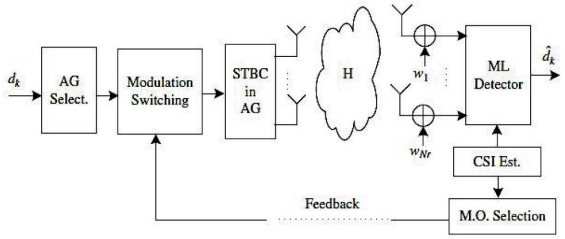


**Fig. 9.** ***Block diagram of ST-ASM transceiver.***

simultaneously in generalized SM to transmit information symbols, thus transmitting multiple data streams. The data rate of a generalized SM system is larger than that of a conventional SM assuming an equal number of available Tx antennas. In ASM, different levels of modulation are selected adaptively based on a low-complexity MO selection criterion for different channel conditions. Thus, we achieve better BER performance while using the same data rate target when using ASM than using conventional SM methods with fixed MO.

To be robust in smart railway scenarios, an ASM scheme was proposed in [124], in which the authors embedded an Alamouti STBC to the high-speed communication system. As shown in Fig. 9, in the ST-ASM system, two antennas are activated to transmit Alamouti code blocks, which are obtained through encoding information symbols. Compared with the conventional SM and ASM, the proposed ST-ASM scheme provides larger diversity gain and better BER performance when using equal data rate and equal transmission power [124]. Similar to the ASM scheme, the MOs assigned to each Tx antenna under the ST-ASM are chosen adaptively according to different channel conditions. Note that there is zero-order modulation present in the antenna group when selecting MOs and no information bits are modulated at the Tx antennas. As such, we consider the ML detection of the ST-ASM.

## B. Fast Time-Varying Channel Estimation in HSR Scenarios

HSR trains run at a speed of around 360 km/h, which leads to rapid changes in both amplitudes and phases of wireless channels. As a result, fast time-varying channel estimators are of vital importance to the wireless transmission systems of HSRs. Classical channel estimators (i.e., LMMSE and LS) may fail to provide satisfactory performance in such highly time-selective scenarios.

Traditionally, there are two widespread ways to estimate the time-varying channel parameters: through a BEM and through a GMM. The former decomposes the channel parameters into a superposition of the time-varying basis functions, which is weighted by time variant coefficients; and the latter tracks channel variation through symbol-to-symbol updating.

BEM expresses the time-varying channel $\mathbf{h} = [h_1, \ldots, h_N]^T$ as $\mathbf{h} = \mathbf{B}_g$, where $\mathbf{B} = [b_1, \ldots, b_Q]$ is the basis matrix, $\mathrm{b}_q (1 \le q \le Q)$ denotes the $q$th basis vector, and $\mathbf{g} = [g_1, g_2, \ldots, g_Q]^T$ contains the BEM coefficients to be estimated. Clearly, the number of the channel parameters to be estimated is decreased from $N$ to $Q$. Often choices of BEMs are complex-exponential BEM and polynomial BEM. However, these BEMs generally have unvaried basis matrices, which may not be optimal for various HSR scenarios such as cuttings, viaducts, and tunnels. Similarly, GMMs use the same model for all HSR scenarios.

It is worth noting that trains run on a fixed track, which is different from mobile users on the ground that can move toward any directions. Therefore, the communication systems of the current train face the same wireless environment with those of the past trains, which can be further exploited to aid channel estimation. Because of this, we propose applying a new channel estimator for HSR that exploits the historical information of past trains to enhance the estimation performance [80], [125]. This channel estimator is referred to as HiBEM.

The underlying principle of HiBEM is that trains follow fixed tracks and that the surrounding geological and wireless environments are almost the same at a fixed position given similar weather conditions. Consequently, the wireless channels of the current and past trains in certain areas are strongly correlated if the transceiver configuration is known and that train speeds are predictable. Based on these observations, HiBEM adopts the first $Q$ eigenvectors of the channel autocorrelation matrix as its basis matrix. The channel autocorrelation matrix is defined as $R_h = E\{hh^H\}$. The eigenvalue decomposition of $R_h$ will yield $R_h = UDU^H$.

Hence, we can construct basis matrix $B$ from $U$, which has its first $Q$ columns as the eigenvectors of $R_h$. The basis matrix $\mathbf{B}$ of HiBEM is chosen as $\mathbf{B} = U(:, 1:Q)$. It has been proven in [148, Sec. III-B] that basis matrix $\mathbf{B}$ is optimal in terms of minimizing the approximation MSE, that is, min $E\|h - B\eta\|^2$, where $E$ denotes average operation and $\eta$ is a coefficient vector.

One key problem is how to compute the channel correlation matrix $R_h$. The wireless channels on HSR are strongly related given a fixed position, which can be used to calculate the correlation matrix $R_h$. Fig. 10 depicts the scheme of calculating $R_h$. Suppose the current train is the $k$th train. We utilize the estimated channels of previous $(k-1)$ trains at the same position to compute the corresponding correlation matrices $R_{hi} (1 \le i \le k-1)$ and average them as

$$\mathbf{R} = E\{\mathbf{R}_{hi}\} = \frac{1}{k-1} \sum_{i=1}^{k-1} \mathbf{R}_{hi}. \tag{4}$$

We utilize $\widehat{R}_h$ to derive the real channel correlation matrix $R_h$.

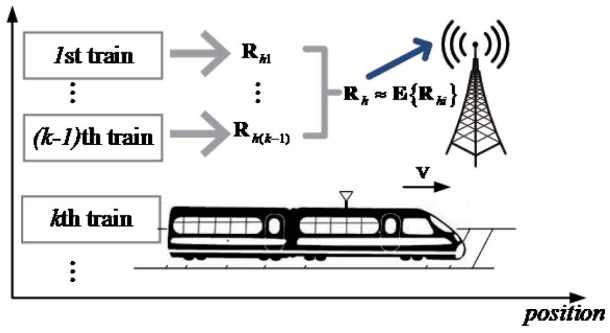


**Fig. 10.** *Scheme of estimation with historical information.*

Fig. 11 compares the estimation performance of various BEMs. We set the channel vector length as 20 and generate random channels through a Jakes model with a maximum Doppler shift of 300 Hz. Then we choose CE-BEM and HiBEM to obtain the channel estimates and their corresponding MSE. For comparison, the estimation MSE of the interpolation method is also plotted. It can be readily checked in Fig. 11 that HiBEM has the minimum approximation error and the error decreases with larger $Q$ value.

Fig. 12 depicts the estimation performance of HiBEM, LS, and LMMSE approaches in terms of their mean square errors; parameter $k$ indicates the number of past trains. We consider an OFDM system with 64 subcarriers, that is, $N = 64$. We set the number of pilots in the LS and LMMSE estimators to 16. The train moves at 360 km/h, which indicates that the train moves 0.1 m in 1 ms or during one LTE frame duration. The communication system is deployed at 3.5-GHz carrier frequency with 100-MHz bandwidth. The transmitted symbols are generated through BPSK modulation. The white Gaussian noise variance is set to 1. It can be shown from Fig. 12 that HiBEM outperforms both LS and LMMSE. It can also be seen that HiBEM, in the case of four trains, achieves almost the same estimation performance compared to the case when the channel correlation matrix is perfectly known at the Rx.

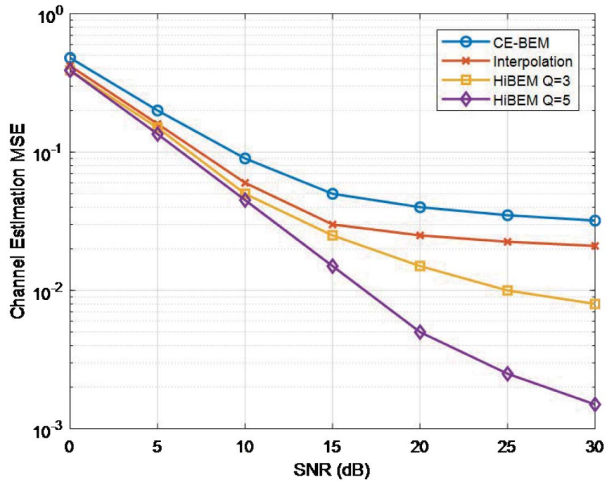


**Fig. 11.** *Performance comparison of various BEMs.*

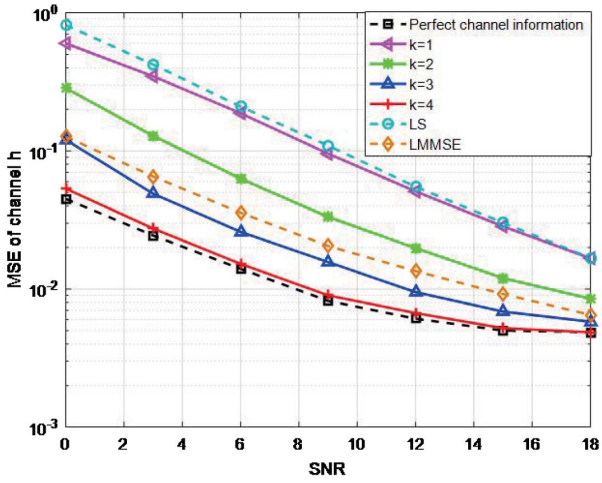


**Fig. 12.** *Estimation performance of HiBEM, LS, and LMMSE approaches.*

## C. Massive MIMO for HSR

When a train arrives in a slow-moving or stationary "in-station communication" scenario, we can upload the temporarily stored data to the train using large-scale mmWave arrays. Accordingly, we can obtain the traffic data, passenger information data, and multimedia data required for the next travel section. As for the hotspot railway areas (e.g., railway marshaling stations and hub areas), we can use other broadband technologies as a supplement. The users in these areas move slowly; thus, we can use high frequency and broadband modulation technologies to provide large-capacity wireless communication services.

Massive MIMO extends the multiuser MIMO concept by dramatically increasing the number of antennas employed at the BS to be able to serve more users simultaneously within the same time-frequency block. Note that massive MIMO does not significantly increase the peak rate for a single user; it inherently obtains its high SE by serving multiple users simultaneously. Fig. 13 illustrates the applications of mmWave massive MIMO under in-station and in-carriage scenarios.

With a large number of antenna units, the centralized massive MIMO BSs of railway systems can use BF to form high-gain and directional narrow-beam targets to track high-speed moving trains, as shown in Fig. 14. By using a high-gain antenna, the SNR and channel capacity improves. Multiuser spatial division multiple access can be achieved by using narrow beams, which will accordingly

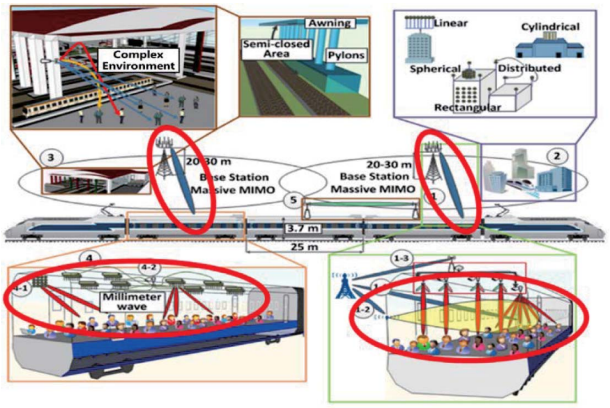


**Fig. 13.** *mmWave massive MIMO under in-station and in-carriage scenarios.*

increase the overall capacity of the BSs. Beam-tracking is utilized to achieve dynamic focusing of the energy to the intended Rx. The distributed massive MIMO configuration inside the carriages can provide high data rate services for multiusers.

The feasibility of using massive MIMO in several HSR scenarios has been investigated in [126]. The key challenges to its practical implementation include the requirement for fast signal processing in light of the high train mobility and the low EE of concentrated large-scale arrays along HSR lines. One of the promising solutions to tackle these key challenges is to utilize a distributed massive MIMO architecture called cell-free massive MIMO [127].

Unlike the conventional collocated massive MIMO, cell-free massive MIMO employs a large number of distributed APs with antennas over a wide area. First, the required net throughput and QoS are quite high inside a very large railway station since the station may contain thousands of passengers. The goal then is to provide the same high date rate to all passengers while using limited transmit power. Second, along the railway lines, cell-free massive MIMO can be deployed to guarantee good coverage without cells or cell edges. Remember that in conventional cell-based systems, the receive signal power drops to a minimum when the UE is near the cell edge, which may lead to low service quality at those locations. However, the QoS of cell-free massive MIMO is uniformly good for all areas. Finally, the frequent handover problem (e.g., hundreds of UEs need to be handed over to the next BS in a very short time) can be eliminated in a wide area.

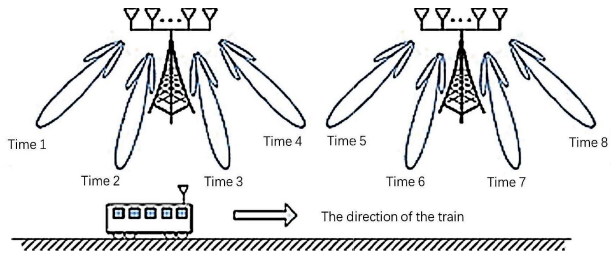


**Fig. 14.** *Block diagram of the beam tracking under railway scenarios.*

We use $q_k$ to denote the information symbol of the $k$th UE and $\rho_u$ to denote the maximum transmit power. The channel response $g_{mk}$ between AP $m$ and UE $k$ is assumed to be Rayleigh fading as $g_{mk} \sim \mathcal{CN}(0, \beta_{mk})$. With the assumption of perfect transceiver hardware, the received signal sequence at the $m$th AP is

$$y_{um} = \sum_{k=1}^{K} g_{mk} \sqrt{\rho_u \gamma_k} q_k + w_{um} \tag{5}$$

where $\gamma_k$ denotes the uplink power control coefficient and $w_{um}$ represents the additive noise. We introduce $\kappa_r$ and $\kappa_t$ to denote the level of hardware impairments in the Rx and transmitter, respectively. Then, the uplink SE of the $k$th UE can be given by

$$R_{uk} = \log_2\left(1 + \frac{\kappa_r \kappa_t A_u}{\kappa_r B_u + \kappa_r C_u - \kappa_r \kappa_t A_u + (1-\kappa_r) D_u + E_u}\right) \tag{6}$$

where

$$\begin{aligned}
A_u &\triangleq \gamma_k \left(\sum_{m=1}^{M} \lambda_{mk}\right)^2 \\
B_u &\triangleq \sum_{k'=1}^{K} \gamma_{k'} \left(\sum_{m=1}^{M} \lambda_{mk} \beta_{mk'} + \rho_p (1-\kappa_r) \sum_{m=1}^{M} c_{mk}^2 \beta_{mk'}^2\right) \\
C_u &\triangleq \sum_{k'=1}^{K} \gamma_{k'} \left(\left|\varphi_k^H \varphi_{k'}\right|^2 + \frac{1-\kappa_t}{\kappa_t \tau}\right) \left(\sum_{m=1}^{M} \lambda_{mk} \frac{\beta_{mk'}}{\beta_{mk}}\right)^2 \\
D_u &\triangleq \sum_{m=1}^{M} \left(\lambda_{mk} \sum_{k'=1}^{K} \gamma_{k'} \beta_{mk'} + c_{mk}^2 (1-\kappa_r) \rho_p \beta_{mk'}^2 \right. \\
&\qquad \left. + c_{mk}^2 \kappa_r \rho_p \beta_{mk'} \left(\tau \kappa_t \left|\varphi_k^H \varphi_{k'}\right|^2 + (1-\kappa_t)\right)\right) \\
E_u &\triangleq \frac{\sigma^2}{\rho_u} \sum_{m=1}^{M} \lambda_{mk}.
\end{aligned} \tag{7}$$

In the following, we numerically investigate the SE of cell-free massive MIMO systems with hardware impairments. We assume $M$ APs and $K$ UEs are independently and uniformly distributed in an area of size $1 \times 1$ km$^2$. We further assume that the number of pilot sequences is equal to the number of UEs, and all UEs are assigned with orthogonal pilots. Therefore, the pilot contamination can be alleviated. The detailed simulation parameters can be found in [128].

Fig. 15 shows the cumulative distribution of the downlink SE per UE for cell-free massive MIMO for the different numbers of APs (e.g., $M = 100, 200, 300$). We assume

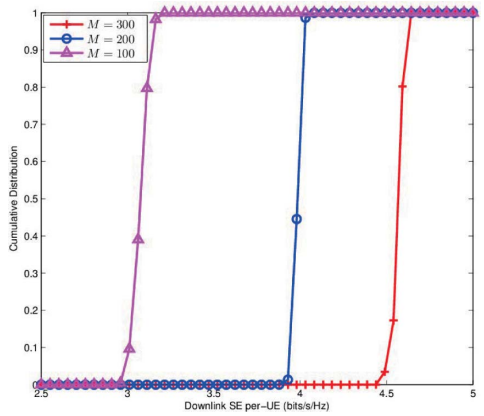


**Fig. 15.** *Cumulative distribution of the downlink SE per UE for cell-free massive MIMO against different number of APs.*

that the number of UEs is 10. Clearly, the SE performance improves significantly in terms of both median and 95%-likely per-user SE for a larger number of APs. It is clear that the power control algorithm proposed in [128] can significantly improve the SE performance of cell-free massive MIMO systems.

However, applying cell-free massive MIMO to HSR scenarios also poses the following major challenges.

1) The fronthaul requirements, in terms of both capacity and power, increase significantly. One should be careful in designing signal processing techniques that reduce data traffic between APs and the central processing unit. The tradeoff between the backhaul requirements and the system performance for cell-free massive MIMO is noteworthy, and thus needs to be further studied in the future.
2) The coherence times of the channel impulse responses in high-mobility scenarios are significantly small. The deep learning-based channel estimation schemes proposed in [129] and [130] can achieve remarkable performance. Thus, the current number of pilots is not enough to enable orthogonal pilots from all APs; yet reusing the same pilots may cause pilot contamination [131]. Therefore, the effect of pilot contamination on future 5G-R needs to be further investigated.

Cell-free massive MIMO and MCBF are two promising technologies for HSR scenarios. However, due to their characters, they can be employed in different conditions. For example, MCBF utilizes several proximate BSs to coordinately serve the UEs onboard high-speed train. There still exists frequent handover between different BS clusters. In contrast, cell-free massive MIMO eliminates handover by using efficient scheduling and high-capacity fronthaul. The complexity of cell-free massive MIMO is more pronounced than the one of MCBF, while achieving better performance. MCBF can be used for low-speed scenarios, such as rolling in and out of a railway station.

### D. mmWave Massive MIMO in Smart Rail Scenarios

mmWave communications employ massive MIMO mainly to achieve directional transmission based on antenna array BF. Accordingly, a major topic that researchers focus on is how to reduce the implementation complexity of mmWave railway communications without significantly degrading its performance. Based on the available velocity and location information that the communication-based train control system provides, we present a new simplified beam-tracking scheme that ensures high-capacity and low-outage probability. Unlike the beam-switching methods standardized in IEEE 802.11ad or the fixed selection of beamwidth as a function of location uncertainty, we can develop and optimize the sum-rate capacity maximization problem by jointly adjusting the beam direction and the beamwidth, and then introduce high flexibility in the beam patterns. Furthermore, we can apply a genetic algorithm to solve the decomposed problems and to obtain the near-optimal beamwidth in order to determine the beam direction [132]. In summary, the beam direction acts as the leader with a predefined azimuth angle, and we aim to maximize the sum-rate capacity. Thus, the joint beam direction and width allocation problem can be treated as a leader–follower sequential decision problem. The beam direction will not change, unless the transmit antenna gain cannot be improved and QoS requirements will not be satisfied by a single beam with width control. When the optimal beamwidth appears to be far from the QoS requirements, a dynamic beam direction adjustment is sought to cope with the transceiver angle, and the BS to MRN link turns the centerline of the main lobe. Thus, the implementation complexity is significantly reduced, while the sum rate remains close to optimal.

We consider a typical system model of BF operating at mmWave for railway wireless communication systems as depicted in Fig. 16. The figure shows that the trackside BS is equipped with a ULA array. A moving relay node is also mounted on top of the train to help enhance the cellular coverage and to avoid vehicle penetration loss [132]. Then, we present the performance evaluation results of our beam-tracking scheme. In the simulations, a typical mmWave small cell in the HSR scenario is considered. Special propagation properties, high mobility, and the Doppler effects are also considered.

For comparison, two typical BF strategies, real-time beam tracking and beam switching, are conducted. Specially, the real-time beam tracking strategy is configured with the beamwidth setting as $10^\circ$ and $25^\circ$ separately, wherein the beam direction is always aligned with the MRN when the QoS requirements cannot be satisfied. Besides, the beam switching scheme in IEEE 802.11ad is also considered for comparison, where $N_b = 3$ candidate beam patterns are predefined with equal angle interval and fixed beamwidth is set as $10^\circ$ and $25^\circ$. In the simu-

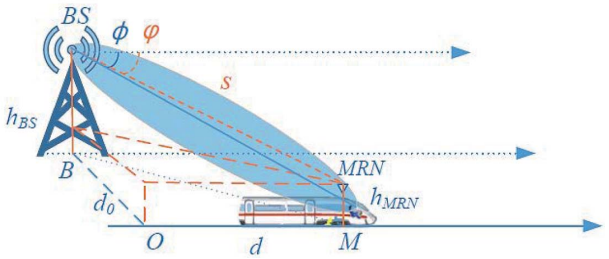


**Fig. 16.** ***mmWave massive MIMO beam-tracking systems for HSR communications.***

**Table 3** Simulation Parameters for Massive MIMO BF

| Parameter | Symbol | Value |
|---|---|---|
| System bandwidth | W | 2160 MHz |
| Background noise | $N_0$ | -174 dBm/Hz |
| Path loss exponent | $\alpha$ | 2 |
| Carrier frequency | $f_c$ | 60 GHz |
| Height of the BS | $h_{BS}$ | 4 m |
| Height of the MRN | $h_{MRN}$ | 2.5 m |
| Perpendicular distance | $d_0$ | 5 m |
| Radius | $R$ | 580 m |
| Threshold | $\gamma th$ | 10 dB |

lations, the proposed strategy is configured with a valid range of a beamwidth from 10° to 25°. The simulation results are depicted in the following four figures, which demonstrate the sum rate and outage of the proposed algorithm under various configurations, including transmission power, velocity of the train, and the train-to-BS distance. The simulation parameters are given in Table 3.

Fig. 17 plots the received SNR from different train positions under three BF schemes. The received SNR decreases as the train moves away from the BS. The rate of decrease is higher when the train is near the BS. When the train moves farther away from the BS, on the other hand, the SNR decreases more slowly. The SNR of beam switching varies significantly with train position since the relative position of the BS and the relay on the train decide the beam direction. The SNRs of the real-time beam tracking and our proposed BF scheme vary less since the beam direction adjusts according to the relative position of the BS and the relay.

Fig. 18 plots the sum rate of the proposed scheme under different train velocities and BS transmission powers.

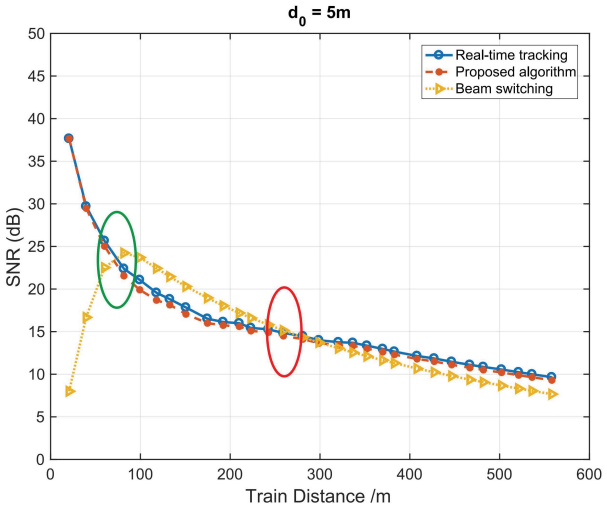


**Fig. 17.** ***Received SNR from different train positions.***

The sum rate is the total data rate between the BS and the relay in the serving cell. Fig. 18 shows that as the velocity of the trains increases, the Doppler effect consequently becomes more serious, and the sum rate decreases. When the transmission power increases, the received SNR increases; thus, the sum rate also increases.

Fig. 19 depicts the sum rate comparison of the three schemes versus train velocity with $P_T = 30$ dBm. We can observe that the sum rate drops quickly when the train speed increases, owing to the high ICI introduced by the more severe Doppler spread.

Fig. 20 displays the outage performance of these BF techniques versus train velocity for $P_T = 10$ dBm. The outage probability rises with the speed of the train, owing to aggravating ICI introduced by Doppler spread. The gap between the proposed algorithm and the real-time tracking is indistinguishable at the speed of 100 km/h, and rises to

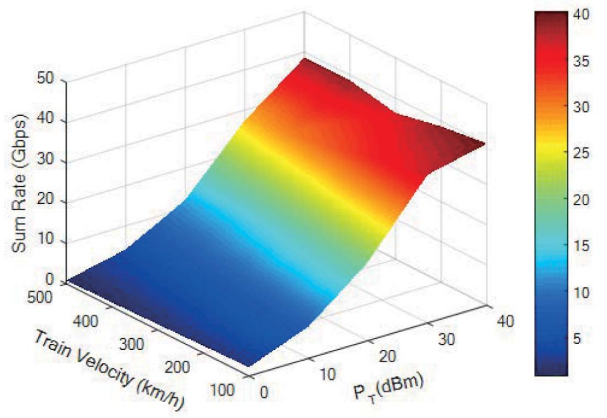


**Fig. 18.** ***Sum rate under different train velocities and BS transmission power.***

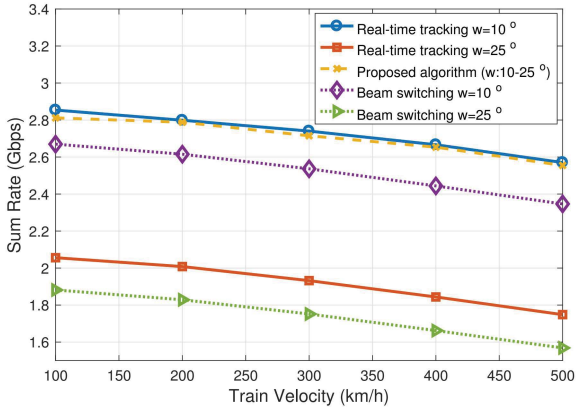


**Fig. 19.** ***Sum-rate comparison of BF strategies at different $v$.***

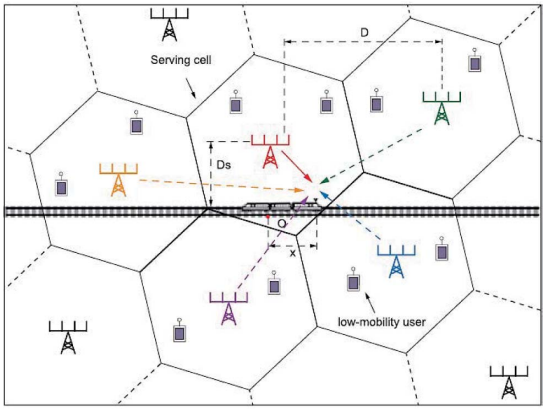


**Fig. 21.** ***HST goes through the mobile cellular system with low-mobility users.***

8.69% at a speed of 500 km/h. However, the counterpart gaps between the proposed one and the beam switching are 73.92% and 47.23%, respectively.

Considering varying packet loss due to the high speed of HSRs and possible mmWave obstacles, the performance of TCP-based applications is also important. It is very interesting to see how common user applications that use TCP might be affected with the current architecture for future smart railways. This will be done in our future work.

### E. Efficient BF for High-Speed Train in Low-Mobility User Existence Scenario

Hybrid BF is a technique in which a large number of antenna elements are connected, via an analog BF matrix in the RF domain, to a smaller number of RF chains that connect to the BBUs [133]. Thus, data transmission can be achieved by a combination of digital BF and analog BF. This technique can drastically reduce cost and energy consumption as compared to employing one RF chain for each antenna element, while resulting in only minor performance loss under most circumstances. It is useful at all frequency ranges but particularly important at mmWave frequencies [134].

We now design a robust and efficient communication scheme with mmWave hybrid BF. First, we establish a movement model for railways. Thereafter, we analyze the change in channel state when the train is moving, and then we design the channel-state monitoring mechanism, the beam tracking mechanism, and spatial diversity. We evaluate the performance of the proposed BF/tracking mechanism and compare it with the system performance without it.

The key issue for future mmWave studies is "How can we ensure a robust and efficient data transmission between BS and the train carriage mmWave links while considering the movement of the train carriage?" Specifically, this issue includes answering the following questions: How can we design an efficient hybrid BF to achieve multistream parallel and efficient transmission of hybrid BF technology? How can we improve the robustness of the connection using spatial diversity technology? How can we adapt the corresponding technologies for the hybrid BF system based on mobile features, such as analog BF, digital precoding, and spatial diversity combining?

Fig. 21 depicts a typical scenario in which a high-mobility train passes through a mobile cellular system that also serves some traditional low-mobility users outside the train. That is to say, the high-mobility train and low-mobility users coexist in a single system. To improve SE, we employ multiantenna BSs to serve the high-mobility train and low-mobility users over the same frequency band at the same time. Therefore, the interuser interference between the train (i.e., the big user) and low-mobility users and the intercell interference among different cells cannot be neglected in HSR communication systems [135].

MCBF is an effective way for mitigating both intercell and intracell interference since it is able to exploit the spatial degree of freedom to concentrate the beamformers to the users of interest. Therefore, MCBF may also be an efficient way of improving the information-transmit

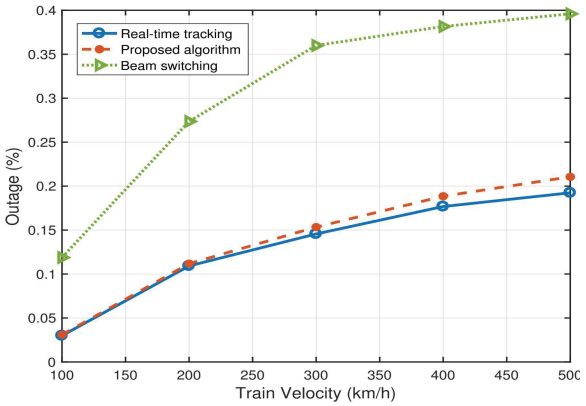


**Fig. 20.** ***Outage analysis of BF strategies at different train speed $v$.***

performance of railway communications with high mobility. Generally, existing MCBF designs are classified into two types: the power minimization-oriented design and the max–min fairness-oriented design. However, both of them may no longer be efficient in HSR communications. First, the primary goal of 5G-R is to increase the data rate for the train passengers rather than to save energy. Besides, BSs along the railway line often have stable and sufficient power supply. Thus, minimizing power is not the main concern of HSR communications. Second, to avoid group handover, all train passengers are aggregated to form a single virtual high-mobility user (i.e., the big user), which requires much higher data rates than low-mobility users have. Thus, in this situation, a max–min fairness between the big user and the low-mobility users is not meaningful. Therefore, existing MCBF designs are not suitable for HSR communications.

Consequently, we propose an improved MCBF design for HSR communication systems that is capable of providing the train with the highest information rate possible. The considered design is formulated into an optimization problem with the objective to maximize the achievable rate of the train under the constraints of minimal rate requirement of low-mobility users and the power budgets of all BSs

$$\begin{aligned}
&\max R_{s,0}\left(w_{s,0},\ldots,w_{N_C,K}\right)\\
&\text{s.t. } R_{s,K}\left(w_{s,0},\ldots,w_{N_C,K}\right)\geq r_{s,k},\quad k=1,\ldots,K\\
&\quad R_{n,K}\left(w_{s,0},\ldots,w_{N_C,K}\right)\geq r_{n,k},\quad k=1,\ldots,K\\
&\quad \sum_{l=0}^{K}\|w_{s,l}\|_2^2\leq P_T\\
&\quad \sum_{l=0}^{K}\|w_{m,l}\|_2^2\leq P_T,\quad n=1,\ldots,N_c
\end{aligned}\tag{8}$$

where $s$ and $n=1,\ldots,N_C$ are the index of the serving BS and the $n$th neighboring BS, respectively. In each cell, $k$ is used to indicate the $k$th user, where $k=0$ indicates the train and $k=1,\ldots,K$ indcates the low-mobility user. Here, $w_{x,y}$ denotes the BF vector from the $x$th BS to its $y$th user and $r_{x,y}$ denotes the minimal rate requirement of $y$th low-mobility user in the $x$th cell. Finally, $P_T$ denotes the power budget of each BS. For the high-speed train, the Doppler effect can be regarded as the source of ICI and the Doppler interference factor can be calculated by

$$P_{\mathrm{ICI}}=1-\int_{-1}^{1}(1-|\tau|)J_0\left(2\pi f_{\mathrm{d}}T_{\mathrm{s}}\tau\right)\mathrm{d}\tau \tag{9}$$

where $f_d$ is the maximum Doppler frequency, $T_s$ is the symbol duration, and $J_0(\bullet)$ is the zeroth-order Bessel function of the first kind. Note that the proposed design involves the fairness among low-mobility users. As the goal of the proposed design is to maximize the achievable rate of the train, the rate requirement constraints on the low-mobility users shall hold with equality when the optimal result is obtained. The reason is that the higher rate of the low-mobility users, the more interference goes to the train. By varying the minimal rate requirements of low-mobility users, a balance between fairness and performance can be achieved.

The considered problem is nonconvex, and an algorithm based on a bisection method is presented in [135] to globally solve it when the perfect CSI of the train and the low-mobility users is available at all BSs. The main idea of the solution approach is to use the bisection method to find the maximum data rate that the train can receive while all constraints can be satisfied.

In practice, due to the high mobility, the CSI of the train is hard to collect. Thus, the imperfect CSI of the train should be taken into consideration. To provide a high quality of wireless coverage, a broad BF approach is presented in [135] for the HSR communication systems. The algorithm presented in [135] can easily achieve a broad BF coverage for the train by taking the channel estimated error of the train into consideration. As the channel estimation error increases (which happens when the train moves with higher velocity), the BF coverage for the train becomes broader and the power consumption at the BSs increases. How to collect the CSI is also of high importance for transmit design, which is, however, beyond the scope of this article. The proposed design can be extended to the scenarios with other channel assumptions and the insights based on this article shall hold true.

The simulation scenario is given in Fig. 21. The number of neighboring cells is two, the number of transmit antennas is six, and the number of low-mobility users in each cell is two. The transmit power at each BS is set to 30 dBm. The required SINR at each low-mobility user is 5 dB. The system bandwidth is 10 MHz, and the noise power spectral density is $-162$ dBm/Hz. The inter-BS distance $D$ is 1000 m, and the distance between the BSs and the track is 300 m. For the train with 100 m/s, the relevant Doppler interference factor is estimated to be $-19.4$ dB. The locations of two BSs are set as (0, 300) and (1000, 300), respectively. We consider the case that the train moves from ($-500$, 0) to (500, 0). For a given location of the train, the distance between each BS and the train can be calculated and the channel between each BS and the train can be estimated. Then, the optimal transmit BF design can be achieved by utilizing the algorithm presented in [135]. Some metrics are considered to evaluate the performance of the proposed design, such as mobile service denoting sum rate achieved at the train between two locations and the average information rate during this period.

Fig. 22 shows the mobile service amount of the train and the low-mobility user. It is observed that the derivative of the service amount of the train first increases and then decreases with regard to the train location whereas that of the low-mobility user does not change. Moreover, increasing the transmit power $P_T$ improves only the service amount of the train. The reason is that our design intends

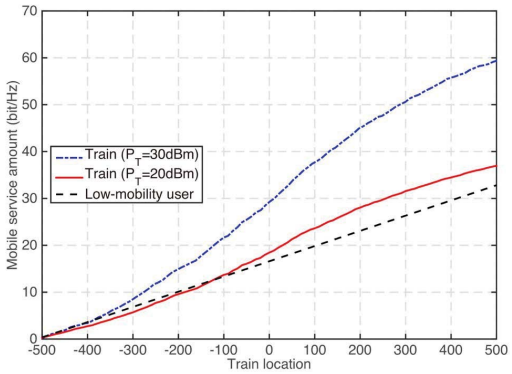


**Fig. 22.** *Mobile service amount versus the position of the train.*

to maximize the received rate of the train. Once the required minimal information rates of low-mobility users are satisfied, all remaining power of the serving BS should be allocated to increase the information rate of the train. Meanwhile, in order to mitigate the interference to the train, the remaining power of the neighboring BSs should not be used.

### F. MWB for Smart Rail Systems

In many cases, both rural and urban, an MWB is often the only technologically or economically feasible backhaul method. Even when an optical fiber is available near a cell site, MWB allows faster deployment until the fiber can be installed. On the other hand, it is expected that 5G will use spectrum above 6 GHz. Therefore, ideally, any spectrum used for 5G access should be flexible enough such that it can also be used for a backhaul. Likewise, spectrum licenses should be flexible enough to allow operators to meet the rollout demand while being capable of using 5G spectrum for backhaul whenever appropriate.

Higher frequency bands, especially millimeter and submillimeter wave bands, often show higher directionality, and thus can be used with more directional antennas, than sub-6-GHz channels. Hence, this makes it easier to use these bands simultaneously for mobile access and backhaul. Yet the specifics of the orthogonalization between these two types of transmissions still need to be further studied.

A related problem arises when the AP on the train aggregates the information from different users, and relays them to the BS. The latter part can be seen as a mobile backhaul from the AP to the BS; clearly in this case an MWB is the only option.

Fig. 23 shows an MIMO-BS only architecture and our proposed MWB architecture. The MIMO-BS is equipped with $N_t$ antennas and is located on the track side, with $N_u$ UEs located inside the carriage. In the MIMO-BS-only architecture, the MIMO-BS directly connects to the UEs in a multiuser MIMO mode using ZF-BF. In the MWB architecture, $N_s$ APs are deployed from above the carriage; the access link can be AP-UE and MIMO-BS-UE. The backhaul link between the AP and the MIMO-BS operates in a multiuser MIMO mode using ZF-BF. The MWB architecture uses a hierarchical joint user association and a backhaul bandwidth-allocation algorithm, which has an appropriate BF capacity model. We then compare the data rates of both architectures using the same parameter setups.

At the same time, the propagation between the $N_t$ BS antennas and the $N_s$ APs can be described through an $N_t \times N_s$ matrix, in which for simplicity the entries of the small-scale fading are assumed to be independent $\mathcal{CN}(0,1)$ random variables. We further assume that the MIMO-BS antenna size $N_t$ is much larger than $N_s$ to keep our problem within the regime of massive MIMO. For convenience, the link between the BS and UEs denotes BS access link, the link between APs and UEs denotes AP access link, and the link between BS and APs denotes backhaul link. We define the set of UEs by $U$ and the set of APs by $S$. Thus, $S0 = S \sqcup 0$ is the set of BS and APs, in which the index 0 is introduced for the BS. We assume a ZF algorithm to reduce the interferences caused by inter-UEs and inter-APs. To fully utilize the advantages of massive MIMO, the MIMO-BS concurrently serves a backhaul link and a BS access link in a downlink time slot. The maximum size of the precoding group that the MIMO-BS can serve is $N_g$ ($N_g \ll N_t$), with $N_g$ larger than $N_s$ such that the backhaul links for all the APs can be accommodated over the same frequency band, where $N_u$ is much greater than $N_g$. Resource sharing must be used when a large number of UEs are associated with the massive MIMO-BS. By allocating different orthogonal frequency-temporal resources to different precoding groups, which consist of maximum concurrent serving terminals, the system has no intergroup interference. After the APs establish the backhaul link, the data are delivered to the UEs from the AP access link. TDMA can be applied to this conventional point-to-point communication scenario.

We consider the mmWave bands 18, 28, and 38 GHz, each with 1-GHz bandwidth. The number of antenna elements at the BS $N_t$ ranges from 100 to 500. The total transmit power of the MIMO-BS is 40 dBm. The transmit power of the APs is 33 dBm. The train has six carriages with

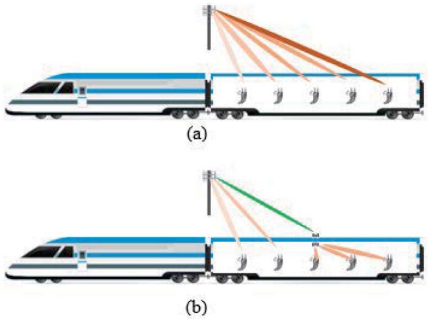


**Fig. 23.** *Considered architectures. (a) MIMO-BS only. (b) MWB.*

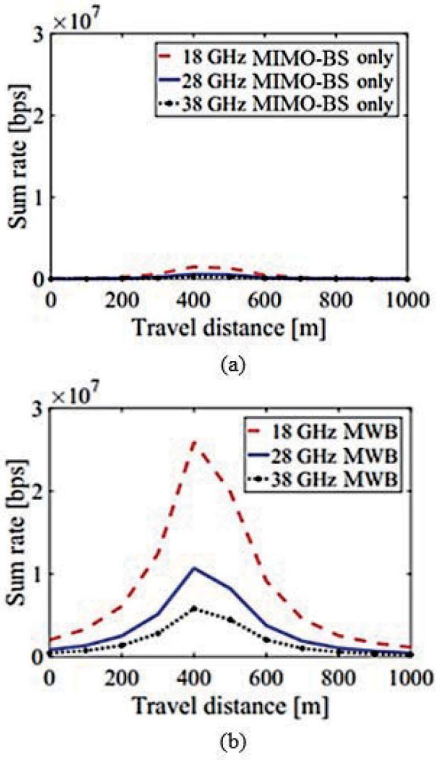


**Fig. 24.** *Influence of train position on sum rate of UEs. (a) The result of the MIMO-BS only architecture; (b) the result of using the MWB architecture.*

80 randomly distributed UEs per carriage. The penetration loss of the wagon is 15 dB at all the three frequencies. The length of each carriage is 20 m, and the total length of the train is 120 m. The train moves along the $x$-axis from 0 to 1000 m, and the MIMO-BS is located at $x = 400$ m, which is half the traveling distance. The 2-D horizontal distance $d_1$ of the MIMO-BS to the track is parallel with the $y$-axis, and $d_1$ ranges from 10 to 100 m. The heights of the MIMO-BS and APs are 20 and 4 m, respectively.

Fig. 24 shows the effect of the train position on the sum rate of UEs ($N_t = 500$, $d_1 = 10$). Fig. 24(a) shows the result of the MIMO-BS only architecture whereas Fig. 24(b) shows the result of using the MWB architecture. In both architectures, the maximum data rate is achieved when the center of the train reaches $x = 500$ m because the average distance from the APs to the MIMO-BS ($d_2$) is the smallest. The larger the $d_2$ is, the less the data rate would be. The data rate of the MWB architecture is higher than that of the MIMO-BS-only architecture. When the frequency increases, the path loss increases, and the data rate decreases.

## G. URLLC for Smart Rail Systems

The core performance metric of URLLC is that the packet error ratio is on the order of $10^{-5}$ under less than 1-ms E2E transmission delay, which is significant for the future development of smart rail communication systems [136]. As one of 5G key technologies, URLLC technology can offer super highly reliable and low E2E latency communication for HSR services, such as safety and control message delivery and surveillance. However, the classical Shannon capacity is accurate only when the codeword blocklength is infinitely long, and thus not applicable to the URLLC systems. Recently, the achievable rate with FBCs over the AWGN channel has been accurately approximated. With received SNR $\rho$, blocklength in units of symbols $m$, and packet error probability (PEP) $\varepsilon$, the transmission rate in bits per channel use is given by [137]

$$\frac{N}{m} = \log_2(1+\rho) - \sqrt{\frac{1}{m}\left(1-\frac{1}{(1+\rho)^2}\right)}\frac{Q^{-1}(\varepsilon)}{\ln 2} + \frac{\mathcal{O}(\log m)}{m} \tag{10}$$

where $N$ denotes the number of bits to be sent. This formula explicitly characterizes the tradeoff between the transmission rate, latency, and reliability. Meanwhile, the theoretical rate of (5) can be practically approached via polar codes with short blocklength. A more recent result given in [138] shows that the rate gap is only 0.025 dB by a simple CRC-polar concatenated scheme with the CRC-aided hybrid decoding method. Hence, the FBC-based formula has been applied to the study of various communication scenarios with strict latency constraints, as in [139].

In Fig. 25, we compare the predicted energy consumptions by invoking the FBC capacity and Shannon capacity of a downlink single-antenna system, in which the energy-efficient resource allocation for a two-user heterogeneous NOMA downlink with an FBC is investigated. The packets may not be aligned thus the successive interference

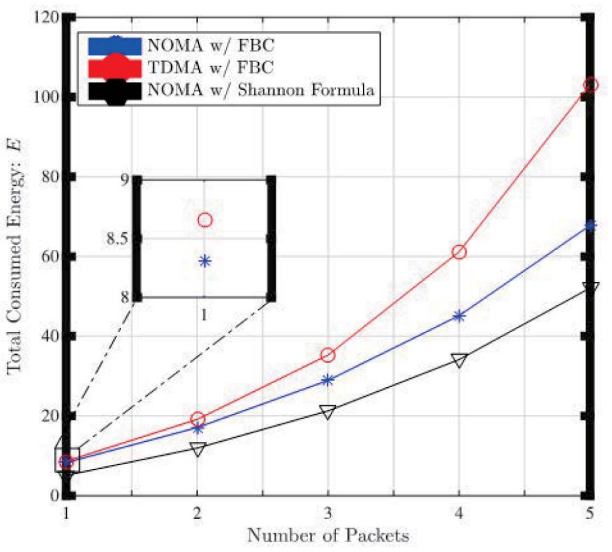


**Fig. 25.** *Energy consumption using FBC and Shannon capacities versus the number of packets, with $m_1 = m_2 = 640$, $P_{max} = 40$ dBm, and 32 bytes/packet.*

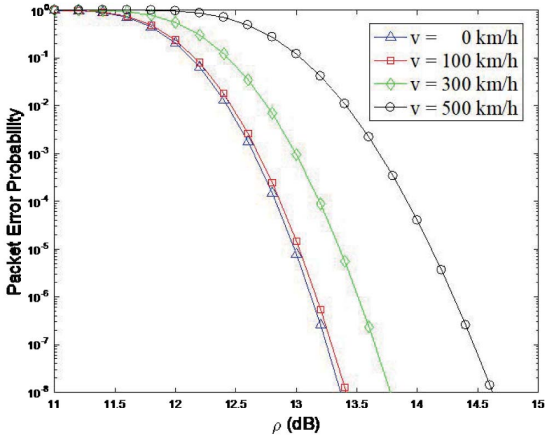


**Fig. 26.** *PEP $\varepsilon$ versus received SNR $\rho$ at different moving speeds, in which the ICI due to the Doppler spread is estimated by its upper bound, N = 128 bytes, m = 256 symbols, $f_c$ = 800 MHz, and $T_S$ = 160 $\mu$s.*

cancellation (SIC) cannot be always perfectly performed, which is a key challenge for NOMA transmission with finite blocklength coding. By fixing the requirements of latency ($m_1 = m_2 = 640$) and PEP $\varepsilon = 10^{-6}$, it shows that the Shannon capacity formula universally underestimates the energy of both the TDMA and NOMA schemes. We can also observe that the performance gains of NOMA, as compared with those of TDMA, increases as the number of transmitted packets increases due to the higher SE from the nonorthogonal superposition coding. This indicates that the performance caused by FBC should be carefully considered in URLLC system designs. Note that in the simulation, we assume QAM modulation with channel code 1/5, the system bandwidth is 1 MHz and the noise power density is set to be $\sigma_1^2 = \sigma_2^2 = -110$ dBm. The interested readers are referred to the details to [139]. Note that the PEP performance can be further improved by adaptive coding.

We can also observe that the performance gains of NOMA, as compared with those of TDMA, increase as the number of transmitted packets increases due to the higher SE from the nonorthogonal superposition coding. This indicates that the performance caused by FBC should be carefully considered in URLLC system designs.

Meanwhile, the FBC capacity degrades significantly under the imperfect CSI. Specifically, in high-speed scenarios, the CSI estimation is challenging even in a typical case where trains run over viaducts; thus, the channel is dominated by the LoS with a large coherence bandwidth. There are three main reasons for this: 1) fast-varying CSI due to high mobility; 2) the ICI of an OFDM system due to the Doppler spread; and 3) limited number of pilot symbols and feedback delay.

In Fig. 26, we compare the PEP performance versus received SNR under different movement speeds. To simplify, we assume that the received SNR $\rho$ is known whereas the ICI power is approximately characterized by its upper bound

$$P_{\text{ICI}} \leq \frac{1}{12}(2\pi f_d T_S)^2 \tag{11}$$

where $f_d = v f_c / C$ denotes the maximal Doppler frequency, $f_c = 800$ MHz, $v$ is the movement speed, $C = 3.0 \times 10^8$ m/s, and $T_S = 160$ $\mu$s. In the case with $N = 128$ bytes and $m = 256$ symbols, we find that the PEP is very sensitive to the received SNR; specifically, it increases as the movement speed increases. If the speed increases from 300 to 500 km/h, then it would require 1-dB SNR gain to achieve the same PEP of $10^{-6}$. This implies that both the channel estimation and the resource allocation are two key issues of the URLLC system design for high-speed applications.

## H. Power Adjustment Enhanced Handover for Smart Railways

Many existing works showed that increasing transmit power is capable of reducing the handover failure probability, but it is power hungry and goes against the indicator of green communications in 5G. Thus, adjusting the power in the time domain may be an efficient way, since it is able to reduce the "uncertainty" of the received signal strength in the handover procedure of HSR communications and does not require more transmit energy during the handover procedure [140], [141]. The basic idea is described as follows.

Consider an HSRC system, where a train is moving through an overlap region of two neighboring cells. The moving users within the train desire to connect with the ground BS via the TRS and some APs installed inside the carriages over a two-hop, that is, user-TRS-BS, communication architecture. When moving users intend to access the Internet via the BS, in the uplink, APs first collect the data from the users and then forward the aggregated data and requests to the BS with the help of TRS. In the downlink, the BS first transmits information to the TRS and then the TRS helps to forward the information to users via the APs. With such a two-hop architecture, as mentioned previously, the heavy overhead of the group handover is mitigated.

To perform the power adjustment, the overlap region is divided into two subregions, as shown in Fig. 27. The first one is from point A to point B, and the second one is from point B to point C, where A and C denote the starting and the ending points of the overlap region, respectively. B is the middle point. In the first subregion, the transmit power of the serving cell is increased and that of the target cell is decreased. In the second subregion, vice versa. To keep the energy consumption of each BS unchanged within the overlap region, the power is adjusted such that the increased amount of energy is equal to the decreased amount. Such a location-based time-domain power adjustment can be realized by estimating

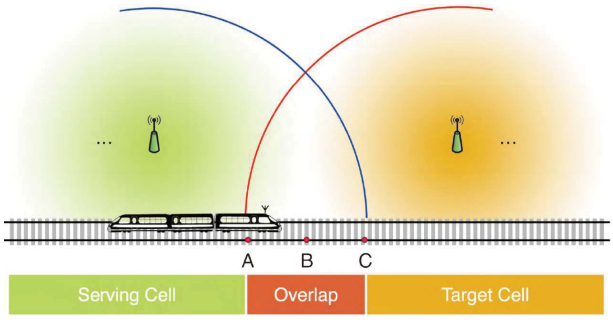


**Fig. 27.** ***Illustration of the subregions associated with handover.***

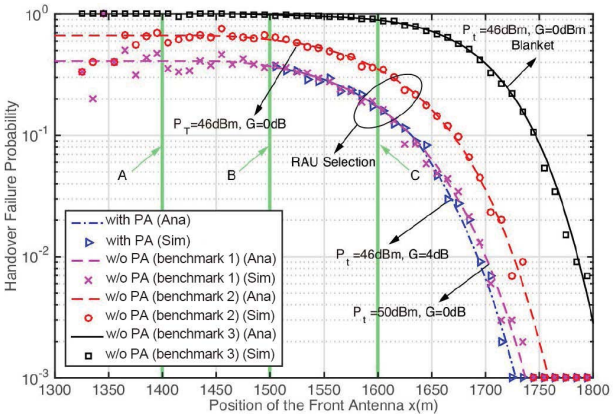


**Fig. 28.** ***Handover failure probability.***

**Table 4** Simulation Parameters

| Parameter | Value |
|---|---|
| Current transmit power of BS | 46dBm |
| Maximum transmit power of BS | 50dBm |
| Adjusted power | 4dB |
| Shadow fading deviation | 4dB |
| Path loss model | $31.5 + 35\log_{10}(d)$ |
| Distance between two cells | 3000m |
| Number of RAUs in one cell | 4 |
| Noise density | -145dBm/Hz |
| Signal threshold | -30dB |
| Predefined margin | 2dB |
| Distance from BS to railways | 100m |
| Distance from RAU to railways | 60m |
| Speed of train | 300m/s |

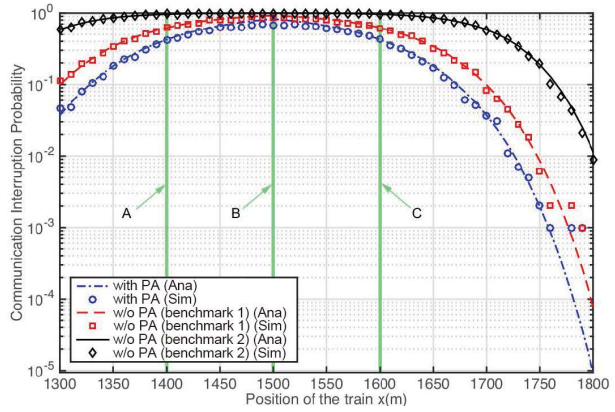


**Fig. 29.** ***Communication interruption probability.***

the channel information based on the position of the train.

To show the handover performance of the power adjustment enhanced method, we simulate and also compare it with other benchmarks with the following parameters as shown in Table 4.

In Fig. 28, both the analytical and the simulation results are provided, where the curves marked with "with PA" and "w/o PA benchmark x" represent the performance of the handover methods with and without power adjustment, respectively. Particularly, benchmark 1 is the handover scheme in traditional cells, which is with no RAU and PA. Benchmark 2 is the blanket transmission-based handover scheme in DAS cells without PA. Benchmark 3 is the RAU selection transmission-based handover scheme in DAS cells without PA. It is shown that by including power adjustment into handover, a much lower handover failure probability is achieved within the subregion from B to C [140]. With power adjustment, the communication interruption probability is greatly decreased in the handover procedure [140], which is shown in Fig. 29.

## VII. FUTURE WORKS

In this article, we have provided an overview of 5G-R emphasizing network architectures, channel models, and key technologies for future smart rails. However, there are still many challenges before we can successfully deploy 5G networks and services in the future railway scenarios. Such railway scenarios include D2D, integrated ground-air-space network, IoT, security, and AI, which are considered as key emerging technologies in 5G networks. In this part, we present some key new technology components and solutions to address the key challenges and requirements for a future smart railways.

### A. D2D for Smart Railways

When the wireless link between the BS and the train fails, operators could consider direct D2D communication as an emergency communication tool that is a well-integrated part of the 5G overall wireless-access solution [142]. Using direct D2D communications could be considered as a means of extending coverage beyond the reach of the conventional infrastructure (device-based relaying). D2D can also provide infrastructure fault-alerting and monitoring mechanism inside train carriages.

Applying D2D to railway scenarios is challenging for several reasons. First, D2D communications should also be possible in scenarios where network coverage is unavailable. As such, a D2D link needs to establish connectivity without- network control/assistance. Second, the D2D link is unstable because trains are highly mobile.

Cooperative D2D have high-speed communications that provides operators with the means for "joint" transmission and/or reception between multiple devices. Note that this can be seen as a kind of coordinated transmission and reception; however, this would be on the device side and not on the network side.

The key point is to see D2D communications as a well-integrated part of the overall wireless access solution. Accordingly, D2D communications should be considered right from the start of the 5G definition, instead of taking it as a later-introduced "add-on" option.

### B. Integrated Ground-Air-Space Network for Smart Railways

Using satellites, airships, unmanned aerial vehicles, and other air platforms in 5G-R scenarios has many advantages [143]. For example, it can provide efficient E2E transmission services for data, such as smart rail network-detection data, early warning information, and remote decision-making. However, achieving integrative information transmission and processing under air, space, train, and ground settings is challenging for 5G-R.

First, a high data-rate wireless connectivity should be built to support the tremendous amount of data that the integrated ground-air-space network will produce. Second, power-limited airships and unmanned aerial vehicles cannot be used under strong rain or wind conditions. To solve these problems, designers should provide an efficient data distribution and resource-sharing scheme under a unified control.

### C. IoT for Smart Railways

In addition to real-time query and tracking the whole trajectory of the train and goods location, the IoT for railways can be developed in 5G-R networks to integrate sensing information of rail infrastructures, including bridges, viaducts, tunnels, leaky feeders, rail gaps, frozen soil, and slope protection [144]. This could be done by installing various sensing measures such as infrared sensors, sound sensors, and temperature sensors.

In recent years, IoT technology has gradually attracted the attention of railway departments, industries, and research institutes all over the world. Developing railway IoT and building railway safety information-guaranteed systems based on IoT are important directions for the in-depth integration of railway information and industrialization.

However, researchers should further explore the mobile communication problems in IoT, and then conduct research on multiple access from a massive number of devices in the presence of high mobility. Likewise, the mechanisms for evaluating and optimizing railway wireless network resource management mechanism should also be further studied. Designers also need to develop and construct a closed-loop management of safety monitoring and control system for railway equipment and facilities, collect real-time information of mobile equipment and fixed facilities, analyze the disposition, and ensure safe operation. Also, monitoring systems for natural disasters (e.g., wind, rain, snow, and earthquakes) and foreign matter intrusion need to be established to monitor the safety of railway operations in real time.

### D. Security for Smart Railways

The operator of the 5G-R system should be able to securely collect information that can enhance user and service experiences via data analytics. Security has been one of the fundamental capabilities that operators should provide to their customers, especially to train operation control systems [145]. 5G-R will support a wide range of applications, from human-based to machine-based communications. Thus, the corresponding technology should be able to deal with huge amounts of sensitive data that need to be protected against unauthorized access, use, disruption, modification, inspection, attack, etc. The importance of providing a comprehensive set of features that guarantee high-level security is a core requirement for 5G-R systems. Therefore, 5G-R should be designed to provide more options beyond node-to-node and E2E security that is available in today's mobile systems. Accordingly, such a design will protect the users' data and prevent or mitigate any possible cyber security attack.

On the other hand, operators should protect railway communications and signaling against electromagnetic attacks, which can identify scenarios and devices that are vulnerable to attacks. Likewise, researchers and developers should conduct a risk analysis of attack scenarios, propose responses, develop solution to detect electromagnetic attacks, and design the right architecture to be resilient to such kind of attack.

### E. AI for Smart Railways

Recently, AI has become popular in the computer field due to its great success in computer vision, natural language processing, automatic speech recognition, and wireless communications [146]. Future smart rail considers

AI as a significant direction in 5G-R networks; AI would enable networks to process a large amount of data, dynamically recognize and adapt to complex scenarios, and satisfy the requirements for high-speed and real-time signal processing capabilities.

Note that most AI methods existing today are data driven. They use the standard neural network structure as a black box and then train it through large amounts of data. By integrating AI into the physical and upper layers, smart rail communication systems would be able to automatically adapt to the properties of signals and propagation scenarios, and thereby obtain efficient deployment. However, the current design of AI schemes for communication networks is simple. To improve the network performance, we should resort to a specialized AI architecture for 5G-R communications that considers the special characteristics of railway scenarios and strict indicators. Above all, with the development of advanced communication technologies, future railways will become smarter and smarter.

### F. Mobile Edge Computing (MEC) for Smart Railways

Under HSR scenarios, the high-speed movement of the train faces the dramatically time-variance of the channel and the rapid reduction of the available bandwidth. The new applications for smart rail, such as automatic train driving, real time management for high-speed rail networks, and real time HD video surveillance, need stronger computing capabilities and lower processing latencies.

MEC is considered as an emerging key technique in 5G communication which extends the ability of cloud computing to the edge of the network and reduces the bandwidth requirement between the train and the BS [147]. MEC for smart rail consists of onboard MEC and trackside MEC. Onboard MEC stores and processes data that are incoming with a high rate (HD video surveillance data, train sensor data) generated during train running. The processed results, which require a small data rate, are fed back and shared with trackside MEC. Trackside MEC analyzes and processes the data of wayside sensors, receives data from onboard MEC, and shares the processed data with the cloud computing node. With the deployment of MEC, the process of smart rail data will become real-time and efficient.

### G. Sixth-Generation (6G) for Smart Railways

As 5G is deployed around the world, 6G becomes a hot topic and has attracted increasing attention from researchers in both academia and industry [148]. The development of 6G will pave the way for diverse QoS requirements, real-time tactile interactions, customized and open services, integration of communication, broadcasting [149]–[151], computing, sensing, control, security, and AI functionalities. The main innovation in 6G network architecture is also required with respect to current 5G network design, such as space-aerial-terrestrial-sea integrated network, full-spectrum and full-dimensional coverage, intelligent self-sensing, learning, optimization, and evolution.

On the other hand, the potential applications for future smart railway networks include autonomous train driving, cooperative train networks, Internet of trains, UHD (4K/8K) train video, train *ad hoc* networks, and ultraaccurate (cm level) train localization. In order to fulfill the requirements of 6G smart railway applications, significant technological innovations are expected.

1) *Cell-Free network architecture:* The high mobility of trains induces frequent handover and huge overhead in conventional cellular networks. In order to guarantee a seamless and high-QoS coverage, it is promising to employ cell-free networks. Moreover, different heterogeneous sub-6G, mmWave, terahertz, and visible light communications enable cell-free networks to provide unprecedented performance. The application of AI techniques in 6G smart railway networks is promising in the context of self-learning, optimization, and evolution.
2) *Novel URLLC techniques:* To guarantee the autonomous train driving over 1000 km/h in 6G (e.g., emerging Hyperloop systems), it is important to further investigate the novel URLLC frame structure under FBC to achieve the tradeoff between ultrareliability and ultralow latency.
3) *Digital twin networks:* The monitoring, prediction and decision-making for the behavior of train drivers (e.g., nervous, drunk, sleepy, and excited) is significant for the security of HSTs. A digital twin network based on 6G can fundamentally evolve the digital profile of the historical and current behavior of the train driver that helps optimize their performance.

Finally, we believe the networking, broadcasting, communications, interaction, and security of future smart railways will be vastly improved by applying the aforementioned techniques, which deserve our effort in the next decade.

## VIII. CONCLUSION

Wireless communication is playing an important role in the future smart railway domain. This article explored a potential solution by leveraging emerging 5G technologies to provide a plethora of services in HSRs, both control and data services. More specifically, we first briefly described the current trend of wireless communications for smart railways. Services and requirements of future smart railways were presented in the following. Moreover, we introduced state of the art, drawbacks, and challenges of existing wireless technologies in supporting the envisioned services, respectively. We proposed a new 5G-R network architecture for smart railways. Channel models, physical layer design challenges and a thorough review

of 5G technologies were investigated in detail, respectively. Finally, we pinpointed some open research issues and promising future directions, which may lead to valuable results. We hope that this article provides a guideline for researchers and practitioners interested in applying novel 5G techniques to significant problems in HSRs. ■

## Acknowledgment

The authors would like to thank Prof. Jiayi Zhang, Chao Shen, Fanggang Wang, Yong Niu, Gongpu Wang, Ke Guan, Danping He, Ruisi He, Minming Ni, Wei Chen, Meilin Gao, Guoyu Ma, Longhe Wang, and Ke Xiong for their simulations and constructive suggestions.

## REFERENCES

[1] European Commission. (2014). *SHIFT2RAILInnovation Program 2*. [Online]. Available: https://shift2rail.org/research-development/ip2

[2] J. Moreno, J. M. Riera, L. D. Haro, and C. Rodriguez, "A survey on future railway radio communications services: Challenges and opportunities," *IEEE Commun. Mag.*, vol. 53, no. 10, pp. 62–68, Oct. 2015.

[3] R. He *et al.*, "High-speed railway communications: From GSM-R to LTE-R," *IEEE Veh. Technol. Mag.*, vol. 11, no. 3, pp. 49–58, Sep. 2016.

[4] Ericsson Mobility Report. (2019). *The Power of 5G Is Here and Will Continue to Spread Across the Globe in the Coming Years*. [Online]. Available: https://www.ericsson.com/en/mobility-report/reports/june-2019

[5] A. Osseiran *et al.*, "Scenarios for 5G mobile and wireless communications: The vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26–35, May 2014.

[6] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[7] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[8] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The Next Generation Wireless Access Technology*. New York, NY, USA: Academic, 2018.

[9] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[10] J. G. Andrews, T. Bai, M. Kulkarni, A. Alkhateeb, A. Gupta, and R. W. Heath, Jr., "Modeling and analyzing millimeter wave cellular systems," *IEEE Trans. Commun.*, vol. 65, no. 1, pp. 403–430, Jan. 2016.

[11] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 173–196, 1st Quart., 2019.

[12] M. Shafi *et al.*, "Microwave vs. millimeter-wave propagation channels: Key differences and impact on 5G cellular systems," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 14–20, Dec. 2018.

[13] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge, U.K.: Cambridge Univ. Press, 2016.

[14] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, "Prospective multiple antenna technologies for beyond 5G," 2019, *arXiv:1910.00092*. [Online]. Available: http://arxiv.org/abs/1910.00092

[15] A. Kaloxylos, "A survey and an analysis of network slicing in 5G networks," *IEEE Commun. Standards Mag.*, vol. 2, no. 1, pp. 60–65, Mar. 2018.

[16] B. Holfeld *et al.*, "Wireless communication for factory automation: An opportunity for LTE and 5G systems," *IEEE Commun. Mag.*, vol. 54, no. 6, pp. 36–43, Jun. 2016.

[17] X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, "Network slicing in 5G: Survey and challenges," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 94–100, May 2017.

[18] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE 77th Veh. Technol. Conf. (VTC Spring)*, Jun. 2013, pp. 1–5.

[19] M. Vaezi, Z. Ding, and H. V. Poor, *Multiple Access Techniques for 5G Wireless Networks and Beyond*. Berlin, Germany: Springer, 2019.

[20] Y. Liang, X. Li, J. Zhang, and Z. Ding, "Non-orthogonal random access for 5G networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4817–4831, Jul. 2017.

[21] H2020 METIS-II Project, D1.1. (Jan. 2016). *Refined Scenarios and Requirements Consolidated Use Cases and Qualitative Techno-Economic Feasibility Assessment*. [Online]. Available: http://metis-ii.5g-ppp.eu/

[22] NGMN Alliance. (Feb. 2015). *NGMN 5G White Paper*. [Online]. Available: https://www.ngmn.org/5g-white-paper/5g-white-paper.html

[23] B. Ai *et al.*, "Challenges toward wireless communications for high-speed railway," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2143–2158, Oct. 2014.

[24] *Methodology for Calculation of Spectrum Requirements for the Terrestrial Component of International Mobile Telecommunications*, document Rec. ITU-R M.1768, Apr. 2013.

[25] *Current and Future Usage of Railway Radiocommunication Systems Between Train and Trackside (RSTT)*, document Rec. ITU-R R15-WP5A, Jun. 2017. [Online]. Available: https://www.itu.int/md/R15-WP5A-C-0920/en

[26] C. Harrison *et al.*, "Foundations for smarter cities," *IBM J. Res. Develop.*, vol. 54, no. 4, pp. 1–16, Jul. 2010.

[27] B. Ning, T. Tang, Z. Gao, F. Yan, F.-Y. Wang, and D. Zeng, "Intelligent railway systems in China," *IEEE Intell. Syst.*, vol. 21, no. 5, pp. 80–83, Sep. 2006.

[28] M. Niculescu, A. Golgojan, A. Bednarz, G. Ivanova, and T. Maly, "Smart rail infrastructure, maintenance, and life cycle costs," *Trans. Prob.*, vol. 9, pp. 109–122, Sep. 2014.

[29] China Mobile Research Institute. (Jun. 2014). *C-RAN White Paper: The Road Towards Green Ran*. [Online]. Available: http://labs.chinamobile.com/cran

[30] J. Wang, H. Zhu, and N. J. Gomes, "Distributed antenna systems for mobile communications in high speed trains," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 4, pp. 675–683, May 2012.

[31] M. K. Müller, M. Taranetz, and M. Rupp, "Providing current and future cellular services to high speed trains," *IEEE Commun. Mag.*, vol. 53, no. 10, pp. 96–101, Oct. 2015.

[32] M. Peng, S. Yan, K. Zhang, and C. Wang, "Fog-computing-based radio access networks: Issues and challenges," *IEEE Netw.*, vol. 30, no. 4, pp. 46–53, Jul. 2016.

[33] A. Gotsis, S. Stefanatos, and A. Alexiou, "UltraDense networks: The new wireless frontier for enabling 5G access," *IEEE Veh. Technol. Mag.*, vol. 11, no. 2, pp. 71–78, Jun. 2016.

[34] C. Pan, M. Elkashlan, J. Wang, J. Yuan, and L. Hanzo, "User-centric C-RAN architecture for ultra-dense 5G networks: Challenges and methodologies," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 14–20, Jun. 2018.

[35] Y. Niu, C. Gao, Y. Li, L. Su, D. Jin, and A. V. Vasilakos, "Exploiting device-to-device communications in joint scheduling of access and backhaul for mmWave small cells," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 10, pp. 2052–2069, Oct. 2015.

[36] M. Gao *et al.*, "On hybrid beamforming of mmWave MU-MIMO system for high-speed railways," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[37] L. Yan, X. Fang, and Y. Fang, "Stable beamforming with low overhead for C/U-plane decoupled HSR wireless networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6075–6086, Jul. 2018.

[38] K. Guan *et al.*, "Towards realistic high-speed train channels at 5G millimeter-wave band—Part I: Paradigm, significance analysis, and scenario reconstruction," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9112–9128, Oct. 2018.

[39] Y. Niu *et al.*, "Relay-assisted and QoS aware scheduling to overcome blockage in mmWave backhaul networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1733–1744, Feb. 2019.

[40] Z. Genc, U. H. Rizvi, E. Onur, and I. Niemegeers, "Robust 60 GHz indoor connectivity: Is it possible with reflections?" in *Proc. IEEE 71st Veh. Technol. Conf.*, May 2010, pp. 1–5.

[41] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116753–116773, 2019.

[42] L. Yan, X. Fang, and Y. Fang, "Control and data signaling decoupled architecture for railway wireless networks," *IEEE Wireless Commun.*, vol. 22, no. 1, pp. 103–111, Feb. 2015.

[43] H. Song, X. Fang, and L. Yan, "Handover scheme for 5G C/U plane split heterogeneous network in high-speed railway," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4633–4646, Nov. 2014.

[44] Z. Liu and P. Fan, "Optimal selection of cooperative RAUs in HSR based distributed antenna systems and mobile relay with massive antennas," in *Proc. Int. Workshop High Mobility Wireless Commun. (HMWC)*, Oct. 2015, pp. 76–80.

[45] NGMN Alliance. (Jan. 2016). *Description of Network Slicing Concept by NGMN Alliance*. [Online]. Available: https://www.ngmn.org/fileadmin/user_upload/161010_NGMN_Network_Slicing_framework_v1.0.8.pdf

[46] IMT-2020 (5G) Promotion Group. (May 2016). *5G Network Architecture Design, White Paper*. [Online]. Available: https://www.imt-2020.cn/en/documents/download/27

[47] 5G Americas. (Nov. 2016). *Network Slicing for 5G Networks and Services*. [Online]. Available: http://www.5gamericas.org/files/3214/7975/0104/5G_Americas_Network_Slicing_11.21_Final.pdf

[48] H. Song, X. Fang, and Y. Fang, "Millimeter-wave network architectures for future high-speed railway communications: Challenges and solutions," *IEEE Wireless Commun.*, vol. 23, no. 6, pp. 114–122, Dec. 2016.

[49] L. Yan and X. Fang, "Reliability evaluation of 5G C/U-plane decoupled architecture for high-speed railway," *EURASIP J. Wireless Commun. Netw.*, vol. 2014, no. 1, p. 127, Dec. 2014.

[50] H. Zhang, P. Dong, W. Quan, and B. Hu, "Promoting efficient communications for high-speed railway using smart collaborative networking," *IEEE Wireless Commun.*, vol. 22, no. 6, pp. 92–97, Dec. 2015.

[51] C.-X. Wang, A. Ghazal, B. Ai, Y. Liu, and P. Fan, "Channel measurements and models for

high-speed train communication systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 974–987, 2nd Quart., 2016.

[52] T. Zhou, H. Li, Y. Wang, L. Liu, and C. Tao, "Channel modeling for future high-speed railway communication systems: A survey," *IEEE Access*, vol. 7, pp. 52818–52826, 2019.

[53] L. Liu *et al.*, "Position-based modeling for wireless channel on high-speed railway under a viaduct at 2.35 GHz," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 4, pp. 834–845, May 2012.

[54] R. He, B. Ai, Z. Zhong, A. F. Molisch, R. Chen, and Y. Yang, "A measurement-based stochastic model for high-speed railway channels," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1120–1135, Jun. 2015.

[55] A. Ghazal, C.-X. Wang, B. Ai, D. Yuan, and H. Haas, "A nonstationary wideband MIMO channel model for high-mobility intelligent transportation systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 885–897, Apr. 2015.

[56] S. Lin *et al.*, "Finite-state Markov modeling for high-speed railway fading channels," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 954–957, Dec. 2015.

[57] X. Yin, X. Cai, X. Cheng, J. Chen, and M. Tian, "Empirical geometry-based random-cluster model for high-speed-train channels in UMTS networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2850–2861, Oct. 2015.

[58] L. Zhou, F. Luan, S. Zhou, A. F. Molisch, and F. Tufvesson, "Geometry-based stochastic channel model for high-speed railway communications," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4353–4366, May 2019.

[59] G. Li *et al.*, "Channel characterization for mobile hotspot network in subway tunnels at 30 GHz band," in *Proc. IEEE 83rd Veh. Technol. Conf. (VTC Spring)*, May 2016, pp. 1–5.

[60] G. Li *et al.*, "Path loss modeling and fading analysis for channels with various antenna setups in tunnels at 30 GHz band," in *Proc. 10th Eur. Conf. Antennas Propag. (EuCAP)*, Apr. 2016, pp. 1–5.

[61] A. Hrovat, T. Javornik, and K. Guan, "Analysis of radio wave propagation at millimeter-wave band in tunnels for 5G communications," in *Proc. 22nd Int. Conf. Appl. Electromagn. Commun. (ICECOM)*, Sep. 2016, pp. 1–5.

[62] Y. Chang, M. Furukawa, H. Suzuki, and K. Fukawa, "Propagation analysis with ray tracing method for high speed trains environment at 60 GHz," in *Proc. IEEE 81st Veh. Technol. Conf. (VTC Spring)*, vol. 5, May 2015, pp. 1–5.

[63] S. Wu, C.-X. Wang, E.-H.-M. Aggoune, M. M. Alwakeel, and Y. He, "A non-stationary 3-D wideband twin-cluster model for 5G massive MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1207–1218, Jun. 2014.

[64] B. Ai *et al.*, "On indoor millimeter wave massive MIMO channels: Measurement and simulation," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1678–1690, Jul. 2017.

[65] V. Kristem *et al.*, "3D MIMO outdoor-to-indoor propagation channel measurement," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4600–4613, Jul. 2017.

[66] J. Zhang, L. Dai, Z. He, S. Jin, and X. Li, "Performance analysis of mixed-ADC massive MIMO systems over Rician fading channels," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1327–1338, Jun. 2017.

[67] J. Zhang, L. Dai, X. Li, Y. Liu, and L. Hanzo, "On low-resolution ADCs in practical 5G millimeter-wave massive MIMO systems," *IEEE Commun. Mag.*, vol. 56, no. 7, pp. 205–211, Jul. 2018.

[68] J. Zhang, L. Dai, X. Zhang, E. Bjornson, and Z. Wang, "Achievable rate of Rician large-scale MIMO channels with transceiver hardware impairments," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8800–8806, Oct. 2016.

[69] J. Zhang, L. Dai, S. Sun, and Z. Wang, "On the spectral efficiency of massive MIMO systems with low-resolution ADCs," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 842–845, May 2016.

[70] J. Zhang, X. Xue, E. Bjornson, B. Ai, and S. Jin, "Spectral efficiency of multipair massive MIMO two-way relaying with hardware impairments," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 14–17, Feb. 2018.

[71] K. Guan *et al.*, "Scenario modules and ray-tracing simulations of millimeter wave and terahertz channels for smart rail mobility," in *Proc. 11th Eur. Conf. Antennas Propag. (EUCAP)*, Mar. 2017, pp. 113–117.

[72] G. Noh, J. Kim, H. Chung, and I. Kim, "Realizing multi-Gbps vehicular communication: Design, implementation, and validation," *IEEE Access*, vol. 7, pp. 19435–19446, 2019.

[73] F. Hasegawa *et al.*, "High-speed train communications standardization in 3GPP 5G NR," *IEEE Commun. Standards Mag.*, vol. 2, no. 1, pp. 44–52, Mar. 2018.

[74] Y. Cui and X. Fang, "A massive MIMO-based adaptive multi-stream beamforming scheme for high-speed railway," *EURASIP J. Wireless Commun. Netw.*, vol. 2015, no. 1, p. 259, Dec. 2015.

[75] J. Zhang, L. Dai, Z. He, B. Ai, and O. A. Dobre, "Mixed-ADC/DAC multipair massive MIMO relaying systems: Performance analysis and power optimization," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 140–153, Jan. 2019.

[76] Y. Cui and X. Fang, "Performance analysis of massive spatial modulation MIMO in high-speed railway," *IEEE Trans. Veh. Technol.*, vol. 65, no. 11, pp. 8925–8932, Nov. 2016.

[77] Y. Cui, X. Fang, and L. Yan, "Hybrid spatial modulation beamforming for mmWave railway communication systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9597–9606, Dec. 2016.

[78] X. Chen and P. Fan, "Low-complexity location-aware multi-user massive MIMO beamforming for high speed train communications," in *Proc. IEEE 85th Veh. Technol. Conf. (VTC Spring)*, Jun. 2017, pp. 1–6.

[79] L. He, X. Fang, H. Li, C. Li, and Y. Liu, "An mmWave beamforming scheme for disaster detection in high speed railway," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Jul. 2016, pp. 1–6.

[80] G. Wang, Q. Liu, R. He, F. Gao, and C. Tellambura, "Acquisition of channel state information in heterogeneous cloud radio access networks: Challenges and research directions," *IEEE Wireless Commun.*, vol. 22, no. 3, pp. 100–107, Jun. 2015.

[81] R. Hadani *et al.*, "Orthogonal time frequency space modulation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017, pp. 1–6.

[82] J.-Y. Zhang, Z.-H. Tan, Z.-D. Zhong, and Y. Kong, "A multi-mode multi-band and Multi-system-based access architecture for high-speed railways," in *Proc. IEEE 72nd Veh. Technol. Conf. (Fall)*, Sep. 2010, pp. 1–5.

[83] M. Cheng, S. Yang, and X. Fang, "Adaptive antenna-activation based beamforming for large-scale MIMO communication systems of high speed railway," *China Commun.*, vol. 13, no. 9, pp. 12–23, Sep. 2016.

[84] D. Fan, Z. Zhong, G. Wang, and F. Gao, "Doppler shift estimation for high-speed railway wireless communication systems with large-scale linear antennas," in *Proc. Int. Workshop High Mobility Wireless Commun. (HMWC)*, Oct. 2015, pp. 96–100.

[85] J. Rodriguez-Pineiro, T. Dominguez-Bolano, J. A. Garcia-Naya, and L. Castedo, "Performance assessment of 5G-candidate waveforms in high speed scenarios," in *Proc. IEEE 27th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Sep. 2016, pp. 1–6.

[86] J. You, Z. Zhong, Z. Dou, J. Dang, and G. Wang, "Wireless relay communication on high speed railway: Full duplex or half duplex?" *China Commun.*, vol. 13, no. 11, pp. 14–26, Nov. 2016.

[87] K. F. Trillingsgaard and P. Popovski, "Downlink transmission of short packets: Framing and control information revisited," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2048–2061, May 2017.

[88] B. Lee, S. Park, D. J. Love, H. Ji, and B. Shim, "Packet structure and receiver design for low-latency communications with ultra-small packets," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.

[89] S.-Y. Lien, S.-L. Shieh, Y. Huang, B. Su, Y.-L. Hsu, and H.-Y. Wei, "5G new radio: Waveform, frame structure, multiple access, and initial access," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 64–71, Jun. 2017.

[90] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Feedback in the non-asymptotic regime," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4903–4925, Aug. 2011.

[91] H. Shariatmadari, S. Iraji, and R. Jantti, "Analysis of transmission methods for ultra-reliable communications," in *Proc. IEEE 26th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Aug. 2015, pp. 2303–2308.

[92] G. Hattab and D. Cabric, "Joint resource allocation and user association in multi-antenna heterogeneous networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–7.

[93] X. Ge, X. Li, H. Jin, J. Cheng, and V. C. M. Leung, "Joint user association and user scheduling for load balancing in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3211–3225, May 2018.

[94] Z. Su *et al.*, "User association and wireless backhaul bandwidth allocation for 5G heterogeneous networks in the millimeter-wave band," *China Commun.*, vol. 15, no. 4, pp. 1–13, Apr. 2018.

[95] E. Kalantari, I. Bor-Yaliniz, A. Yongacoglu, and H. Yanikomeroglu, "User association and bandwidth allocation for terrestrial and aerial base stations with backhaul considerations," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–6.

[96] J. Quittek, P. Bauskar, T. BenMeriem, A. Bennett, and M. Besson, "Network functions virtualisation (NFV)-management and orchestration," ETSI NFV ISG, Sophia Antipolis, France, White Paper, 2014.

[97] IMT-2020 Promotion Group. *5G Network Architecture Design*. Accessed: May 21, 2020. [Online]. Available: http://www.catr.cn/kxyj/qwfb/bps/201606/P020160624527039498728.pdf

[98] *Requirements for the Radio Interface(s) for International Mobile Telecommunications*, document Rec. ITU-R M: 1034-1, 2000. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.1034-1-199702-I!!PDF-E.pdf

[99] U. Mobile, *Universal Mobile Telecommunications System (UMTS); Selection Procedures for the Choice of Radio Transmission Technologies of the UMTS*, document TR 101, 1998.

[100] WINNER Group, "WINNER II channel models," Tech. Rep. IST-WINNER2, WINNER 2008. [Online]. Available: https://www.cept.org/files/8339/winner2%20-%20finalreport.pdf

[101] B. Ai *et al.*, "Radio wave propagation scene partitioning for high-speed rails," *Int. J. Antennas Propag.*, vol. 2012, no. 1, pp. 1072–1075, 2013.

[102] K. Guan *et al.*, "Excess propagation loss modeling of semiclosed obstacles for intelligent transportation system," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 8, pp. 2171–2181, Aug. 2016.

[103] R. He, Z. Zhong, B. Ai, J. Ding, Y. Yang, and A. F. Molisch, "Short-term fading behavior in high-speed railway cutting scenario: Measurements, analysis, and statistical models,"

*IEEE Trans. Antennas Propag.*, vol. 61, no. 4, pp. 2209–2222, Apr. 2013.

[104] R. He, Z. Zhong, B. Ai, and K. Guan, "Reducing the cost of high-speed railway communications: From the propagation channel view," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 4, pp. 2050–2060, Aug. 2015.

[105] Samsung Networks. (Oct. 2014). *Samsung Electronics Sets 5G Speed Record at 7.5 Gbps, Over 30 Times Faster Than 4G LTE*. Samsung Tomorrow. [Online]. Available: https://www.samsung.com/global/business/networks/insights/news/samsung-electronics-sets-5g-speed-record-at-7-5gbps-over-30-times-faster-than-4g-lte/

[106] J. Kim, H. S. Chung, I. G. Kim, H. Lee, and M. S. Lee, "A study on millimeter-wave beamforming for high-speed train communication," in *Proc. IEEE ICTC*, Oct. 2015, pp. 1190–1193.

[107] J. Yang, B. Ai, D. He, L. Wang, Z. Zhong, and A. Hrovat, "A simplified multipath component modeling approach for high-speed train channel based on ray tracing," *Wireless Commun. Mobile Comput.*, vol. 2017, no. 333, pp. 1–14, Oct. 2017.

[108] D. He *et al.*, "Channel measurement, simulation, and analysis for high-speed railway communications in 5G millimeter-wave band," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 10, pp. 3144–3158, Oct. 2018.

[109] Q. Wang, B. Ai, R. He, K. Guan, J. Li, and D. Fei, "Parameter estimation using SAGE algorithm based on massive MIMO channel measurements," in *Proc. IEEE Int. Symp. Antennas Propag. (APSURSI)*, Jun. 2016, pp. 5358–5674.

[110] *Effects of Building Materials and Structures on Radiowave Propagation Above About 100 MHz*, document ITU-R, P.2040, Jul. 2015. [Online]. Available: https://www.itu.int/rec/R-REC-P.2040-1-201507-I/en

[111] X. Zhang, N. Sood, J. Siu, and C. D. Sarris, "Calibration of a 3-D ray-tracing model in railway environments," in *Proc. IEEE Int. Symp. Antennas Propag. USNC/URSI Nat. Radio Sci. Meeting*, Jul. 2015, pp. 89–90.

[112] T. Kawanishi *et al.*, "Proposal of a new working document of a draft new apt report on millimeter-wave band railway radiocommunication systems between train and trackside and its work plan," in *Proc. 20th Meeting APT Wireless Group*, 2016, pp. 1–12.

[113] Z. Lai, H. Song, P. Wang, H. Mu, L. Wu, and J. Zhang, "Implementation and validation of a 2.5D intelligent ray launching algorithm for large urban scenarios," in *Proc. 6th Eur. Conf. Antennas Propag. (EUCAP)*, Mar. 2012, pp. 2396–2400.

[114] I. Wald, T. J. Purcell, J. Schmittler, C. Benthin, and P. Slusallek, "Realtime ray tracing and its use for interactive global illumination," *Eurographics STAR*, vol. 1, no. 3, p. 5, 2003.

[115] D. M. Rose and T. Kurner, "An analytical 3D ray-launching method using arbitrary polygonal shapes for wireless propagation prediction," in *Proc. IEEE 80th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2014, pp. 1–6.

[116] D. He, B. Ai, K. Guan, L. Wang, Z. Zhong, and T. Kurner, "The design and applications of high-performance ray-tracing simulation platform for 5G and beyond wireless communications: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 10–27, 1st Quart., 2019.

[117] J. Nuckelt, M. Schack, and T. Kurner, "Geometry-based path interpolation for rapid ray-optical modeling of vehicular channels," in *Proc. IEEE EuCAP*, Apr. 2015, pp. 1–5.

[118] K. Guan, D. He, A. Hrovat, B. Ai, Z. Zhong, and T. Kurner, "Challenges and chances for smart rail mobility at mmWave and THz bands from the channels viewpoint," in *Proc. 15th Int. Conf. Telecommun. (ITST)*, May 2017, pp. 1–5.

[119] K. Guan *et al.*, "Scenario modules, ray-tracing simulations and analysis of millimetre wave and terahertz channels for smart rail mobility," *IET Microw., Antennas Propag.*, vol. 12, no. 4, pp. 501–508, Mar. 2018.

[120] C. U. Bas *et al.*, "Real-time millimeter-wave MIMO channel sounder for dynamic directional measurements," *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 8775–8789, Sep. 2019.

[121] X. Lin *et al.*, "Measurement based ray tracer calibration and channel analysis for high-speed railway viaduct scenario at 93.2 GHz," in *Proc. IEEE Int. Symp. Antennas Propag. USNC/URSI Nat. Radio Sci. Meeting*, Jul. 2017, pp. 617–618.

[122] M. Di Renzo and H. Haas, "Bit error probability of SM-MIMO over generalized fading channels," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1124–1144, Mar. 2012.

[123] Z. Li, F. Wang, and Z. Zhong, "Multi-antenna transmission schemes in high speed railway communication," in *Proc. 5th IET Int. Conf. Wireless, Mobile Multimedia Netw. (ICWMMN)*, 2013, pp. 115–118.

[124] S. Wang, F. Wang, and Z. Li, "Space-time-coded adaptive spatial modulation in wireless MIMO communication systems," *Frequenz*, vol. 69, nos. 7–8, pp. 335–339, Jan. 2015.

[125] X. Wang, G. Wang, R. Fan, and B. Ai, "Channel estimation with expectation maximization and historical information based basis expansion model for wireless communication systems on high speed railways," *IEEE Access*, vol. 6, pp. 72–80, 2018.

[126] B. Ai *et al.*, "Future railway services-oriented mobile communications network," *IEEE Commun. Mag.*, vol. 53, no. 10, pp. 78–85, Oct. 2015.

[127] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, and L. Hanzo, "Cell-free massive MIMO: A new next-generation paradigm," *IEEE Access*, vol. 7, pp. 99878–99888, 2019.

[128] J. Zhang, Y. Wei, E. Bjornson, Y. Han, and S. Jin, "Performance analysis and power control of cell-free massive MIMO systems with hardware impairments," *IEEE Access*, vol. 6, pp. 55302–55314, 2018.

[129] Y. Jin, J. Zhang, S. Jin, and B. Ai, "Channel estimation for cell-free mmWave massive MIMO through deep learning," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 10325–10329, Oct. 2019.

[130] Y. Jin, J. Zhang, B. Ai, and X. Zhang, "Channel estimation for mmWave massive MIMO with convolutional blind denoising network," *IEEE Commun. Lett.*, vol. 24, no. 1, pp. 95–98, Jan. 2020.

[131] H. Liu, J. Zhang, X. Zhang, A. Kurniawan, T. Juhana, and B. Ai, "Tabu-search-based pilot assignment for cell-free massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 2286–2290, Feb. 2020.

[132] M. Gao *et al.*, "Dynamic mmWave beam tracking for high speed railway communications," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Apr. 2018, pp. 278–283.

[133] A. F. Molisch *et al.*, "Hybrid beamforming for massive MIMO: A survey," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 134–141, Sep. 2017.

[134] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[135] Y. Lu, K. Xiong, P. Fan, and Z. Zhong, "Optimal multicell coordinated beamforming for downlink high-speed railway communications," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9603–9608, Oct. 2017.

[136] N. DOCOMO. (Jul. 2014). *Docomo 5G White Paper-5G Radio Access: Requirements, Concept, and Technologies*. [Online]. Available: https://www.nttdocomo.co.jp/english/binary/pdf/corporate/technology/whitepaper_5g/DOCOMO_5G_White_Paper.pdf

[137] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[138] J. Piao, K. Niu, J. Dai, and C. Dong, "Approaching the normal approximation of the finite blocklength capacity within 0.025 dB by short polar codes," *IEEE Wireless Commun. Lett.*, early access, Mar. 17, 2020, doi: 10.1109/LWC.2020.2981432.

[139] Y. Xu, C. Shen, T.-H. Chang, S.-C. Lin, Y. Zhao, and G. Zhu, "Transmission energy minimization for heterogeneous low-latency NOMA downlink," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1054–1069, Feb. 2020.

[140] J. Lu, K. Xiong, X. Chen, and P. Fan, "Toward traffic patterns in high-speed railway communication systems: Power allocation and access selection," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 12273–12287, Dec. 2018.

[141] Y. Lu, K. Xiong, P. Fan, Z. Zhong, and B. Ai, "The effect of power adjustment on handover in high-speed railway communication networks," *IEEE Access*, vol. 5, pp. 26237–26250, 2017.

[142] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.

[143] G. Gur, "Spectrum sharing and content-centric operation for 5G hybrid satellite networks: Prospects and challenges for space-terrestrial system integration," *IEEE Veh. Technol. Mag.*, vol. 14, no. 4, pp. 38–48, Dec. 2019.

[144] O. Jo, Y.-K. Kim, and J. Kim, "Internet of Things for smart railway: Feasibility and applications," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 482–490, Apr. 2018.

[145] Y. Wu, A. Khisti, C. Xiao, G. Caire, K.-K. Wong, and X. Gao, "A survey of physical layer security techniques for 5G wireless networks and challenges ahead," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 679–695, Apr. 2018.

[146] R. Li *et al.*, "Intelligent 5G: When cellular networks meet artificial intelligence," *IEEE Wireless Commun.*, vol. 24, no. 5, pp. 175–183, Oct. 2017.

[147] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, and D. Sabella, "On multi-access edge computing: A survey of the emerging 5G network edge cloud architecture and orchestration," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1657–1681, 3rd Quart., 2017.

[148] K. David, J. Elmirghani, H. Haas, and X.-H. You, "Defining 6G: Challenges and opportunities [From the guest Editors]," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 14–16, Sep. 2019.

[149] J. Song *et al.*, "Key technologies and measurements for DTMB—A system," *IEEE Trans. Broadcast.*, vol. 65, no. 1, pp. 53–64, Mar. 2019.

[150] *Frame Structure Channel Coding and Modulation for a Second Generation Digital Terrestrial Television Broadcasting System (DVB-T2)*, ETSI Standard EN 302 755, 2012. [Online]. Available: https://www.etsi.org/deliver/etsi_en/302700_302799/302755/01.03.01_60/en_302755v010301p.pdf

[151] L. Fay, L. Michæl, D. Gómez-Barquero, N. Ammar, and M. W. Caldwell, "An overview of the ATSC 3.0 physical layer specification," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 159–171, Mar. 2016.

## ABOUT THE AUTHORS

**Bo Ai** (Senior Member, IEEE) graduated from Tsinghua University, Beijing, China, with the honor of Excellent Postdoctoral Research Fellow in 2007. He received the master's and Ph.D. degrees from Xidian University, Xi'an, China, in 2002 and 2004, respectively.

He was a Visiting Professor with the Electrical Engineering (EE) Department, Stanford University, Stanford, CA, USA, in 2015. He is currently working as a Full Professor and a Ph.D. Advisor with Beijing Jiaotong University, Beijing, where he is also the Deputy Director of the State Key Laboratory of Rail Traffic Control and Safety and the International Joint Research Center. He is one of the main responsible people for Beijing (urban rail operation control system) International Science and Technology Cooperation Base, and the Backbone Member of the Innovative Engineering-Based jointly granted by the Chinese Ministry of Education and the State Administration of Foreign Experts Affairs. He has authored or coauthored eight books and published over 300 academic research articles in his research area. He holds 26 invention patents. He has been the research team leader for 26 national projects and has won some important scientific research prizes. Five of his articles have been the ESI highly cited articles. He has been notified by the Council of Canadian Academies (CCA) that, based on Scopus database, he has been listed as one of the Top 1% authors in his field all over the world. He has also been feature interviewed by IET Electronics Letters. His research interests include the research and applications of channel measurement and channel modeling and dedicated mobile communications for rail traffic systems.

Dr. Ai is a Fellow of the Institution of Engineering and Technology (IET). He is also an Editorial Committee Member of *Wireless Personal Communications* journal. He has received many awards, such as the Distinguished Youth Foundation and Excellent Youth Foundation Award by the National Natural Science Foundation of China, the Qiushi Outstanding Youth Award by the Hong Kong Qiushi Foundation, the New Century Talents by the Chinese Ministry of Education, the Zhan Tianyou Railway Science and Technology Award by the Chinese Ministry of Railways, and the Science and Technology New Star Award by the Beijing Municipal Science and Technology Commission. He is also a Distinguished Lecturer of the IEEE Vehicular Technology Society, a Vice Chair of the IEEE VTS Beijing Chapter, and the Chair of the IEEE BTS Xi'an Chapter. He was the co-chair or the session/track chair for many international conferences. He is also an Associate Editor of the IEEE ANTENNAS AND WIRELESS PROPAGATION LETTERS and the IEEE TRANSACTIONS ON CONSUMER ELECTRONICS. He is the Lead Guest Editor for Special Issues on the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE ANTENNAS AND PROPAGATIONS LETTERS, and the *International Journal on Antennas and Propagations*.

**Andreas F. Molisch** (Fellow, IEEE) received the Dipl.Ing., Ph.D., and Habilitation degrees from the Vienna University of Technology, Vienna, Austria, in 1990, 1994, and 1999, respectively.

He spent the next ten years in industry at FTW, Austria, AT&T (Bell) Laboratories, USA, and Mitsubishi Electric Research Laboratories, Cambridge, MA, USA, where he rose to the Chief Wireless Standards Architect. In 2009, he joined the University of Southern California (USC), Los Angeles, CA, USA, as a Professor, where he founded the Wireless Devices and Systems (WiDeS) Group. In 2017, he was appointed as the Solomon Golomb–Andrew and Erna Viterbi Chair. Overall, he has published four books (among them the textbook *Wireless Communications* currently in its second edition), 20 book chapters, 250 journal articles, and 350 conference papers. He is also an inventor of 60 granted (and more than 20 pending) patents and a coauthor of some 70 standard contributions. His research interests revolve around wireless propagation channels, wireless systems design, and their interaction. Recently, his main research interests include wireless channel measurement and modeling for 5G and beyond 5G systems, wireless video distribution, hybrid beamforming, ultrawideband/time-of-arrival-based localization, caching at the wireless edge, and novel modulation/multiple access methods.

Dr. Molisch is a Fellow of the National Academy of Inventors, the American Association for the Advancement of Science, and the Institution of Engineering and Technology (IET). He is also a member of the Austrian Academy of Sciences. He has been an editor of a number of journals and special issues, the general chair, the technical program committee chair, or the symposium chair of multiple international conferences, and the chairman of various international standardization groups. He is also an IEEE Distinguished Lecturer. He has received numerous awards, among them the IET Achievement Medal, the Technical Achievement Awards of the IEEE Vehicular Technology Society (Evans Avant-Garde Award), the IEEE Communications Society (Edwin Howard Armstrong Award), the Technical Field Award of the IEEE for Communications, and the Eric Sumner Award.

**Markus Rupp** (Fellow, IEEE) received the Dr.Ing. degree from the Technische Universität Darmstadt, Darmstadt, Germany, in 1993.

He served as the Dean from 2005 to 2007 and 2016 to 2019 and the Head of the Institute of Telecommunications, Technische Universität Wien, Vienna, Austria, from 2014 to 2015, where he has been a Full Professor since 2001. He has authored or coauthored more than 650 scientific articles and patents on adaptive filtering, wireless communications, rapid prototyping, and automatic design methods.

Dr. Rupp was with the Board of Directors (BoD) of EURASIP from 2004 to 2011, including two years serving as the President of the society.

**Zhang-Dui Zhong** (Senior Member, IEEE) is currently a Professor and an Advisor of Ph.D. candidates with Beijing Jiaotong University, Beijing, China. He is also the Director with the School of Computer and Information Technology and a Chief Scientist with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University. He is also the Director of the Innovative Research Team, Ministry of Education, and a Chief Scientist of the Ministry of Railways, Beijing. He has authored or coauthored seven books, five invention patents, and more than 200 scientific research articles in his research area. His research interests include wireless communications for railways, control theory and techniques for railways, and global system for mobile communications for railway (GSM-R) systems. His research has been widely used in the railway engineering, such as Qinghai-Xizang Railway, Datong-Qinhuangdao Heavy Haul railway, and many high-speed railway lines of China.

Dr. Zhong is an Executive Council Member of the Radio Association of China and the Deputy Director of the Radio Association of Beijing. He received the Mao Yisheng Scientific Award of China, the Zhan Tianyou Railway Honorary Award of China, and the Top Ten Science/Technology Achievements Award of Chinese Universities.