Abstract—Since the introduction of fifth-generation new radio (5G-NR) in Third Generation Partnership Project (3GPP) Release 15, swift progress has been made to evolve 5G with 3GPP Release 18 emerging. A critical aspect is the design of massive multiple-input multiple-output (MIMO) technology. In this line, this paper makes several important contributions: We provide a comprehensive overview of the evolution of standardized massive MIMO features from 3GPP Release 15 to 17 for both time/frequency-division duplex operation across frequency range in (FR)-1 and FR 2-1. We analyze the progress on channel state information (CSI) frameworks, beam management frameworks and present enhancements for uplink CSI. We shed light on emerging 3GPP Release 18 problems requiring imminent attention. These include advanced codebook design and sounding reference signal design for coherent joint transmission (CJT) with multiple transmission/reception points (multi TRPs). We discuss advancements in uplink demodulation reference signal design, enhancements for mobility to provide accurate CSI estimates, and unified transmission configuration indicator framework tailored for FR 2-1 bands. For each concept, we provide system level simulation results to highlight their performance benefits. Via field trials in an outdoor environment at Shanghai Jiaotong University, we demonstrate the gains of multi-TRP CJT relative to single TRP at 3.7 GHz.

Index Terms—3GPP, 5G-NR, beamforming, codebooks, coherent joint transmission, CSI, massive MIMO.

I. INTRODUCTION

THE utilization of massive multiple-input multiple-output (MIMO) technology has been one of the defining features of fifth-generation new radio (5G-NR) systems [1], [2]. With the advent of active antenna arrays, configuring large numbers of antenna elements (tens to hundreds) at radio base stations (BSs) has become feasible from a design and implementation standpoint. Commonly, uniform planar arrays (a.k.a. uniform rectangular arrays) spanning two-dimensions (2D) are employed at BSs, to serve multiple users (UEs) within the same time-frequency resource via spatial beamforming across the azimuth and elevation domains [3]. Beamforming allows the BS to form multiple electronically steerable beams in the desired direction(s) of UE(s) to enhance coverage and spectral efficiency performance. Irrespective of the frequency band of operation, to ensure such higher performance in the downlink, 5G-NR systems use link adaptation to select a suitable modulation and coding rate (MCS) accommodating for the time-varying propagation channel conditions between the BS and UE(s) [4]. With time-division duplex (TDD) operation, the downlink transmission is adapted based on uplink channel estimates, assuming reciprocity between the uplink and downlink channels, respectively. In contrast, for frequency-division duplex (FDD) operation, link adaptation is based on the uplink feedback subsequent to the channel estimation in the downlink. The estimation and acquisition of channel knowledge at both the BS and UEs is referred to as channel state information (CSI) [5]. For both TDD and FDD operation, a CSI report in 5G-NR primarily comprises of three components: (1) rank indicator (RI), (2) precoding matrix identifier (PMI), and (3) channel quality indicator (CQI) [6].

The required massive MIMO features and functionality starkly differ between the two Third Generation Partnership Project (3GPP) classified frequency ranges (FR), i.e., FR-1 which spans from 410 to 7,125 MHz and FR 2-1 which spans from 24,250 to 52,600 MHz, respectively [2], [5], [7]. For systems operating in the FR-1 band, digital beamforming is typically implemented at the BS, where a large array of $N_t$ dual-polarized transmit antennas is used to generate $N_b$ digitally controlled beams ($N_b \ll N_t$) at the array baseband (prior to up-conversion). Here, each beam constitutes a reference signal (RS) [4]. Contrary to this, in the FR-2-1 band, almost all current commercial BSs only support analog beamforming, generating $N_b$ analog beams (after up-conversion) from the available $N_t$ transmit antennas, each constituting a RS resource. FR 2-1 bands are primarily restricted to analog beamforming due to the implementation complexity and energy consumption issues arising from supporting wider carrier bandwidths at much higher center frequencies relative to FR-1 bands. In particular, lower power amplifier (PA) efficiencies (typically 10 to 20%) and the need for a dedicated digital-to-analog converter(s) (DAC(s)) for each antenna element limits the scope of FR 2-1 band transmission to localized per-antenna phase shifts with complimentary power divider and combiner circuits [7].

For FR-1 bands, the first 5G-NR specification, i.e., 3GPP Release 15, supports low (Type-I) and high (Type-II) resolution

1Typically, the state-of-the-art literature denotes the FR-1 band as conventional microwave bands, while the FR 2-1 bands are referred to as millimeter-wave (mmWave) bands, see e.g., [2] and references therein.
codebooks (a.k.a. Type-I and Type-II CSI) for beamforming to single and multiple UEs [8]. For multiuser scenarios, Type-II codebooks are intended to facilitate quantized channel eigenvector feedback, which is especially beneficial for combating inter-UE interference while spatially multiplexing data streams to multiple UEs. 3GPP Release 16 built on the codebooks introduced in Release 15 and introduced enhanced Type-II codebooks which reduced the associated feedback overheads via the use of spatial domain compression, while facilitating transmission of 4 spatial streams per-UE relative to 2 streams in Release 15 [9]. For both TDD and FDD systems, 3GPP Release 17 codebooks exploit the joint angle-delay reciprocity in the downlink and the uplink. By doing so, the spatial-domain/frequency-domain compression operation inherent in Release 16 enhanced Type-II codebooks can be shifted to the BS, thereby reducing UE computational load in exchange with additional BS processing complexity. For FDD systems, while the amplitude and phase per-propagation path are generally not downlink and uplink reciprocal, the BS can employ angle-delay information obtained from uplink measurements to beamform using specific CSI-RS. Further information about exploitation of joint angle-delay reciprocity can be found in [6] and [10]. For downlink transmission in the FR-2-1 bands, CSI-RS and synchronization signal blocks (SSBs) are used, where a signal transmitted along a beam is received at the UE(s) by selecting one of the candidate receive beams that best matches the transmitted beam in terms of the received signal-to-noise ratio (SNR) of the UE(s). A host of 3GPP procedures to facilitate this operation is referred to as beam management (see [11]). While 3GPP Release 15 specifies beam management sufficient to support fixed wireless access services, its overheads and latency become a bottleneck for moderate-to-high-speed UE mobility, which is addressed by 3GPP Release 16 and 17, respectively [12]. For both FR-1 and FR 2-1 bands, CSI enhancements for the uplink are also heavily investigated with proposals for different codebook vs. non-codebook approaches [13] and the possibility of utilizing different demodulation reference signal (DMRS) configuration [14].

Despite the above advances, to the best of our knowledge, the existing literature does not holistically capture most aspects of massive MIMO evolution for both TDD and FDD systems, across both FR-1 and FR 2-1 bands from the beginning of Release 15 to Release 18. We bridge this gap by addressing different aspects of massive MIMO evolution in a continuum.

In parallel to the above developments, 3GPP Release 16 and 17 standardized the so-called multi-transmission/reception point (multi-TRP) functionality [15]. This helped to increase the reliability of 5G NR-NS systems by providing spatial macro diversity so that if one propagation path is blocked, an alternative path from a second TRP can be used for data transmission. In particular, the transmission configuration indicator (TCI) state was designed to play a key role in multi-TRP transmission, providing information required to identify and track a reference signal for downlink reception. In the uplink, a counterpart of TCI state is spatial relation, which identifies a reference signal to be used as an uplink beam, especially for FR 2-1 bands. In the case of multi-TRP transmission, more than one TCI state or spatial relation is signaled to UE for transmission/reception of a given channel. The 3GPP Release 16 multi-TRP transmission mechanisms support both coherent and non-coherent joint transmission, a.k.a., CJT and NCJT, for TDD and FDD systems, respectively [15]. The 3GPP Release 16 multi-TRP transmission mechanisms support both coherent and non-coherent joint transmission, a.k.a., CJT and NCJT, for TDD and FDD systems, respectively [15]. CJT combines all transmit antennas within the cooperating set of distributed antenna arrays into a single distributed array, such that the transmissions are synchronized and phase coherent across the cooperating set. The inverse of CJT is NCJT. This results in significant downlink data rate increases (particularly for UEs that are closer to the radio cell-edge) alongside an increase in system reliability.

In the 3GPP Rel-18, MIMO techniques continue to be enhanced to fulfill the request for evolution of new radio (NR) deployments [16]. The Rel-18 MIMO evolution targets at both downlink and uplink enhancements, which facilitates the use of multiple and large antenna arrays for both FR1 and FR2. In particular, the 3GPP Rel-18 MIMO evolution focuses on the following several agenda items: multi-TRP enhancements including extending the Rel-17 single-TRP based unified TCI framework to multi-TRP scenarios and support of two timing advances (TAs); CSI enhancements for coherent joint transmission and UE with medium and high mobility; support of larger number of orthogonal demodulation reference signal (DMRS) ports for multi-user MIMO (MU-MIMO); support of 8 antenna ports in uplink and simultaneous multi-panel transmission to better enable advanced UEs such as customer premise equipment (CPE), fixed wireless access (FWA) devices and vehicular UEs.

Nonetheless, for CJT systems operating in the FR-1 band under FDD mode, optimal design of beamforming codebooks from a CSI overhead minimization and spectral efficiency maximization viewpoint remains an open problem. This is in addition to selecting an optimal TRP coordination set for enhancing the overall system performance for joint transmission. To address this gap, we propose a modification to the contemporary 3GPP Release 16 and 17 oversampled and orthogonal 2D discrete Fourier transform (DFT) beams used in codebook design, where we incorporate channel feedback based on spatial and frequency domain basis. Our idea is to utilize the channel correlation in either frequency or spatial domain to seek a sparse basis for an approximate representation of the overall beamforming matrix which helps to reduce the required overheads and helps to improve performance. Likewise, for CJT systems operating in the TDD mode for FR-1 bands, the strict requirements on the CSI accuracy drives the need for new solutions. This is since the beamformers designed at the remote radio heads (RRHs) of the coordinated cells are required to ensure that the phase of the received signals from different cells add up constructively based on the acquired CSI. In such a case, an important factor that limits the CSI accuracy is the inter-cell interference. The current interference suppression methods rely on whitening the colored interference by low correlation sequences during the least squares (LS) operation at the BS. However, the
sequence correlation is limited by its length and is difficult to further improve. In such cases, to achieve better SNR for CSI acquisition, we propose a modification to the original CSI acquisition framework using sounding reference signals (SRS). Specifically, we extend the concept of SRS frequency hopping, leading to shorter root sequences being used. To resolve the inter-cell SRS interference/collision (especially for the case with limited SRS resources in the system), we introduce SRS interference randomization to improve the SRS-based channel estimation performance in scenarios with strong inter-cell interference. Via system-level simulations (SLS),\(^3\) we later demonstrate the gains of using the abovementioned proposals for both FDD and TDD systems. To the best of our knowledge, this paper is the first to make multiple advances for CJT operation within the FR-1 bands.

In addition to the above, to demonstrate the gains achieved by CJT in a TDD system with single and multi-TRP, we present results from a field measurement conducted in an outdoor urban environment at Shanghai Jiao Tong University, Shanghai, China. The measurements were carried out at 3.5 GHz center frequency across a bandwidth of 20 MHz, where both single and multi-UE capabilities were measured. The single UE measurement was carried out with a 5 kilometer/hour (km/h) UE mobility along a defined trajectory, whereas the multi-UE measurements focused on the performance enhancements delivered at the cell edge. Both cases compared the resulting UE spectral efficiency performance with multi-TRP relative to a single TRP case as a baseline. The transmission TRP set consisted of 2 TRPs (intra-site), while the interfering TRP set consisted of 4 TRPs (intra and inter-site). Each BS was equipped with 64 transmit/receive radio chains while each UE (Huawei Mate30) was equipped with 2 transmit and 4 receive radio chains. To the best of our knowledge, such an extensive measurement campaign has not previously been carried out for analyzing 3GPP compliant CJT, where the literature contains examples of measurement results in more simplified scenarios \([17], [18]\).

For FR 2-1 bands, the literature is sparse on the 3GPP Release 17 beam management enhancements in high mobility scenarios (channel coherence time proportional to UE velocities exceeding 100 km/h), see e.g., \([6]\) and references therein. To bridge this gap, we present SLS in a dense urban highway (DUH) for beam management according to 3GPP Release 17 for the multi-TRP case. The motivation behind a unified TCI framework to reduce latency and overhead of beam indication is introduced, thereby enhancing system performance especially in high mobility scenarios. With a pre-defined path, the UE was designed to moving at a speed of 120 km/h (33.3 m/sec) and the UE’s throughput was sampled every 1 m (30 ms), with 100 sample points i.e., total simulation duration is 3 seconds. SLS performance is evaluated for beam indication using downlink control information (DCI), with a latency of 0.5 ms and a block error rate (BLER) of 1% and using medium access control element (MAC-CE), with a latency of 3 ms and a BLER of 10%.\(^3\)

Collectively, the key contributions of this paper can be summarized as follows:

- We present a holistic overview of massive MIMO evolution in 3GPP standardization for both FR-1 and FR 2-1 frequency bands, across FDD and TDD configurations from Release 15 to 17. These form the basis for the current MIMO study item of Release 18, a.k.a. 5G-Advanced. Our overview includes the evolution of CSI acquisition frameworks and techniques; Enhancements towards higher-order massive MIMO support on the uplink via CSI and DMRS enhancements; CSI enhancements for mobility; and beam management frameworks. To the best of our knowledge, this is the first paper to address different aspects of massive MIMO evolution in a continuum.

- For FDD CJT systems with multi-TRPs, we consider the problem of optimal design of beamforming codebooks to minimize CSI overheads and maximize per-UE spectral efficiency. In this line, we propose high resolution multi-TRP codebook based on 3GPP Release 16 and 17 codebooks, to match the channel property of coordinated transmission in multi-TRP case. Equivalently, for TDD-based CJT with multi-TRPs, we extend the concept of SRS frequency hopping and randomization, resolving the inter-cell SRS interference/collision while improving SRS-based channel estimation performance in scenarios with strong inter-cell interference. For both cases, we present SLS results for quantifying the resulting performance gains relative to a single TRP system.

- We present field measurements to demonstrate the gains achieved by CJT with single and multi-TRPs. The measurements were conducted in an outdoor urban environment at Shanghai Jiao Tong University, at 3.5 GHz with a bandwidth of 20 MHz, where both single and multi-UE capabilities were measured. The single UE measurement was carried out with a 5 km/h UE mobility, whereas the multi-UE measurements considered four cases around on the performance enhancements delivered at the cell edges. Both cases compared the resulting UE spectral efficiency performance with multi-TRP relative to a single TRP case as a baseline. To the best of our knowledge, such an extensive measurement campaign has not been carried out for analyzing 3GPP compliant CJT previously.

- Via SLS, we demonstrate the enhancements on offer by 3GPP Release 17 beam management framework for FR 2-1 bands with high UE mobility. In particular, we show how a unified TCI framework reduces latency and overheads, improving system performance.

**Notation:** Upper and lower boldface letters represent matrices and vectors respectively. The $M \times M$ identity matrix is denoted as $I_M$. The Hermitian transpose and ceiling operations are denoted by $(\cdot)^H$ and $\lceil \cdot \rceil$, whereas the statistical expectation is denoted by $E(\cdot)$. Moreover, $\left\| \cdot \right\|$ and $| \cdot |$ denote the vector and scalar norms, while $a \otimes b$ denotes the Kronecker product of two vectors $a$ and $b$. Finally, we denote $O(1)$ as an order one term, while $\text{FFT}(\cdot)$ and $\text{IFFT}(\cdot)$ represent the fast Fourier transform and inverse fast Fourier transform,

---

\(^3\)A SLS spans multiple layers of the open systems interconnection stack (physical, data link, access control) and implements 3GPP recommended procedures to compute UE or BS performance in a system setup as a network.
respectively. All other mathematical notation outside of what is stated here is defined within the text as necessary.

II. MASSIVE MIMO TECHNOLOGY EVOLUTION IN 5G-NR

In this section, we discuss the evolution of massive MIMO technologies in 5G-NR systems, including the typical massive MIMO architecture at radio BSs, CSI reporting and acquisition frameworks for FDD and TDD systems, respectively, codebook-based beamforming, support in 3GPP for higher-order massive MIMO transmission/reception (serving many UEs), and mmWave enhancements in the FR 2-1 bands.

A. Massive MIMO BS Architecture in 5G-NR

For both the FR-1 and FR 2-1 bands, three categories of BSs have been standardized by the 3GPP in TS 38.104; namely BS Type 1-C, 1-H and 1-O/2-O [19]. Most commercial FR-1 BSs are of type 1-H, where the radiation characteristics are defined over the air (OTA) and the conducted characteristics are defined at the boundary between the physical antenna ports (inclusive of radio distribution network) and transceiver antenna units. In contrast, commercial BSs for FR 2-1 bands are of the type 1-O/2-O, where the conducted characteristics are inclusive of the physical antenna ports and transceiver units.

Taking the FR-1 bands as an example, the massive MIMO structure in terms of its architecture and design is depicted in Fig. 1. Such an architecture is common for 3GPP compliant commercial BSs and consists of three key components: (1) Dual-polarized micro-strip (patch) antenna elements; (2) Sub-array(s); and (3) An overall antenna array composing of multiple sub-arrays. In general, the deployed antenna elements are capable of simultaneously transmitting and receiving two mutually orthogonal signals leveraging maximum diversity in the polarization domain. Naturally, the element cross-polarization discrimination governs the spatial path isolation between the horizontal (H) and vertical (V) polarizations without increasing the form-factor of the element. Considering the fact that the azimuth and zenith angles would be confined to a range governed by the distribution of the UEs, and that FR-1 BSs are typically deployed on rooftops or masts, coverage across the full zenith range of $-90^\circ$ to $90^\circ$ (usually) not required. To this end, a few adjacent antenna elements (typically in the elevation direction) are grouped together and are driven by a single up/down-conversion chain (including one PA and low-noise amplifier). This is known as a sub-array. An example of a $3 \times 1$ sub-array with elements spaced $d_x$ apart in elevation in a massive MIMO radio is shown in Fig.1a). Since only one digital beamforming weight is applied within a given sub-array, the adjustment of a sub-array’s steering angle requires an additional phase shifter network. The overall array then composes of multiple sub-arrays as its elements/components. As shown in Fig. 1a, four vertical columns of sub-arrays and four horizontal rows of sub-arrays create the 2D uniform planar array.

Such an architecture can be utilized by transmitting UE-specific beams towards individual receive antennas of the UE via a beamforming weight vector. Based on the three array components discussed above, the fundamental characteristics which determine 3GPP compliant commercial massive MIMO BS performance can be defined as follows:

- **UE-specific beam gain:** The maximum antenna gain of a UE-specific beam is expressed via a sum of its array gain and the sub-array gain, respectively. The array gain is proportional to the number of sub-arrays in a given array, and the sub-array gain is proportional to the number of elements grouped within a given sub-array, as well as the inter-element spacing between adjacent elements. The antenna gain towards the array broadside is governed by the antenna aperture of the massive MIMO BS, irrespective of its form factor.
- **Narrower beamwidths:** From array theory fundamentals, large arrays can yield narrower beams, which help to reduce the interference levels to other co-scheduled UEs who may be falling within the array’s sidelobes. Naturally, the beamwidth is inversely proportional to the length of the massive MIMO array in both azimuth and elevation domains.
• Elevation angular coverage: The sub-array pattern does not change dynamically unless an additional phase shifter network is added. Thus, the operational range of the elevation steering angle is limited to the sub-array beamwidth — often referred to as the angular coverage of the massive MIMO system. Since a sub-array usually comprises of a single elevation column, as shown in Fig. 1 a) and b), the elevation angular coverage is inversely proportional to the length of the sub-array.

The above discussed massive MIMO architecture is used in 3GPP standardization to develop the necessary CSI acquisition and control/data signal transmission/reception frameworks. In what follows, we discuss some of these aspects, as well as their evolution across the different 3GPP standards releases from the beginning of Release 15 to Release 18.

B. CSI Reporting and Acquisition for FDD and TDD Systems

The premise of beamforming, and more generally the evolution of massive MIMO in 3GPP standardization, is heavily dependent on the quality and accuracy of the acquired CSI at the BS and UEs. For FR-1 bands, FDD systems use CSI-RS as a signaling mechanism where the BS facilitates downlink transmission by sending a digitally beamformed signal towards the UE, which is received using digital combining across the UE’s receive antennas. The BS derives its beamforming weights from CSI reports calculated by the UE according to a pre-defined codebook of beamforming vectors/matrices. In this line, 3GPP Release 15 supported Type-I (low) and Type-II (high-resolution) CSI codebooks, respectively [6]. Type-II CSI is intended to facilitate refined resolution quantized channel eigenvector feedback, which is tailored for spatially multiplexing data to multiple UEs, enabling higher-order massive MIMO operation. While these are sufficient to support fixed wireless access services, its overheads and latency become a bottleneck for moderate-to-high-speed UE mobility, which is addressed by Release 16 and 17, respectively. Additionally, Release 15 lacks support for carrier aggregation, which is alleviated in Release 16 and 17.

A typical CSI report consists of RI, PMI and CQI, and each possible value of PMI corresponds to one possible beamformer configuration. To this end, the set of all possible values of PMIs corresponds to a set of beamformers, referred to in the 3GPP terminology as beamforming codebook, which the UEs can select between when reporting PMI. In the sequel, we discuss the components and design of Type-I and Type-II codebooks, a.k.a., Type-I and II CSI, and present their evolution across the different 3GPP releases.

1) Type-I Codebooks/CSI: Type-I codebook/CSI primarily targets scenarios where a single UE is scheduled within a given time-frequency resource, potentially with transmission of a relatively large number of layers in parallel, i.e., higher-order spatial multiplexing. Type-I CSI composes of two sub-types of CSI, a.k.a., Type-I Single Panel CSI and Type-I Multiple Panel CSI, respectively, corresponding to different codebooks assuming single and multiple panel BS array configurations [13]. In the case of Type-I single panel CSI, the beamforming matrix, \( W \), is then given by \( W = W_1 W_2 \) with information about the selected \( W_1 \) and \( W_2 \) reported separately as different parts of the PMI [6], [11]. Note that \( W_1 \) is designed to capture the long-term frequency-independent characteristics of the propagation channel. To this end, one \( W_1 \) is selected and reported for the entire CSI reporting bandwidth. In contrast, \( W_2 \) is designed to capture short-term frequency-dependent characteristics of the channel.

2) Type-II Codebooks/CSI: Relative to Type-I, Type-II codebooks/CSI provides channel information with significantly higher spatial granularity. Primarily, targeting the multi-UE scenario, it was intentionally limited to a maximum of rank two in Release 15. This was later extended to rank four in Release 16. The mathematical structure of Type-II CSI follows \( W = W_1 W_2 \) with \( W_1 \) defining a set of beams that is reported as part of the CSI. For each beam, \( W_2 \) is designed to provide an amplitude and phase control. We now discuss the evolution of Type-II CSI towards enabling multi-UE capabilities in Release 16.

3) 3GPP Release 16 Type-II Enhanced Codebook: The reporting of a relatively large number of combining coefficients on a per-sub-band basis yields large CSI reporting overheads for FDD massive MIMO systems. Hence the Enhanced Type-II codebook/CSI in Release 16 enables the utilization of frequency domain correlation across the sub-bands to reduce signaling overheads. While doing this, Release 16 offers a 2× improvement in the frequency domain granularity of PMI reporting. This is jointly achieved via the concept of frequency-domain units, where a unit corresponds to either a sub-band or half a sub-band in conjunction with the compression operation. The BS is provided with a recommended beamformer for a frequency-domain unit, relative to one beamformer per-sub-band for Release 15 Type-II CSI. For a given layer, \( \ell \), the Release 16 Type-II CSI can be expressed as

\[
\begin{bmatrix}
    \mathbf{u}_{\ell}^{(0)} \mathbf{u}_{\ell}^{(1)} \cdots \mathbf{u}_{\ell}^{(F-1)}
\end{bmatrix} = W_1 \left( \overline{W}_{2,\ell} W_{\ell,\ell}^H \right),
\]

(1)

where \( f \in F \) is the total number of frequency-domain units to be reported. Note that the left-hand side of (1) describes the set of beamforming vectors for the complete set of frequency-domain units for a given layer. For a total of \( L \) layers, there are \( L \) such beamforming vectors, each consisting of \( F \) different configurations. The actual reported beamformer which maps the involved massive MIMO layers to antenna ports for a given frequency-domain unit, \( x \), can then be computed as

\[
W^{(x)} = \begin{bmatrix} \mathbf{v}_{0}^{(x)} & \mathbf{v}_{1}^{(x)} & \cdots & \mathbf{v}_{L-1}^{(x)} \end{bmatrix}.
\]

(2)

The structure of \( W_1 \) is as given in (1) and is the same for all frequency-domain units and layers, respectively. The size of the compression matrix \( W_{\ell,\ell}^H \) is \( Z \times F \), which comprises of a set of row vectors from a DFT basis and provides a transformation from the frequency-domain of dimension \( F \) into a smaller delay domain of dimension \( Z (Z < F) \). We note that \( W_{\ell,\ell}^H \) is frequency independent and is reported separately for each layer \( \ell \). The number of rows of \( W_{\ell,\ell}^H \) is \( Z = \lfloor p(f/\phi) \rfloor \), where \( \phi \) is the number of frequency-domain units per-sub-frame and \( p \) is a tunable parameter which controls the amount of compression. Furthermore, \( \overline{W}_{2,\ell} \) maps a smaller version of delay domain to the beam domain, implying a
To enable better trade-offs in UE complexity, CSI reporting overhead and performance, Release 17 Type-II CSI is introduced by utilizing joint angle and delay reciprocity between the downlink and uplink channels. Specifically, the BS will employ the angle-delay information obtained from the uplink channel to weight CSI-RS. As a result, based on the CSI-RS weighted by angle and delay pairs. By applying this frequency domain information on CSI-RS, it is equivalent to the BS shifting the delay of $p$-th angle-delay pair to a specific delay position, $\tau$. As a result, based on the CSI-RS weighted by angle and delay information, UE can obtain the coefficients according to the specific delay position. Followed by this, the UE can calculate the linear combination coefficient, $C_p$, for the $p$-th angle-delay pair on the CSI-RS for port $p$ based on a specific frequency information, avoiding the need to find the appropriate frequency-domain vectors.

For TDD systems, channel reciprocity holds between the downlink and uplink, and CSI acquisition is based on SRS initiated by the UE(s) towards the BS. Specifically, Release 15 defines the basic SRS functions for UEs equipped with up to 4 receive antennas. In practice, given that the number of receive radio chains at the UE is generally larger than the transmit radio chains, an antenna switching mechanism is supported. For instance, assuming that BS has 64 transmit/receive chains (a.k.a. 64T64R), and the UE has 2 transmit and 4 receive chains (a.k.a. 2T4R), via the principle of reciprocity, the BS needs CSI of the channel $H$ having dimensions of $4 \times 64$ in order to design the downlink beamformer for the UE. The UE could sound the 4 receive chains across two different times to let the BS derive $H_1$ of the first two receive chains and $H_2$ of the remaining two receiving chains separately, where the dimension of both $H_1$ and $H_2$ is $2 \times 64$, and $H = \begin{bmatrix} H_1^T & H_2^T \end{bmatrix}^T$. As shown in Fig. 2, two orthogonal frequency-division multiplexing (OFDM) symbols are occupied by SRS from the UE, and each of the symbols contains the SRS transmitted via 2 of the 4 receive chains. Note that, each receive radio chain corresponds to an SRS port, and different SRS ports should thus occupy orthogonal physical resources.

In contrast to Release 15, Release 16 introduced more flexible antenna switching mechanisms for SRS overhead reduction. For example, in Release 15, the 2T4R UE would be configured with 4 SRS ports, while in Release 16, the UE has the flexibility to fall back to 1T2R configuration, which means that only 2 SRS ports are needed to acquire CSI. The channel corresponding to the two remaining receive radio chains are not transmitted and can be derived using the correlation between the transmit antennas and the antennas that are not transmitting associated with the respective radio chains. Advancing further, Release 17 increases the antenna switching mechanism to 8 receive radio chains.

In addition to the above, Release 15 also supported the SRS frequency hopping mechanism for coverage enhancement. As shown in Fig. 3, the UEs would transmit SRS on only part of the sounding bandwidth each time and using multiple times to sound the whole system bandwidth. To this end, the power spectrum density is increased compared with the case when SRS is transmitted on whole band within one time slot. In Fig. 3, the BS can estimate the CSI across the whole bandwidth after hopping 4 times. Release 17 further progressed the SRS spectral efficiency enhancement for frequency hopping. For instance, partial SRS is defined for each frequency hopping bandwidth, which allows each UE to transmit SRS only on a part of each frequency hopping bandwidth, while the BS could still reconstruct the full bandwidth channel. Since the re-use factor is 4 (Fig. 3 R17 partial SRS), 4 times the number of UEs can be supported relative to Release 15. This facilitates higher-order massive MIMO operation.

### C. mmWave Applications

In addition to the massive MIMO evolution aspects presented for the FR-1 band, a new key technology introduced in 5G-NR is the so-called beam-based operation. This is especially relevant for operation of systems in the FR 2-1 bands, i.e., above 24 GHz center frequency. As is well known from fundamentals of wave propagation [21], [22], higher frequencies suffer from high propagation loss. However, given the shorter wavelength, it is more feasible to pack more elements in a practical form-factor. Larger arrays provide narrower transmission and receptions beamwidths with higher gain to overcome the higher propagation loss.
3GPP Release 15 introduced multi-beam operation, where a beam refers to a spatial domain transmission filter or the spatial domain reception filter. Beam-based operation includes beam acquisition, beam maintenance and tracking, and beam failure recovery. During the initial access phase of a link, a UE identifies a SSB (which includes primary synchronization signal, secondary synchronization signal, and the physical broadcast channel (PBCH)) that is received with a reference signal received power (RSRP) that exceeds a preset threshold. Each SSB is transmitted using a corresponding spatial domain transmission filter. Thus, the SSB selection determines the beam to use for subsequent communication between the UE and BS. The UE selects a physical random-access channel (PRACH) occasion and a PRACH preamble that is associated with the SSB to use for transmission of the preamble, hence implicitly indicating the beam to the BS.

During the beam maintenance phase, for downlink beam indication and measurement, the reference signal can contain non-zero power channel state information reference signal (NZP CSI-RS) and/or SSB. Here, downlink beam indication is performed via TCI states. Each TCI state includes a TCI state identifier (ID) and quasi-co-location (QCL) information. The QCL information includes a QCL type and a source RS. The QCL information associates one signal (e.g., data or control) with another (e.g., a source RS) that shares a same set of channel statistics such as Doppler/delay properties and spatial receive filtering pre-defined by a QCL type. The source RS determines the beam to use. For instance, the UE can be indicated by a scheduling DCI – e.g., via the TCI field – a TCI state comprising a SSB as the QCL source RS with the corresponding QCL type set to ‘D’ – i.e., spatial domain property – for PDSCH reception. Based on the indicated TCI state, the UE can then assume the same spatial domain receive filter as that used for receiving the SSB (e.g., obtained during the initial access phase) to receive the PDSCH. This would allow the UE to transmit to and receive from the base station without explicitly knowing how the “beams” are implemented at the base station.

Due to the fundamentally different nature of system operation relative to FR-1 bands, drastically different CSI acquisition framework is developed and optimized in 3GPP from Release 15 to 17 [4], [6]. This concludes our evolution summary and leads the discussion into the massive MIMO technologies employed in Release 18 (and beyond) study item considerations.

III. MASSIVE MIMO TECHNOLOGIES FOR 3GPP RELEASE 18 (5G-ADVANCED) AND BEYOND

The massive MIMO concepts standardized in 3GPP Release 15, 16 and 17 present an ideal opportunity for further development in Release 18, where the massive MIMO evolution is likely to transition from a network-centric architecture to a UE-centric architecture. In this section, we discuss the key massive MIMO technologies for Release 18 and beyond across both FR-1 and FR-2 bands, including CJT for multi-TRP operation, uplink CSI enhancements, enhancements in the DMRS design, CSI for mobility enhancements, and enhancements on mmWave systems.

Fig. 4. Illustration of CJT for multi-TRP-based systems.

Fig. 5. Illustration of intra-site inter-cell and inter-site inter-cell CJT scenarios.

A. CJT for Multi-TRP for FDD Systems

A typical scenario of CJT by multiple TRPs is illustrated in Fig. 4. There is a CSI measurement TRP set (TRPs within the dashed red lines inside the black circle), and a CJT TRP set (generally same as CSI measurement TRP set or a subset of CSI measurement TRP set). A fundamental problem is the determination of optimal CSI measurement TRP set and CJT TRP set, chosen to satisfy some optimization constraints in a UE-centric manner. It is assumed that sufficient backhaul connectivity is in place for the all TRPs, and the CSI measurement TRP set and CJT TRP set can be selected in a UE-centric fashion. The CSI measurement TRP set is configured by RRC based on the RSRP difference with the serving cell, such that the TRPs with the strongest RSRP are included in the CSI measurement set. Furthermore, each UE needs to measure the CSI of TRPs within the CSI measurement TRP set and report the measurement to the BS. Subsequently, the BS can determine the CJT TRP set for each UE depending on system scheduler and CSI.

The above system can function in several practical deployment scenarios tailored for CJT operation: In the first case, three RRHs are distributed and connected to the BS within a cell as shown in Fig. 6. Compared to the single-TRP layout, the additional three RRHs in each cell are located at the midpoints between the centers of its associated site and neighboring sites, respectively, and the antenna angle of each TRP is facing towards the center region of the associated cell. In the second scenario, inter-cell CJT exists, where multiple cells form a CJT transmission set can be considered. Fig. 5 shows two representative scenarios of intra-site inter-cell and inter-site CJT inter-cell scenarios with three-cell CJT transmission sets.

In such scenarios, to select an optimal TRP coordination set and CSI measurement TRP set, while attempting to achieve...
While CJT brings attractive merits, it also poses higher requirements on the accuracy of CSI obtained on SRS, since the beamformers of the coordinated cells (cells belongs to CJT TRP set) should assure that the phase of received signals from these cells is constructive. In order to perform idealized joint orthogonal SRS resource allocation across the CJT TRPs to avoid the SRS interference, the requirement of orthogonal SRS resources would be multiplied compared with that of single-TRP operation, and the feasibility in UE-centric coordination is also doubtful. In that case, a more practical and effective way to manage SRS interference is to perform SRS resource allocation by serving TRP. Since the inter-cell interference can heavily limits the channel estimation accuracy based on SRS, SRS interference randomization among the TRPs should be performed.

Current interference suppression method is to whiten the colored interference by low correlation sequences during LS operation at BS. However, the sequence correlation would be quite large especially for the case with short sequence length which may incur large interference. For example, the sequence length would be very limited if frequency hopping is performed for coverage enhancement. In CJT scenario, the issue is more serious. As shown in Fig. 6a), the SRS of UE1 to its coordination cell suffers severe interference from a near located UE2 since the RSRP of UE2’s SRS can be much higher than that of UE1. Figure 8b) shows the issue in a system-level simulation. 4 dB loss of normalized mean squared error (NMSE) can be seen for the coordinated cell in CJT compared with the single-cell operation, which means the channel coefficient derived by the coordinated cell cannot be as accurate as that in single-TRP operation and CJT performance in that case cannot be maintained.

The NMSE for each SRS channel estimation used in this section is defined as:

\[
\text{NMSE}_{p} = \frac{\sum_{k,s} \left| \hat{H}_{p,k,s} - H_{p,k,s} \right|^2}{\sum_{k,s} \left| H_{p,k,s} \right|^2},
\]

where \( \hat{H}_{p,k,s} \) and \( H_{p,k,s} \) are the estimated channel coefficient and ideal channel coefficient corresponding to SRS port \( p \), TRP receiving antenna \( k \) and subcarrier \( s \). In the simulation, 4 UEs per cell is assumed and totally 21 cells are setup. One 4-port SRS (for sounding 4 receive radio chains) is transmitted by each UE for all of its coordinated cells. Each UE has one serving cell and multiple coordinated cells, the coordinated cells are determined to satisfy that the RSRP gap between the serving cell and the coordinated cell should be no more than 10 dB. The physical resource of the SRS is allocated by each serving cell. For example, orthogonal physical resources and same root sequence are used for each UE in the same serving cell. All cells share the same set of SRS resources, and different cells would use different root sequences. The blue line in Fig. 6b) collects the MSE from all of the serving cells via SRS channel estimation, and the red line collects the MSE from all of the coordinated cells. The yellow line assumes that no interference but only the noise is suffered by the coordinated cells. To resolve the issue, especially for the case with limited SRS resource in the system, SRS
interference randomization should be introduced to improve the SRS channel estimation performance under the scenario with strong interference.

Assuming that one target UE and one strong interference UE transmit SRS to a TRP on the same physical resource, the received SRS signal model at the $n$-th transmission in the point of view of the TRP can be written as

$$y_n (m) = r_n (m) h_n (m) + r'_n (m) h'_n (m) + w_n (m) = e^{j \alpha_n m} \bar{r} (m) h_n (m) + e^{j \alpha'_n m} \bar{r}' (m) h'_n (m) + w_n (m),$$

(5)

where $(m)$ denotes the $m$-th element of a sequence mapping to different subcarriers in frequency domain, and $m = 0, 1, \ldots, M$, where the $M$ is the number of subcarriers occupied by the SRS. $y_n$ and $w_n$ are the received SRS sequence and the noise. $r_n$ and $r'_n$ are the SRS sequence of the target UE and that of the interference UE. $h_n$ and $h'_n$ are the target channel and the interference channel coefficient in frequency domain. $\bar{r}$ and $\bar{r}'$ are the sequence used by target UE and the interference UE. $\alpha_n$ and $\alpha'_n$ are the cyclic shift (CS) values for the two UEs. The CS value can be set between 0 to $2\pi$. By using different CS values in the same physical resource elements, the two SRSs can be orthogonal.

To estimate the target channel, $y_n$ is multiplied by the conjugate of $r_n$, then we have

$$\bar{y}_n (m) = y_n (m) r^*_n (m) = h_n (m) + r'_n (m) r^*_n (m) h'_n (m) + r^*_n (m) w_n (m).$$

(6)

Note that $r^*_n (m) r'_n (m) h'_n (m) = e^{j (\alpha'_n - \alpha_n) m} \bar{r} (m) \bar{r}' (m) h'_n (m)$.

Then we transform the above signal to the delay domain, giving

$$\bar{y}_n^D (\tau) = \text{IFFT} (\bar{y}_n) = h_n^D (\tau) + I_n^D (\tau) + w_n^D (\tau),$$

(7)

where $h_n^D (\tau), I_n^D (\tau)$, and $w_n^D (\tau)$ are the target channel, interference and noise in the delay domain, respectively. Furthermore, we denote $I_n^D (\tau) = \text{IFFT} (e^{j \alpha_n \tau} \bar{r}' h'_n)$ to be the delay-domain interference without adding and removing CS. Note that, the dimensions of $\bar{r}, \bar{r}'$ and $h'_n$ are all $M \times 1$. Obviously, $I_n^D (\tau)$ is not whitened. Since a frequency domain cyclic phase shift is a delay shift in the delay domain,

$$I_n^D (\tau) = I_n^D (\tau + \frac{\tau_{\max}}{\alpha_{\max}} (\alpha'_n - \alpha_n))$$

(8)

where $\tau_{\max}$ is the pre * (defined maximum CS value for a SRS RE, $\alpha_{\max}$ is the CS value used for SRS of each UE. By time filtering across multiple SRS transmissions (here we take time-average as an example), we calculate the statistical power-delay profile (PDP) of the received signal

$$\text{PDP}_{\bar{y}} (\tau) = \frac{1}{N} \sum_{n=1}^{N} |\bar{y}_n^D (\tau)|^2.$$  

(9)

For a large $N$, assuming that the target signal and the interference are independent,

$$\text{PDP}_{\bar{y}} (\tau) = \frac{1}{N} \sum_{n=1}^{N} |h_n^D (\tau)|^2 + \frac{1}{N} \sum_{n=1}^{N} |I_n^D (\tau)|^2 + \sigma_n^2 + o(1)$$

(10)

where $\sigma_n^2$ is the noise variance, $O(1) \rightarrow 0$ when $N \rightarrow \infty$. Due to the fact that the channel consists of only a few paths and the path delays vary slowly, the power of PDP$_h$ $(\tau)$ concentrates on only a few strong delay taps. If the power of PDP$_I$ $(\tau)$ spreads over the entire delay domain, the interference and noise and be reduced by finding the strong delay taps of PDP$_{\bar{y}} (\tau)$. However, PDP$_I$ $(\tau)$ may also have strong delay taps in practice, which causes severe channel estimation error.

To solve the problem, CS hopping is needed to whiten PDP$_I$ $(\tau)$.

Without CS hopping, $\alpha'_n - \alpha_n$ is constant among different SRS transmissions. Then

$$\text{PDP}_I (\tau) = \frac{1}{N} \sum_{n=1}^{N} |I_n^D (\tau + \text{constant})|^2.$$  

(11)

PDP$_I$ $(\tau)$ varies with $\tau$, so it is not white in the delay domain. By contrast, by adopting CS hopping, since $(\alpha'_n - \alpha_n)$ mod $\alpha_{\max}$ is uniformly distributed in $[0, \alpha_{\max})$, for example, if infinity transmission times is assumed, each tap in PDP has equal probability to be occupied, in that case, PDP$_I$ $(\tau)$ can be whitened when $N$ is large:

$$\text{PDP}_I (\tau) = \frac{1}{N} \sum_{n=1}^{N} |I_n^D (\tau + \frac{\tau_{\max}}{\alpha_{\max}} (\alpha'_n - \alpha_n))|^2.$$
\[
\frac{1}{\alpha_{\text{max}}} \int_{0}^{\tau_{\text{max}}} \left( \int_{D} \left( \tau + \frac{\tau_{\text{max}}}{\alpha_{\text{max}}} \alpha \right)^2 d\alpha \right) \bar{I}_D(\tau) d\tau.
\]

We can see that PDP_I(\tau) is independent to \tau, which means that PDP_I(\tau) is whitened by adopting CS hopping. Therefore, the strong taps of the target signal can be found correctly with the time filtering. For example, the taps with larger power than the power of whitened interference taps added by noise power can be easily found. The performance of proposed CS hopping is presented in Fig. 9. The sequence length used by SRS is 288 (assuming that the bandwidth of SRS is 48RBs with 6 subcarriers per RB), and the power of interference is 6dB larger than the target signal. In Fig. 7a), we sum the error of estimated tap compared with the ideal tap for each channel estimation. Both the CDF of error of the strongest tap and the first three strongest taps are presented. It can be observed that via CS hopping, the taps of target signal can be found more accurately due to the interference is whitened. According to the MSE of channel estimation in Fig. 7b), we can further see that channel estimation can be performed more accurately compared with the legacy case.

In TDD, thanks to the channel reciprocity, the uplink channel estimation can be used to infer downlink channels. This feature enables BS to reduce the feedback overhead. However, due to RF mismatch in the uplink and downlink paths (at receivers and transmitters), utilizing the uplink channels for downlink channels requires periodic calibration among receive and transmit antennas of RF networks at the BS. In general, a BS has an on-board calibration mechanism in its own RF network to calibrate its antenna panels with multiple receiver and transmitter antennas. The on-board calibration mechanism can be performed via small-power reference signal transmission and reception from/to the RF antenna network of base station in order to measure the gain and phase differences among transceivers in the same RF unit. Thus, it can be done by hardware implementation in a self-contained manner. However, it can be difficult to perform such an on-board calibration with non-collocated TRPs. Instead of on-board calibration, an over-the-air signaling mechanism to calibrate receive and transmit antennas among non-co-located TRPs is instrumental. In this regard, an RS-based mechanism for a UE-assisted calibration can be used to report calibration coefficients across TRPs.

### B. Enhancements on Uplink MIMO

The allowed maximum transmission layers are essential for the uplink spectrum efficiency. With the emergence of advanced UEs types such as hubs for FWA, router-like devices for customer premise equipment (CPE), industrial robotics, and automobiles, larger form factors become feasible for UE antenna architectures. This facilitates an increased number of transmission layers, translating to higher peak data rates. In 5G-NR, DMRS is used for channel estimation and demodulation of physical channels [13]. Two types of DMRS are standardized in 3GPP for NR use till Release 17, namely Type-I DMRS and Type-II DMRS, as shown in Fig. 8. Type-I DMRS is denser in the frequency domain compared to Type-II, and Type-I and Type-II DMRS support up to 8 and 12 orthogonal DMRS ports. However, in 3GPP Release 18 (5G-Advanced) and future releases, more receive antennas would be employed at each BS and joint multi-TRP reception would be enabled with more advanced network deployments. To this end, more orthogonal DMRS ports are needed for more uplink transmission layers. Increasing the DMRS ports without increasing DMRS transmission overhead is a major challenge while considering DMRS design.
space-frequency form with dimension of \( \mathcal{H}(s, f, t) \) means the union channel matrix in 
\( \mathcal{X}^L_\alpha \), can be transformed into angular-delay 
Fig. 9. Multiplexing DMRS ports for different UEs in the delay domain.

To increase the number of DMRS ports, multiplexing DMRS in the delay domain is a promising solution. The data channel for each UE may have considerable delay spread variability, and DMRS from multiple UEs can be transmitted in the same time-frequency resource by moving the signal in the delay domain, as shown in Fig. 9. The DMRS ports from different UE are transmitted in the same time-frequency resource, while can be distinguished in the delay domain. Shift in the delay domain is equivalent to multiplexing DFT code in the frequency domain. To extend the number of orthogonal DMRS ports from 12 to 24, a simple scheme is doubling the length of orthogonal cover code (OCC) in 5G from 2 to 4.

C. CSI Enhancements for Mobility

In existing CSI measurement and feedback procedures till 3GPP Release 17, the BS generally assumes that the channel condition remains unchanged within the CSI reporting period, and the BS uses the latest reported channel CSI for downlink transmission beamforming. However, the performance in scenarios involving mobility is greatly degraded due to more rapid CSI expiration. Here, the channel may vary rapidly, and the CSI expiration significantly deteriorates the system performance. In addition, this occurs worse in multi-UE scenarios, as the CSI mismatches will introduce the inter-UE interference when multi-UE pairing is performed by the system scheduler.

The above motivates new solutions for CSI enhancements with mobility in 3GPP Release 18. In massive MIMO transmission, a channel consists of multiple paths, and the time-varying characteristic of a single path can be uniquely described by the exponential function proportional to its Doppler frequency:

\[
\mathbf{H}(s, f, t) = \sum_{i=1}^{L} \alpha_i e^{j2\pi v_i t} \mathbf{\theta}_i \otimes \tau_i, \quad (13)
\]

where \( \mathbf{H}(s, f, t) \) means the union channel matrix in space-frequency form with dimension of \( P \times N_f \) for slot \( t \). Note that \( \alpha_i e^{j2\pi v_i t} \) denotes the channel coefficient for \( i \)-th path, where \( \alpha_i \) is irrelevant to mobility and \( v_i \) denotes the equivalent Doppler shift for \( i \)-th path. Furthermore, \( \mathbf{\theta}_i (P \times 1) \) and \( \tau_i (1 \times N_f) \) are angle and delay steering vectors for \( i \)-th path, respectively. The time-frequency channel matrix, \( \mathbf{H}(s, f, t) \) at slot \( t \), can be transformed into angle-delay domain by 2-Dimension IDFT, shown as yellow rectangular in Fig. 10.

\[
\mathbf{C} = \mathbf{W}_f^H \mathbf{H}(s, f, t) \mathbf{W}_f, \quad (14)
\]

where \( \mathbf{C} \) is angle-delay domain complex coefficient matrix with each coefficient (each red box in Fig. 10) corresponds to one angle-delay pair. \( \mathbf{W}_f (P \times P) \) and \( \mathbf{W}_D (N_f \times N_f) \) are spatial and frequency basis respectively. The Doppler information can be extracted from the complex coefficients of continuous slots for each angle-delay pair, as shown in Fig. 10.

As the Doppler frequency does not change rapidly in short time duration, CSI reporting in the mobility scenarios can be enhanced by predicting the time-varying channel based on the Doppler frequency of each path. The procedure for CSI enhancement in mobility case can be summarized as 3 steps.

\begin{itemize}
  \item **Step 1.** BS sends \textit{continuous N CSI-RS} to UE:
    \begin{itemize}
      \item The time gap between two adjacent CSI-RS transmissions can be denoted by \( \Delta t \), which is determined by Doppler spread. The whole CSI-RS occupies \( N \Delta t \), which is determined by the required Doppler resolution.
    \end{itemize}
  \item **Step 2.** Doppler extraction and CSI prediction:
    \begin{itemize}
      \item Step 2.1: \( \alpha_i \) and the Doppler information \( v_i \) \((i = 1, 2, \ldots, L)\) can be extracted by \( N \) union channel matrices in space-frequency form estimated from \( N \) CSI-RSs in Step 1.
      \item Step 2.2: UE can predict the CSI information in future \( M \) slots based on \( e^{j2\pi v_i t} \) and consistent \( \alpha_i \). The \( \mathbf{H}(s, f, t) \) can be predicted by \( \alpha_i \) and \( e^{j2\pi v_i t} \) with new \( t \) value in future \( M \) slots.
    \end{itemize}
  \item **Step 3.** UE performs CSI compression on the predicted CSI via Doppler domain.
    In Step 3 above, for compressing the predicted CSI \((P \times N_f)\) in future \( N_{\text{slot}} \) slots, the codebook for CSI reporting can be expressed as follows considering CSI compression across spatial domain, frequency domain, and time domain:
    \[
    \mathbf{W} = \mathbf{W}_1 \mathbf{W}_2 (\mathbf{W}_f \otimes \mathbf{W}_D)^H, \quad (15)
    \]
    where \( \mathbf{W} \) means the predicted CSI or precoder in future \( N_{\text{slot}} \) slots with dimension of \( P \times (N_f \times N_{\text{slot}}) \), \( \mathbf{W}_1, \mathbf{W}_f \) and \( \mathbf{W}_D \) means the spatial domain basis, frequency domain basis and
time domain basis for CSI compression. \( W_1 \) and \( W_f \) reuse Release 16 or Release 17 design. \( W_1 \) is a \( P \times 2L \) matrix consisting of \( 2L \) DFT based spatial bases, \( W_f \) is an \( N_f \times M \) matrix composed of \( M \) DFT based frequency bases. \( W_D \) is an \( N_{\text{slot}} \times T \) matrix composed of \( T \) DFT – based time bases and \( W_2 \) is the space- frequency-time combination complex coefficients with dimension \( 2L \times (M \cdot T) \). Only a part of coefficients with stronger power are selected from all \( 2L \times (M \cdot T) \) coefficients for reporting.

\section*{D. Enhancements on mmWave/ FR 2-1 Bands}

In 3GPP Release 15 and 16, a common framework is shared for CSI acquisition (FR-1) and beam management (FR 2-1). While the complexity of such a framework is justified for CSI in FR-1, it makes beam management procedures less efficient in FR 2-1. Furthermore, as aforementioned, the beam management procedures can be different for different channels. Having different beam indication/update mechanisms increases the complexity, overhead, and latency of beam management. Such drawbacks are especially troublesome for high mobility scenarios (such as highway/vehicular use cases at FR 2-1) and/or scenarios requiring large number of configured TCI states. A TCI state here is used for beam indication, providing an association between a TCI state ID and a source reference signal associated with a beam. These drawbacks motivated a streamlined beam management framework for multi-beam operations and procedures that is common for data and control, and uplink and downlink channels. This framework, is referred to as the unified TCI framework, first introduced in Release 17 for single TRP operation and is being enhanced in Release 18 for multi-panel /multi-TRP operation.

The unified TCI framework supports signaling a unified TCI state to the UE, where a unified TCI state is, as illustrated in Fig 16: a downlink or a Joint TCI state, where a downlink TCI state is applied to downlink channels and signals, and a joint TCI state is applied to downlink channels and signals, and/or an uplink or a joint TCI state, where an uplink TCI state is applied to uplink channels and signals. The unified TCI state is applied for at least UE dedicated-channels, which are channels that are transmitted to a single-UE or channels transmitted from a UE. The unified TCI state framework is designed to support downlink receptions and uplink transmissions in the UE with a joint (common) beam indication for downlink and uplink by leveraging beam correspondence. The unified TCI state framework also supports downlink receptions and uplink transmissions in the UE with separate beam indications for downlink and uplink, for example to mitigate maximum permissible exposure.

The network configures the UE up to two lists of a TCI states. The first list of TCI states is used for downlink and joint beam indication, and the second list of TCI states is used for uplink beam indication. In addition to including spatial relation information, joint or UL TCI states also include power control information and path-loss reference signal associated with the TCI state.

The network activates a set of up to 8 TCI state code points by MAC-CE signaling. A TCI state code point can be (1) a downlink TCI state; (2) an uplink TCI state; (3) a joint TCI state; or (4) a pair of downlink TCI state and uplink TCI state. A TCI state code point is signaled to the UE in a downlink-related DCI Format (e.g., DCI Format 1_1 or DCI Format 1_2) using the “transmission configuration indication” field. The downlink-related DCI format may or may not include a downlink assignment. The latter is a DCI Format used specifically for beam indication. As illustrated in Fig. 12, the TCI state signaled by the MAC-CE is applied after a beam application time from the end of uplink channel conveying hybrid automatic repeat request acknowledgement (HARQ-ACK) feedback associated with the DCI format conveying the TCI state.

The multi-TRP transmission holds promise for improving link reliability (by adding redundant communications paths for spatial diversity), system capacity (by aggregating resources from physically non-collocated TRPs), and coverage. Operating at mmWave frequencies can exploit the full potential of multi-TRP transmission; this, however, would require highly efficient beam management procedures. In Release 17, various design aspects related to beam management including beam measurement/reporting and beam failure recover (BFR) were specified for the multi-TRP operation. The corresponding TCI framework, however, still based on Release 15/16, is inefficient and complex especially when more than one TCI states (or TRPs) are indicated. Release 18 enhances and extends the unified TCI framework to multi-TRP transmission. In the Release 17 unified TCI framework, a TCI codepoint corresponding to a single TCI state or pair of DL and UL TCI states is indicated by the DCI field "Transmission Configuration Indication" in the corresponding DCI Format. To extend the framework to multi-TRP operation, a TCI codepoint is enhanced to correspond to two joint TCI states, or two pairs of DL and UL TCI states for two TRPs.
IV. EVALUATION RESULTS FOR MASSIVE MIMO TECHNOLOGIES IN 3GPP RELEASE 18 AND BEYOND

In the section, we provide system-level simulation results for the key enhancements discussed in the earlier section and the evolutions discussed in previous sections. In addition, we present results from a real-world trial on multi-TRP CJT in single and multi-UE scenarios.

A. Simulation Results for Aspects of Massive MIMO

1) Performance Evaluation of CJT: Using the parameters given in Table II in the Appendix section of the paper, Figure 13 and 14 depicts SLS results on the performance gains of CJT over single-TRP transmission with Release 17 Type-II enhanced codebook. For each TRP, the overhead for the proposed CJT codebook is the same as Release 17 Type-II codebook overhead. A total of 57 sectors (19 sites) are modelled and it is assumed that 30 UEs are randomly dropped in each sector. A non-full buffer model, FTP traffic model 1 with packet size 0.5 Mb [21], is assumed and arrival rate $\lambda$ is adjusted for a given range of resource utilization (RU) 70%. Each TRP and UE is assumed to be equipped with 32 and 4 antenna elements respectively. Up to 3 cooperated TRPs for CJT TRP set is assumed. Other simulation assumptions are described in Table II and the simulation procedure followed is exemplified by Fig. 26 in the Appendix section. For comparing performance of multi-TRP CJT transmission vs single-TRP transmission, user perceived throughput (UPT) is used as a performance metric. The UPT for a UE is defined as

\[
r = \frac{\sum_{i=1}^{n} S_i}{\sum_{i=1}^{n} T_i}, \quad (16)
\]

where $S_i$ is the size of $i$-th data burst transmitted to the UE and $T_i$ is the time between the arrival of the first packet of the $i$-th data burst by a given data-traffic model and the reception of the last packet of the $i$-th data burst at the UE [21].

As shown in Figure 13, the CJT from multi-TRP yields significant performance improvements over single-TRP transmission, which are more than 50% in terms of mean UPT. In addition to the performance gain of CJT over Rel-16/17 codebook, SINR comparison of inter-site CJT and single-TRP transmission are provided in Fig. 14. As shown in Fig. 14, compared with S-TRP, significant SINR improvement (about 5 dB) can be achieved by CJT with CSI enhancement. The performance gain basically comes from two aspects. One aspect is the performance gain achieved by CJT compared with no cooperation among multi-TRP, which is due to inter-cell coherent transmission and inter-cell interference suppression. Another aspect is the CJT codebook enhancement to match with the channel property of multi-TRP joint transmission.

2) Performance Evaluation of Uplink DMRS Enhancement: Uplink system-level simulation results for different DMRS design is performed using the SLS conditions in Table III in the Appendix, where coherent joint reception (CJR) across all TRPs is assumed. 8R for each TRP and 4T for each UE are assumed. The baseline is the type II DMRS in 5G with up to 12 orthogonal DMRS ports allowed. The DMRS scheme in 5G-A employs the proposed extending OCC length scheme for DMRS enhancement. The throughput under 24 layers in 3GPP Release 18 are $1.75 \times$ compared with the 12 layers specification limitation in 5G. With CJR among all TRPs, uplink multi-user interference can be mitigated, and more than 20 layers can be paired simultaneously.

3) Performance Evaluation of Mobility Enhancements: Fig. 15 shows throughput simulation results for CSI prediction with proposed CSI enhancement for mobility. The baseline is conventional CSI reporting with Release 16 Type-II. The new codebook structure with time domain basis is used for the proposed scheme for CSI compression on the predicted CSI/precoder in future slots. Based on simulation results, it can be observed that CSI prediction based on Release 16 type II codebook can achieve about 15% performance gain compared with Release 16 Type II conventional.

B. Preliminary Field Test Results

Preliminary field test results for coherent joint transmission with multi-TRP are provided and discussed in this section, including single-user (SU) field test and multi-UE field test., where the detailed parameter is shown in Table I. In the field testing, interference cells are also assumed and only two TRPs for coherent joint transmission are used.
Fig. 15. Throughput (Mbps) of UE-side prediction with Release 16 conventional and Doppler-based codebooks.

TABLE I
FIELD TRIAL SPECIFICATIONS AND PARAMETERS

| Parameter               | Value                                      |
|-------------------------|--------------------------------------------|
| Location                | Shanghai Jiao Tong University             |
| Environment             | Urban outdoor                              |
| Duplex Mode             | TDD                                        |
| BS Configuration and Transmit Power | 64 transmit (T) 64 receive (R), 46 dBm        |
| UE Configuration        | 2 transmit (T) 4 receive (R) with SRS-based antenna switching (Huawei Mate30) |
| Center Frequency        | 3.5GHz                                     |
| Bandwidth               | 20MHz                                      |
| Transmission TRP Set    | 2 TRPs with intra-site (Normal direction of the antenna as shown in yellow arrow) |
| Interfering TRP Set     | 4 TRPs with inter-site and intra-site (Normal direction of the antenna as shown in red arrow) |

The SU field test scenario is shown in the Fig. 16, where the drive route line is shown in white solid line. The UE velocity is 5km/h and the interfering sites are shown with white dotted lines. Along with drive route, spectrum efficiency (transmission rate vs. bandwidth) for CJT and single TRP transmission are tested and collected.

Fig. 17 shows the CDFs of spectrum efficiency comparison of CJT and single TRP transmission along with drive route line. CJT achieve consistent UE experience with almost same spectrum efficiency along the drive route. Compared with single TRP transmission, more than 30% performance gain for cell edge (5%) could be obtained by CJT.

We also test the multi-UE performance with CJT, and the multi-UE test scenario is shown in Fig 18. To verify the performance gain for different scenarios, 4 test cases are considered as shown in Fig. 19, where different UE locations with different RSRP gap are considered.

As shown in Fig. 19, all the candidate UE locations are divided into 2 groups of which one is marked with red color, and the other is marked with blue color. During MU field test, one red UE location and one blue UE location are chosen for 2-UE test. For each group, four regions are defined. Region 1 means UE received RSRP difference from different TRPs is less than 3 dB. Region 2 means the RSRP difference is between 3 dB and 10 dB. Region 3 is with the RSRP difference between 10 dB and 15 dB. Region 4 means UE received RSRP difference from different TRP is more than 15 dB. It can
readily be predicted that the beneficial region is the region with small RSRP gap, where each TRP would have a big impact on the received signal at UE side.

The performance gain of CJT with multi-UE pairing over single TRP transmission for the 4 different cases is shown in Fig. 20. It is observed up to 98% gain can be achieved by CJT with 2 TRPs with MU. From Fig. 20, it also can be verified that UEs in the small RSRP difference region obtain more performance gain with CJT. With CJT, large beamforming gain can be achieved, and MU interference can be suppressed significantly.

The performance gain mainly comes from two aspects:

✓ Larger antenna ports and larger beamforming gain: antenna ports among multiple TRPs are combined in CJT case and equivalent antenna array is enlarged. Thus, CJT would be beneficial for MU pairing in high load and SU high order transmission in low load.

✓ Higher capability on interference suppression: joint scheduling and transmission among multiple TRPs can be performed in CJT case, and the inter-cell interference can be controlled and suppressed significantly. Thus, SINR during data transmission for the UEs would be enlarged by CJT significantly, especially in high load case.

It should be noted that time-frequency calibration and clock synchronization is assumed in simulation and well performed for field test among multiple TRPs for CJT operation in this paper. In practical system, time-frequency calibration and clock synchronization may be loosed due to implementation complexity and cost, thus the performance gain by CJT may be reduced.

C. Simulation Results for mmWave/ FR 2-1 Bands

In this section, the simulation results for the 3GPP Release 17 beam management enhancements are presented and analyzed. As described earlier, the motivation of the unified TCI framework introduced in 3GPP Release 17 for beam management is to reduce the latency and the overhead of beam indication, thereby enhancing system performance especially in high-speed scenarios. To quantify these benefits, we simulated and analyzed two scenarios, first is a high-speed train (HST) scenario, illustrated in Fig. 21 the details of which are described in [6], second is a dense urban highway (DUH) scenario illustrated in Fig. 22. In the highway scenario, the UE is moving between point $P$ and point $Q$. The UE is moving at a speed of 120 kmph (33.3 m/sec), UE’s throughput is sampled every 1 m (30 ms), with 100 sample points i.e., simulation duration is 3 seconds.

Performance is evaluated for beam indication using DCI, with a latency processing time of 0.5 ms and a BLER of 1%, and using MAC-CE, with a latency processing time of 3 ms and a BLER of 10%. DCI processing happens within the physical layer, while the MAC-CE processing happens in the MAC layer requiring additionally processing steps to transfer the received data from the physical layer to the MAC layer and back to the physical layer to apply the indicated beam and hence the longer processing time with MAC-CE based beam indication. For DCI-based beam indication, after the PDCCH is received the beam in updated after two slots at 60 kHz (i.e., 0.5 ms), this is to account for the modem processing timeline and the fact that beams are updated at slot boundaries. For MAC-CE based beam indication, the MAC-CE is received in a slot (PDCCH + PDSCH), the data is transferred from
the physical layer to the MAC layer where the MAC-CE is parsed and the beam indication is sent back to the physical layer to apply, this processing is assumed to take 3 ms. In the simulations it is assumed that the beam is applied without consideration to the HARQ-ACK timing corresponding to the beam indication channel.

For the HST scenario, Fig. 23 illustrates the benefit of using DCI with the unified TCI state framework over legacy (5G-NR Release 16) MAC-CE for beam indication. As can be seen from the figure, on average, DCI using the unified TCI state framework provides a 10% spectral efficiency (SE) gain over MAC-CE. The benefit using DCI for beam indication with low latency is most apparent as the UE moves close to the RRH, when the rate of angular change and hence beam change is fastest as the UE moves past the nearest point to the gNB (BS).

In the 3GPP assumptions, the MAC-CE processing time is the sum of the PDSCH processing time, the HARQ-ACK transmission time and time between the end of HARQ-ACK and the application of the indicated TCI state. The total MAC-CE processing time is greater than 3 ms, hence the relative performance of MAC-CE based beam indication to DCI based beam indication would be even worse than what are shown in the above figures. Though simulation settings and results throughout this section are for single-TRP based framework – i.e., single TCI state indication, they are also applicable to multi-TRP based operation – i.e., multi-TCI state indication. This is because the performance metrics considered here – i.e., both the latency and overhead reductions – are brought by the beam indication signaling medium itself, not the difference between deploying a single TRP and multiple TRPs.

V. CONCLUSION

In this paper, we provided our considerations for massive MIMO Evolution towards 5G-Advanced in Release 18. To begin with, we present a comprehensive review of massive MIMO evolution for TDD and FDD from earlier releases, i.e. from Releases 15 to 17. Based on a commercial massive MIMO architecture, fundamental characteristics like beamforming gain, beamwidth and angular coverage are achieved by a structure of antenna sub-arrays. CSI reporting and acquisition framework specified in earlier releases were further elaborated in the paper. Type I single panel and multiple panel codebooks can provide low resolution CSI quantization for SU-MIMO. On the other hand, Type II codebook can provide high resolution CSI quantization typically used for MU-MIMO and was enhanced in multiple releases by taking advantage of frequency coherency or channel reciprocity to compress codebook parameter like combination coefficients. We explained fundamental SRS mechanisms like antenna switching and frequency hopping, considering implementation constraints of UE RF and power. For FR2 scenario, mmWave application has motivated multi-beam operation and beam maintenance via managing and indicating TCI state or SRS ID.
Starting from above overview and moving forward toward Release 18, we presented our considerations to address critical issues of NR system design:

- To support FDD CJT for multi-TRPs, we presented new codebook structure to support coherent reception at UE with precise CSI reporting and acquisition across TRPs. By taking advantage of both spatial and frequency correlation as Type II codebook, channel information from multi-TRPs can be compressed and reported jointly.

- To support TDD CJT for multi-TRPs, inter-cell SRS interference is problematic when using SRS for downlink channel estimation. We demonstrated that with random CS hopping applied to SRS sequence transmission, inter-cell interference can be averaged and whitened so that targeted SRS sequence and associated MIMO channel can be estimated more precisely with time filtering.

- To support uplink massive MIMO for higher order uplink MU-MIMO, we presented a design of increasing the number of orthogonal DMRS ports by simply doubling OCC length.

- To support better CSI acquisition in mobility, we presented a new codebook structure to allow the UE to perform CSI prediction based on historical CSI measurements. Assuming that Doppler frequency can be static within a time window, it is feasible to compress the latest and predicted CSI in a CSI reporting to reduce reporting overhead and achieve reasonable accuracy simultaneously.

- To support mmWave application for multi-TRPs, we proposed to extend unified TCI state framework to support mapping more TCI states to a TCI codepoint in DCI format to reduce beam switch latency.

---

**TABLE II**

**SLS ASSUMPTIONS FOR CSI ENHANCEMENTS**

The terminology of all parameters is consistent with 3GPP specifications and can be found from [6], [8], [9], [10], [11], [12], [13], [14], [15], [19], [23] as well as the 3GPP references quoted within the table.

| Simulation Parameter | Value |
|----------------------|-------|
| **Scenario**         | Urban Macro |
| **Frequency Range**  | Intra-cell CJT scenario: |
|                      | - FDD: 2.1GHz, SCS=15kHz |
|                      | - TDD: 3.5GHz, SCS=30kHz |
| **Channel model**    | From 3GPP TR 38.901 |
| **Antenna setup and port layouts at BS** | Intra-cell CJT scenario: |
|                      | - FDD: 32 ports: (M, N, P, Mg, Ng; Mp, Np) = (8,8,2,1,1,2,8), |
|                      | - TDD: 64 ports, (M, N, P, Mg, Ng; Mp, Np) = (8,4,2,1,1,4,8). |
|                      | (dH, dV) = (0.5, 0.8) |

| **Antenna setup and port layouts at UE** | 4RX: (M, N, P, Mg, Ng; Mp, Np) |
|                                          | = (1,2,2,1,1,1,1), (dH,dV) = (0.5, 0.5) |
| **BS power and height**                  | 46dBm and 25m |
| **UE antenna mode**                      | From 3GPP TR 36.873 |
| **Traffic mode**                         | Non-full buffer mode |
| **Modulation**                           | Up to 256QAM |
| **Channel estimation**                   | Non-ideal channel estimation with LS operations |

SLS are demonstrated to support above physical layer optimization. It is clear that CJT is essential for 5G-advanced as it
TABLE III
SLS Assumptions for Uplink MIMO. The Terminology of all Parameters Is Consistent With 3GPP Specifications and Can Be Found From [6], [8], [9], [10], [11], [12], [13], [14], [15], [19], [23]

| Simulation Parameter | Value |
|----------------------|-------|
| Center Frequency     | 2.6 GHz |
| Scenario             | IOT, TRP #18, UE #6 per TRP |
| TRP antenna configuration | (M, N, P, Mg, Ng; Mp, Np) = (1, 2, 1, 1, 1; 2), (dH, dV) = (0.5, 0.8) |
| UE antenna configuration | (1, 1, 2, 1, 1; 1), (dH, dV) = (0.5, 0.5) |
| DMRS                  | Type 2 DMRS, double-symbol |
| Channel estimation   | Non-ideal channel estimation with LS operations |
| Traffic model        | Full buffer |
| UE distribution      | 100% indoor |

can provide profound performance gain with better codebook design and SRS CS hopping. System throughput for uplink massive MIMO with more orthogonal uplink DMRS ports can be roughly doubled. CSI prediction and associated CSI compression at the UE can compensate inaccurate CSI for mobility. For FR2 and high mobility, unified TCI framework can provide fast beam switch by DCI so as to throughput gain, especially around those nearest points to the BS, for optimal beam tracking in mobility. Last but not least, we presented TDD field measurements to demonstrate the feasibility and performance gain of CJT. Both SU-MIMO and MU-MIMO were tested in the field. Compared to single TRP based transmission, CJT can significantly improve UE experience at cell edge to achieve ubiquitous spectrum efficiency across network. For MU-MIMO in CJT, different MU pairing strategies were also compared. Whilst 3GPP Release 18 has just begun as a Work Item, we believe this paper can provide useful and practical optimization techniques for the deployment of massive MIMO, which will be the foundation of future 3GPP RAN standards.

APPENDIX
See Fig. 26, Tables II, and III.

REFERENCES
[1] J. G. Andrews et al., “What will 5G be?” IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
[2] M. Shafi et al., “5G: A tutorial overview of standards, trials, challenges, deployment, and practice,” IEEE J. Sel. Areas Commun., vol. 35, no. 6, pp. 1201–1221, Jun. 2017.
[3] S. Parkvall, E. Dahlman, A. Furuskar, and M. Frenne, “NR: The new 5G radio access technology,” IEEE Commun. Standards Mag., vol. 1, no. 4, pp. 24–30, Dec. 2017.
[4] V. Ramiredy, M. Grossmann, M. Landmann, and G. D. Galdo, “Enhancements on type-II 5G new radio codebooks for UE mobility scenarios,” IEEE Commun. Standards Mag., vol. 6, no. 1, pp. 35–40, Mar. 2022.
[5] Physical Layer Procedures for Data in 5G-NR, document TS 38.214, V.16.6.0, 3GPP, Jun. 2021.
[6] E. Onggosanusi, M. Zhang, Y. Zhang, and J. Kang, “Enhancing 5G MIMO core spectral efficiency with higher-resolution multi-user MIMO and multi-beam operation,” IEEE Commun. Standards Mag., vol. 6, no. 1, pp. 20–26, Mar. 2022.
[7] H. Tatari, M. Shafi, A. F. Molisch, M. Dohler, H. Sjoland, and F. Tufvesson, “6G wireless systems: Vision, requirements, challenges, insights, and opportunities,” Proc. IEEE, vol. 109, no. 7, pp. 1166–1199, Jul. 2021.
[8] Study on Scenarios and Requirements for Next Generation Access Technologies (Release 16), document TR 38.913, V.16.0.0, 3GPP, Jul. 2020.
[9] CSI Enhancements for MU-MIMO, document R1-1909536, 3GPP RAN1, Samsung, Aug. 2019.
[10] Further Enhancements on MIMO for NR, document RP-202024, Samsung 3GPP RAN1, Feb. 2020.
[11] E. Onggosanusi et al., “Modular and high-resolution channel state information and beam management for 5G new radio,” IEEE Commun. Mag., vol. 56, no. 3, pp. 48–55, Mar. 2018.
[12] Enhancements on Multi-Beam Operation, document R1-2007763, 3GPP RAN1, ZTE, Oct. 2020.
[13] E. Dahlman, S. Parkvall, and J. Skold, 5G NR—The Next Generation Wireless Access Technology, Oxford, U.K.: Academic, 2018.
[14] Discussion on CSI Enhancements for Release 17, document R1-2007592, 3GPP RAN1, Huawei, HiSilicon, Oct. 2020.
[15] M. Khoshnevisan, K. Jayasinghe, R. Chen, A. Davydov, and L. Guo, “Enhanced reliability and capacity with multi-TRP transmission,” IEEE Commun. Standards Mag., vol. 6, no. 1, pp. 13–19, Mar. 2022.
[16] NR MIMO Evolution for Downlink and Uplink, document R1-213598, Samsung, 3GPP RAN1, Dec. 2021.
[17] J. Yuan, Y. Liu, Y. Hu, G. Xu, and J. C. Zhang, “Distributed FD-MIMO (D-FD-MIMO): From concept to field test,” in Proc. IEEE Radio Wireless Symp. (RWS), Jan. 2022, pp. 86–89.
[18] T. Wirth et al., “Modular concepts for practical massive MIMO implementations,” in Proc. IEEE Radio Wireless Symp. (RWS), Jan. 2020, pp. 60–67.
[19] Base Station (BS) Radio Transmission and Reception (Release 17), document TS 38.104, V17.0.0, 3GPP, Jun. 2021.
[20] 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.
[21] 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.
[22] H. Tatari, K. Haneda, A. F. Molisch, M. Shafi, and F. Tufvesson, “Standardization of propagation models for terrestrial cellular systems: A historical perspective,” Int. J. Wireless Inf. Netw., vol. 28, no. 1, pp. 20–44, Mar. 2021.
[23] G. Lee, M. S. Rahman, and E. Onggosanusi, “CSI feedback for distributed MIMO,” in Proc. IEEE Wireless Commun. Netw. Conf. (WCNC), Apr. 2022, pp. 2154–2159.
Kunpeng Liu received the B.S. and M.S. degrees from the Dalian University of Technology, Dalian, China, in 2005 and 2008, respectively. After graduation, she has been working with the Department of Wireless Product Line Network Research, Huawei. She is currently a Technical Expert and the Project Leader with the Multi-Antenna Standards Team. She has led the development of several massive MIMO technologies as the core 5G technologies in 3GPP standards, and has more than 100 granted patents. Her research interests include antenna architecture design, multi-antenna precoding, pilot design, multi-TRP cooperation, channel state information measurements, and signal processing algorithm design.

Min Zhang received the Ph.D. degree in telecommunication from Australian National University. Since 2016, he has been a Senior Standards Expert for 5G and 6G at Huawei Technologies Research and Development Ltd., U.K. Before working at Huawei, he was a Senior Standards Engineer and a Wireless Researcher at Nokia and Alcatel-Lucent. As a 3GPP RAN1 standards delegate, he has mainly worked on LTE and NR PHY design and specifications since 3GPP Release 11, as Rapporteur of the LTE standard and feature leads of NR for some MIMO work items. He has co-authored many papers in IEEE conferences, magazines, and journals, mainly focusing on state-of-the-art MIMO technology and implementation.

Leiming Zhang (Member, IEEE) received the B.S. degree from Nankai University in 2007 and the Ph.D. degree from Peking University in 2012. He is currently an Expert for research and standardization with the Wireless Network Research Department, Huawei Technologies Company Ltd., and the Manager of 3GPP 5G/SG advanced research and standards projects. He is the author of a number of papers and articles. His research interests include massive MIMO technologies, reference signals design, multi-TRP cooperation, and massive MIMO design for higher frequencies. He was a TPC Member of ICC’14 and ICC’15. He was a recipient of the Academic Top Ten Award from the School of Information Science and Technologies, Peking University. He was a Reviewer of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.

Gilwon Lee received the B.S., M.S., and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2010, 2012, and 2016, respectively. Since 2019, he has been working as a Staff Standard Engineer with Samsung Research America, Plano, TX, USA. Prior to joining Samsung, he was a Post-Doctoral Fellow in electrical and computer engineering with The University of Texas at Austin, TX, USA, and a Software Engineer with Intel Corporation, Santa Clara, CA, USA. His research interests include signal processing for large-scale MIMO, millimeter wave MIMO, and advanced MIMO technologies for beyond 5G and 6G systems.

Emad N. Farag received the B.S. and M.S. degrees in electrical engineering from Ain Shams University, Cairo, Egypt, in 1990 and 1994, respectively, and the Ph.D. degree in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 1997. Since November 2019, he has been with Samsung Research America, working on 5G NR standardization. His current research interests include SL communication, positioning, and MIMO.

Dalin Zhu received the B.Eng. degree in information engineering from the Beijing University of Posts and Telecommunications (BUPT) in 2007, the M.Sc. degree in electrical and computer engineering from Kansas State University in 2009, and the Ph.D. degree from the Department of Electrical and Computer Engineering, The University of Texas at Austin, in 2019. He is currently a Research Staff Engineer with Samsung Research America (SRA) working on 3GPP standardization. Before joining SRA, he worked as a Research Staff Member and the Project Manager at NEC Laboratories China (NLC) focusing on various topics related to wireless communications and standardization, and as a Staff Engineer at the Samsung Research and Development Institute China-Beijing (SRC-B) working toward new waveform design, non-orthogonal multiple access schemes, and mmWave communications. During the Summer of 2016, 2017, and 2018, he worked as an Intern at AT&T Labs, Austin, TX, USA, and Huawei Technologies, Rolling Meadows, IL, USA, focusing on 5G NR simulation and research.

Eko Onggosanusi received the Ph.D. degree in electrical engineering from the University of Wisconsin–Madison in 2000. He is currently the Senior Director of Standards with the Standardization and Multimedia Innovation (SMI) Laboratory, Samsung Research America. Prior to joining Samsung in 2014, he was the Manager at Texas Instruments. He has worked on cellular standards and algorithm/system designs especially HSPA and LTE systems. He also works on 5G NR (New Radio Interface). Having been a 3GPP RAN1 delegate since 2005, he has contributed to numerous components of LTE and NR physical layer specification. He has been the 3GPP rapporteur of the LTE EBF/FD-MIMO, LTE eFD-MIMO, Rel-16 NR-eMIMO, Rel-17 NR-FeMIMO, and Rel-18 NR Evolved MIMO work items. He has co-authored a number of papers in conferences and peer-reviewed journals. He is an inventor of numerous patents in wireless communications.
Mansoor Shafi (Life Fellow, IEEE) received the B.Sc. (Eng.) degree in electrical engineering from the University of Engineering and Technology at Lahore, Lahore, Pakistan, in 1970, and the Ph.D. degree in electrical engineering from The University of Auckland, Auckland, New Zealand, in 1979. From 1975 to 1979, he was a Junior Lecturer with The University of Auckland. He then joined the New Zealand Post Office, Wellington, New Zealand, which later evolved to Telecom New Zealand, and recently to Spark New Zealand. He is currently a Telecom Fellow (Wireless) at Spark New Zealand and an Adjunct Professor with the School of Engineering, Victoria University of Wellington, Wellington, and the School of Engineering, University of Canterbury, Christchurch, New Zealand. He is a Delegate of New Zealand to the meetings of ITU-R and has contributed to a large number of wireless communications standards in the ITU-R and 3GPP RAN 4. His research interests include radio propagation, design, and performance analysis for wireless communication systems, especially antenna arrays, multiple-input multiple-output (MIMO), cognitive radio, and massive MIMO and mmWave systems. He has authored over 200 articles in these areas. He received the IEEE Communications Society Public Service Award in 1992 “For Leadership in the Development of Telecommunications in Pakistan and Other Developing Countries.” He was made a member of the New Zealand Order of Merit, Queen’s Birthday Honors 2013, for Services to Wireless Communications. He has co-shared two IEEE prize-winning papers: the Best Tutorial Paper Award of the IEEE Communications Society in 2004 (co-shared with D. Gesbert, D.-S. Shiu, A. Naguib, and P. J. Smith) for the paper, From Theory to Practice: An Overview of MIMO Space Time Coded Wireless Systems, IEEE JSAC, April 2003, and the IEEE Donald G Fink Award 2011 (co-shared with A. Molisch and L. J. Greenstein) for their paper in the PROCEEDINGS OF THE IEEE April 2009, Propagation Issues for Cognitive Radio. He has been the Co-Chair of the ICC 2005 Wireless Communications Symposium, a Co-Guest Editor for three previous JSAC editions, the PROCEEDINGS OF THE IEEE, and the IEEE Communications Magazine. He has held various editorial and TPC roles in the IEEE journals and conferences. Most recently, his paper titled 5G: A Tutorial Overview Of Standards, Trials, Challenges, Deployment, and Practice has received the IEEE JSAC Best Tutorial Paper Award in 2022.

Harsh Tataria (Member, IEEE) received the B.E. and Ph.D. degrees from the Victoria University of Wellington, New Zealand, in 2013 and 2017, respectively. He is currently a Researcher in radio communications system standardization at Tataria Consulting. Since then, he has held several positions in academia and industry in the U.K., USA, Sweden, and New Zealand.