Received December 21, 2020, accepted January 2, 2021, date of publication January 18, 2021, date of current version January 28, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3052462

A Survey on 4G-5G Dual Connectivity: Road to 5G Implementation

MAMTA AGIWAL[1], HYEYEON KWON[2], SEUNGKEUN PARK[2], (Member, IEEE), AND HU JIN[3], (Senior Member, IEEE)

[1]Department of Electrical Engineering, Sejong University, Seoul 05006, South Korea
[2]Electronics and Telecommunications Research Institute (ETRI), Daejeon 34129, South Korea
[3]Division of Electrical Engineering, Hanyang University, Ansan 15588, South Korea

Corresponding author: Hu Jin (hjin@hanyang.ac.kr)

This work was supported in part by the Institute of Information and Communications Technology Planning and Evaluation (IITP) Grant funded by the Korea Government (MSIT) (Development of Frequency Analysis Technology for the Virtuous Circulation of Radio Resource) under Grant 2017-0-00109, and in part by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government (MSIT) under Grant NRF-2017K1A3A1A19071179.

ABSTRACT Rising popularity of 5G communications is making tremendous demands on the cellular network operators for providing true 5G services to the users. With limited numbers of 5G users initially, the investments for 5G services can be very high. In the early stage of 5G deployments, the 5G cells would not be lavishly spread and there would be 5G coverage holes. The operators can provide seamless services to the 5G users by inter working with the existing 4G Long-Term Evolution (LTE) network. The 5G inter working with fully deployed LTE would not only provide fast and seamless coverage but would also provide economic viability to the network operators. In this paper we survey and consolidate the 4G-5G inter working solutions that can assist in attaining the insight about various inter working possibilities and their challenges. It is important that a network operator is able to optimize its deployed infrastructure while being able to guarantee fast and seamless transition to 5G for its subscribers. To this regard, we evaluate the performance and radio resource management challenges for different 4G-5G dual connectivity options proposed by 3rd Generation Partnership Project (3GPP) standardization. We also discuss spectrum sharing possibilities between 4G and 5G wireless networks. Finally, various research challenges and discussions on path for migration to 5G standalone networks are also presented.

INDEX TERMS 4G, 5G, new radio (NR), dual connectivity, spectrum sharing, deployment options.

I. INTRODUCTION

Spanning over several unprecedented requirements, services and application, the 5G and beyond technologies are expected to enable not only the hyper digitization but they also present novel avenues for economical industrial growth. The 5G mobile communication system offer to fundamentally transform the role of telecommunication technologies in a society driving it into a one with pervasive connectivity [1]. Unlike the previous generations, the transition to 5G systems is much more than the enhancement of legacy LTE/LTE-A networks for it brings about extensive evolution of radio as well as the overall system architecture. The development of 5G networks considers novel spectral bands that can be broadly classified into two ranges: (i) $1-6$ GHz and (ii) above 6 GHz. The 3GPP (3rd Generation Partnership Project) standardization body has already specified the New Radio (NR) access technology and the new 5G core (5GC) network that incorporate the aforesaid spectrum [1]. The network operators must upgrade their networks to meet the demands of their new 5G subscribers while at the same time ensuring that the LTE subscribers' requirements are not compromised till they decide to make the transit. Over the past the network operators have intensively invested in the legacy LTE networks [2] and have already accelerated the network deployment and/or upgradation with the intend of full scale 5G-NR standalone roll out in the future. However, over the transition, the 5G subscriptions can not be expected to become very high instantaneously. Thus, the network operators can neither anticipate huge returns initially nor can they invest in the 5G network deployment at the unreasonable rate. We believe that the migration to 5G should be carefully evaluated and planned so

The associate editor coordinating the review of this manuscript and approving it for publication was Filbert Juwono.

that the network operators are benefited in the transition as well as beyond it. Moreover, they should not suffer economical impediments due to already deployed LTE infrastructure and LTE subscriptions. In this article we analyse the dual connectivity options that can be exploited by the network operators with the goal of smooth and 3GPP compliant transition to standalone 5G networks from their already established 4G LTE network.

In legacy networks, the dual connectivity was introduced to significantly increase the per-user throughput as well as to achieve enhanced mobility robustness by enabling the user equipment (UE) to simultaneously connect to two enhanced node Bs (eNBs) [3]. The same concept of dual connectivity has presented itself as one crucial feature for the smooth migration to 5G such that the UE can connect to both, the 4G base station (eNB) and the 5G base station (gNB), at the same time. This dual connectivity to both 4G and 5G manifest numerous advantages mainly stemming from the fact the 5G-NR systems are based on mmWave signals which are highly susceptible to blockage and path loss [4]. The channel quality on a mmWave link is extremely intermittent and rapid path switching would be common with frequent link failures [4]. In such an environment, the 4G-5G dual connectivity serves much greater purpose of maintaining reliable link quality without the requiring immediate investments on several gNB deployments. While the research works on dual connectivity in the legacy 4G networks were focused on (i) signalling overhead reduction (ii) throughput improvement (iii) out of order delivery (iv) multiple latency, the 4G-5G dual connectivity portrays different set of challenges like (i) fall back to 4G on mmWave link failure (ii) resource scheduling over different technologies (iii) mobility over different radio access technologies (RATs) (iv) coverage gaps in the two deployments (v) user plan and control plane management (vi) channel coding (vii) frequency reuse, etc. Mainly, in 4G-5G dual connectivity, UEs face different conditions on two different technologies. Despite all these new challenges, the network operators should be able to provide tight integration between 4G and 5G so that all the UEs experience seamless service over both 5G as well as 4G LTE [5]. Thus, in this paper we have specifically surveyed the 4G-5G dual connectivity in order to facilitate better understanding about the migration to 5G standalone wireless networks. We first analyse the 4G-5G dual connectivity background based on 3GPP standardization [6]. Through intensive survey and analysis we present the thorough understanding of the 4G-5G dual connectivity in terms of (i) deployment (ii) performance analysis (iii) radio resource management and (iv) spectrum sharing. More precisely the contributions of the paper are as follows:

- We first discuss the 3GPP compliant deployment Options 1 to 7 for 5G networks while laying special focus on multi radio dual connectivity (MR-DC). We specifically highlight characteristics for Options 3, 3a and 3x that correspond to 4G-5G dual connectivity. These options present opportunity for the inter-system backward compatibility. While investigating their advantages and disadvantages we bring about the clarity in terms of control plane and user plane deployment characteristics.
- We analyse the performance of 4G-5G interworking in terms of data rates, latency, mobility, reliability and energy efficiency as these are are key features of 5G communications [7].
- The efficient radio resources management (RRM) over two interworking but different nodes (or base stations) in MR-DC would help increase the system performance. Thus, we analyse the RRM objectives in 4G-5G dual connectivity so that the promised quality of service is provided to the subscribers.
- The study of frequency efficiency in 4G-5G dual connectivity also becomes important to leverage the advantages of costly spectrum. Effective use of frequency is not only necessary for the emerging mmWave spectrum but also for the sub GHz spectrum used in legacy networks for which the network operators must have already made huge investments. Thus, we also investigate 4G-5G dual connectivity from the spectrum sharing perspective.
- Subsequently, we discuss the steps that lead towards the ultimate deployment of the standalone 5G networks from the current possibilities. We specifically elaborate deployment Option 4 as it is the natural next step for migration to 5G networks. Finally, we delineate several open issues and research challenges that need to be addressed for this transition.

The rest of the paper is organized as follows: In Section II, we present the literature survey and research background on dual connectivity with a special focus on MR-DC deployment options for 5G. In Section III, we describe the performance parameters that can be used to compare different MR-DC deployment options. Section IV presents the radio resource management for the MR-DC options. In Section V, we delineate details on the spectrum sharing between 4G and 5G so that the mobile operators can maximize the advantages of both the LTE spectrum and the mmWave spectrum. In Section VI, we present the analysis of Option 4 that supports dual connectivity based on 5G core network and also present research challenges. Finally, conclusion is drawn in Section VII.

II. DUAL CONNECTIVITY BACKGROUND

In this section we highlight the 3GPP compliant architectural options identified from the legacy 4G LTE to full fledged 5G standalone. However, it is notable that all the deployment options would not be practically feasible. We specifically analyze Options 3, 3a and 3x as they endorse 4G-5G dual connectivity. These options are of major current relevance since all LTE deployment can not transit to 5G overnight and involves huge cost. More importantly, Options 3, 3a and 3x present opportunity for the inter-system backward compatibility. However, we first discuss dual connectivity

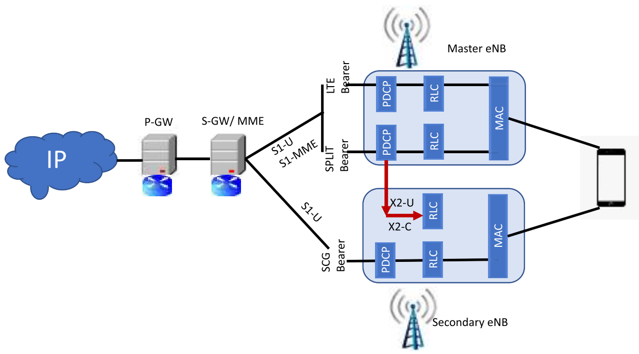


FIGURE 1. Dual connectivity in LTE.

in 4G LTE as the fundamentals of 4G-5G dual connectivity are embedded in it.

A. DUAL CONNECTIVITY IN 4G LTE

Mobile operators have always been attracted towards capacity boosting methods that are alternative to procuring costly spectrum from the government. Small cell deployment in high density areas have been one particularly popular solutions for achieving higher mobile network capacity [8]–[10]. The small cell deployment along with the macro infrastructure led to the idea of dual connectivity. 3GPP Release 12 introduces the dual connectivity feature such that a UE can simultaneously transmit and receive data on several component carriers belonging to two different cell groups (or eNBs) [11], [12]. The dual connectivity technology offers significant improvement in the per-user throughput, especially at the cell edges. It also provides mobility robustness and load balancing [13]. As shown in Figure 1, in LTE dual connectivity, the UE connects to master cell group (MCG) via MeNB (master eNB) and secondary cell group via SeNB (secondary eNB), simultaneously [12]. The interface between MeNB and SeNB facilitates the communication between the two. It can be based on non-ideal backhaul and is called as the X2-interface [14]. The X2-interface facilitates an efficient resource management in dual connectivity [13]. While an ideal backhaul is characterised by very high throughput and very low latency as provided by a dedicated point-to-point connection that uses optical fiber, the non-idle backhauls (like microwaves) are more market and cost friendly [15].

In LTE dual connectivity as shown in Figure 1, the UE sees only one S1-MME (Mobility Management Entity) connection for its *control plane* functionalities. After coordinating the radio resource management (RRM) functions from both MeNB and SeNb over the X2-interface, the MeNB generates the final Radio Resource Control (RRC) message for the UE. This makes it easier for the UE RRC entity to operate as it has to monitor from and reply to only one entity at the MME [3]. Moreover, if there are frequent SeNB changes the signalling overheads towards the core network are reduces as there is only one S1-MME connection (to MeNB) per UE [16]. While the link between UE and SeNB is controlled by MeNB only, the *user plane* architecture supports the following two options:

- The user plane connection from serving gateway (S-GW) can terminates at both MeNB and SeNB. This option is analogous to carrier aggregation in legacy networks where one UE is simultaneously connected to two different carriers.
- The user plane terminates at MeNB and subsequently a split bearer manifests link between MeNB and SeNB through X2-interface. This linkage is especially useful when considering traffic off-loading in legacy networks [12].

Thus, based on the aforesaid two options, the user traffic may split at either the S-GW or at the MeNB. In a non-ideal backhaul it is difficult to split MAC-PDUs since they are generated in real time [16]. Thus, the 3GPP release 12 proposed (packet data convergence protocol) PDCP PDUs level split at the MeNB and the bearer level split at the S-GW. These major LTE dual connectivity features, such as (i) simultaneous connectivity to two nodes (or base stations), (ii) X2-interface, (iii) only one control plane connection (iv) user plane through both the nodes and (v) split at the PDCP PDUs, discussed above are also adopted for coupling between the legacy 4G LTE and the new 5G air interface 5G to obtain 4G-5G dual connectivity scenarios. [6]

B. 4G-5G DUAL CONNECTIVITY

The 5G New Radio (NR) has become extremely popular in last few years. The study and standardization of 5G has been accomplished in a record time by 3GPP [17]. The study on vision of communication for 2020 and beyond is summarized in ITU-R M.2083 by International Telecommunication Union-Radio communication (ITU-R) sector. It highlights three major directions (i) enhanced Mobile Broadband (eMBB) (ii) massive Machine Type Communication (mMTC) and (iii) ultra Reliable Low Latency Communication (uRLLC). Thus, the communication for 5G and beyond would be driven by the diversity of challenging 5G requirements. One of the major challenge is the provision of very high throughput to address the steep increase in wireless traffic volume. The mmWave bands have emerged as the possible enablers of ultra high throughput. While the sub GHz frequencies are very fragmented and limited, the network operators can be allocated with larger chunks of GHz bandwidths in mmWave frequency range [7]. However, use of frequencies above 6 GHz presents several issues like, high isotropic pathloss, attenuation by moisture, attenuation by foliage, smaller range of wireless link and blockage from common objects and infrastructure [18]–[20]. Thus, wireless links using mmWaves may provide very high throughput but at the cost of link quality which may vary over several factors, like path-loss, deafness, and blockage [20]. On the other hand, the widely deployed legacy LTE system provides reliable coverage over larger distance.

Moreover, unlike the previous generation, 5G-NR requires not only a new access network but also a new core network.

TABLE 1. Comparison between 3, 3a and 3x: Deployment Requirements.

	Option 3	Option 3a	Option 3x
Utilization of resources over LTE and 5G	Possible with same bearer MSG split bearer.	Not possible with same bearer Two DRB bearers: MSG and SCG.	Possible with same bearer SCG split bearer.
User Plane latency	Additional user plane latency over Xx-interface.	No additional user plane latency.	Additional user plane latency over Xx-interface.
Backhaul requirements	Xx-interface should offer sufficient capacity and latency from 5 to 30 ms.	No additional requirements.	Xx-interface should offer sufficient capacity and latency from 5 to 30 ms.
Processing capacity requirements	Additional PDCP processing capacity requirements at the eNB.	No additional PDCP processing capacity requirements.	Additional PDCP processing capacity requirements at the gNB.
Traffic ofloading	Dynamic offloading controlled by eNB.	MME involved. Very static.	Dynamic offloading controlled by gNB.

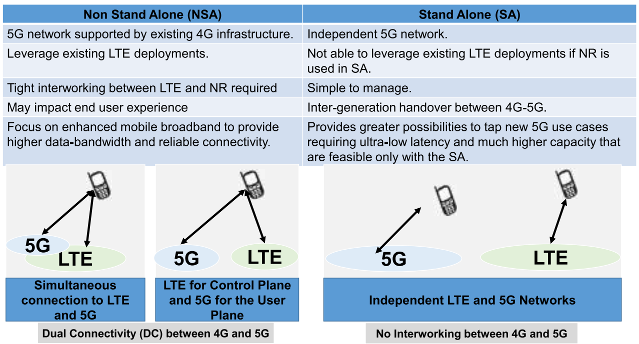
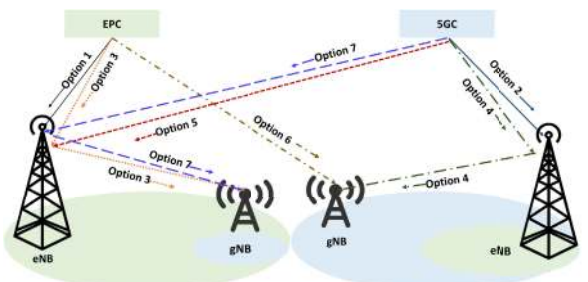


FIGURE 2. Non Standalone and standalone deployment.



	Option1	Option2	Option3	Option4	Option5	Option6	Option7
Core	EPC	5GC	EPC	5GC	5GC	EPC	5GC
MN	eNB	gNB	eNB	gNB	eNB	gNB	eNB
SN	NA	NA	gNB	eNB	NA	NA	gNB

NA: Not applicable

FIGURE 3. Deployment options.

3GPP has defined a new 5G core network (5GC) along with new radio access (5G-NR). Integration of elements from legacy LTE network is also possible in different configurations of 5G deployment. As a result there are a number of deployment options that consider different combinations of 5G-NR, 5GC, LTE core (EPC) and LTE access (E-UTRA). These options are shown in Figure 3. However, it is to be noted that not all options would be practical [11]. When only one radio access technology is used, it is a standalone (SA) scenario as in Options 1, 2, 5 and 6. In non-standalone (NSA) as shown by Options 3, 4 and 7, NR radio cells and LTE radio cells inter work, using dual connectivity, to provide a combined radio access to UEs. In such a dual connectivity scenario, the core can either be the 5GC or LTE's evolved packet core (EPC) [1]. Figure 2 highlights the differences between NSA and SA architectures.

1) DEPLOYMENT FLEXIBILITY

3GPP defines multi-radio dual connectivity (MR-DC) such that UEs can utilize resources from two different nodes (eNB or gNB) that are connected via non-ideal backhaul using Xx-interface [6]. In MR-DC, one node provides NR access and the other can provide either NR access or E-UTRA. Out of the two nodes, one acts as a Master Node (MN) and is connected to the core through an S-interface. The other is considered as Secondary Node (SN). When the eNB connected to EPC acts as an MN, then the MR-DC is called as E-UTRA-NR Dual Connectivity (EN-DC) [6]. In EN-DC, the gNB may or may not connect to EPC. Given the widespread LTE deployment and high capital expenses of 5GC and 5G-NR, EN-DC has appeared to be the choice of mobile operators for the initial 5G demonstrations and subscriptions.

However, in future it is expected that the 5GC core would be connected to a gNB or an enhanced eNB which can act as a MN. 3GPP defines three configurations for such MR-DC:

- NGEN-DC where the enhanced eNB acts as an MN and is connected to the 5GC. The gNB acts as an SN.
- NE-DC where the gNB connected to 5GC acts as an MN while the enhanced eNB acts as an SN.
- NR-DC in which the NR-NR dual connectivity is supported such that both MN and SN are gNBs only.

It can be seen from Figure 3, that the EN-DC corresponds to Option 3 of the deployment scenario. Similarly, while NGEN-DC corresponds to Option 4, NE-DC maps to Option 7 of the deployment possibilities. It is natural for the initial 5G wireless network to be built in NSA mode based on fundamentals of EN-DC given the popularity and wide spread deployment of LTE networks. In this paper we are focused on understanding the migration to 5G from the network operator's perspective and, thus, we first analyze Option 3 (EN-DC). Subsequently, in the section on future directions, we delineate the discussion on Option 4 that manifests MR-DC by again considering both gNB and eNB but with the 5GC.

Considering the user perspective in EN-DC, the UE can exchange data with the 5G gNB while also manifesting an LTE connectivity. For this, the UE may (i) simultaneously connect to both LTE and NR or (ii) may derive *control plane* functionalities from LTE network and *user plane* functionalities from 5G networks as shown in Figure 2 [12]. Thus, EN-DC itself can manifests multiple possibilities. Figure 4

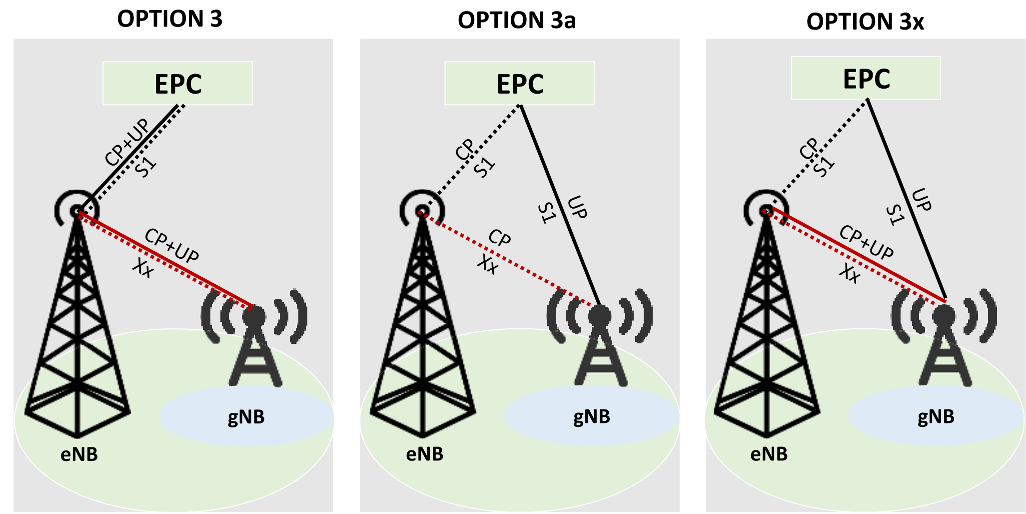


FIGURE 4. Options 3, 3a and 3x.

gives an overview of the 3GPP specified possibilities for EN-DC i.e. Options 3, 3a and 3x. They are explained as follows:

1) **Option 3** In this option, all data (even the one intended for 5G-NR) is traversed to and from the eNB that acts an MN. The gNB bears no direct link to the EPC. The key features are:
   - The S1-U that stands for the user plain interface is anchored at the eNB as shown in Figure 4.
   - Mobility signaling between LTE's eNB and the NR's gNB is not visible to the EPC.
   - S1-U is split at the eNB.
   - There is increase in the load on the eNB as it is required to process both LTE as well as the 5G user traffic.
   - Xx-interface needs to support control plane as well as the 5G user traffic, the flow control and the latency requirements.
   - Limited impact on the UE mobility interruption (NR to LTE) which occurs due to split/switched bearer in eNB i.e. traffic sent over LTE when the UE is outside the coverage of NR.
2) **Option 3a** In Option 3a, both the eNB and the gNB can directly communicate to the EPC. The X2-interface bears only the control signals and no data is shared over it. The key features are:
   - The S1-U is anchored at gNB.
   - The mobility signaling between the legacy eNB and the gNB is visible to EPC due to S1 path switch between them.
   - S1-U is not split, nor does it delivers any data over NR. The eNB is not required to handle additional load or flow control.
   - Xx-interface is used for control plane traffic only.
   - Impact due to UE mobility interruption (NR to LTE) is high as there is a requirement of S1 path switch from gNB to eNB.
3) **Option 3x** In this option, the gNB is directly connected to the EPC, allowing the smooth flow of data traffic between the two. The eNB can also be forwarded a part of the data over the X2-interface. The key features are:
   - The S1-U is anchored at the gNB.
   - Mobility between eNB (LTE) and gNB (NR) is visible to EPC due to S1 path switch.
   - S1-U splits at the gNB.
   - The eNB can transmit some fraction of user data.
   - Xx-interface needs to support control plane, split user traffic, flow control and strict latency requirements.
   - Impact of UE mobility interruption (NR to LTE) is limited due to split/switched bearer i.e. traffic sent via Xx-interferce to LTE when UE is outside the coverage of NR. S1 path switch also leads to some impact.

2) CONTROL PLANE IN MR-DC

A bearer is referred to as a virtual connection between the UE and the core that facilities transport of data with specific quality-of-service (QoS) attributes. There are two types of radio bearers (i) signalling radio bearer (SRB) that supports the control plane data and(or) (ii) data radio bearer (DRB) that are employed for transfer of user plane packets [21]. From the radio perspective, 5G-NR is expected to support the radio bearer concept. For this 5G-NR has introduced a new access stratum sublayer in PDCP referred to as service data adaptation protocol (SDAP) sublayer [11]. It is responsible for QoS identification in data packets and the mapping of QoS to DRBs. It is notable that due to the limitation of the current EPC, SDAP can only be supported by 5GC.

Radio resource control (RRC) layer accounts for the control plane functions. The functions performed by RRC are like broadcasting of reference signals and system information, mobility management, configuration of lower layer protocols, measurement and configuration reporting, etc [22]. In MR-DC, for each UE, there can only be a single control plane connection towards the core network and it is through the MN as a general principal [1], [6]. It provides a more robust system as MN is responsible for the maintenance of the RRC connections including the dual connectivity configuration. The MN is mainly responsible for generating/ sending RRC messages to the UE [22]. The control plane signalling and coordination between MN and SN is achieved by the Xx-interface.

The prime responsibility of MN and SN in MR-DC is the allocation of the radio resources. For this, both MN and SN can have their own RRC entities but the UE can have only one RRC state. The RRC protocol data units (PDUs) from SN can be transported to UE via MN using SRB1 and SRB2 which are the defined signal bearers in legacy 4G networks. However, new bearer SRB3 can be configured in 5G-NR and it can be established so that SN can directly send RRC PDUs to the UE without traversing them through the MN [6]. It is especially useful when the UE needs to provide the measurement report for mobility within SN. While transporting the RRC PDUs from the SN, the MN can not alter the SN provided configurations. It is important to note here that the initial SN RRC configuration is always sent to UE by MN only using SRB1. The subsequent reconfiguration has no such limitations. In MR-DC, split SRB is supported. However, it allows duplication of RRC PDUs that are generated by the MN only and does not support duplication of SN generated RRC PDUs. Thus, we can infer that in MR-DC most of the control plane functionalities are provided by the MN. The SN can support functions like measurement reporting so that fast interaction can take place between UE and SN when required, for example in mobility management with SN. In EN-DC, control plane signalling to EPC is always through the eNB. For all the Options 3, 3a and 3x, the control plane is similar and terminates at the LTE eNB which is the MN in EN-DC.

3) USER PLANE IN MR-DC

From the network's perspective, the user plan connectivity to the core entity can either have the MN terminated data radio bearer (DRB) or SN terminated DRB. The data to/from the UE can be transported over (i) master cell group (MCG) bearer, (ii) secondary cell group (SCG) bearer or (iii) *split bearer* [22]. Each of these bearer can terminate either at MN or SN. The radio bearer setup determines the radio protocol for MN, SN and the UE [11]. The MCG and SCG bearers can utilize resources while following the radio protocols located only in MN and SN, respectively [11]. However, the split bearer configuration allows the UE in an NSA MR-DC to utilize resources from two distinct schedulers (i) one in the MN and (ii) the other in the SN. While one node can provide LTE-access, the other facilitates the NR-access [1]. Thus, the split bearer can transmit some data to the UE via MN and some via SN [23]. In the MCG split bearer, the user plane data from the core is split at the MN. As an example, when the MSG split bearer is utilized for the data downlink, then the decisions on the packet routing are made by MN based on considerations such as traffic, channel conditions, buffer status and the capacity of the non-ideal backhaul interface [11]. Option 3 of EN-DC deployment is an example of MCG split bearer. SCG split bearer is also supported in 5G networks [24]. In SCG split bearer, the user plane data from the core gets split at the SN, even though the MN is connected to the core [11]. The split of data for MCG and SCG split bearers takes place at the PDCP layer. Figure 4 shows the radio protocol options for Options 3, 3x and 3a. The EN-DC deployment option employing SCG bearer is referred to as 3x [11]. We would like to again emphasize here that there can be different connectivity options for the user plane in MR-DC but the control plane terminates only at one node (eNB or gNB). For Option 3 in Figure 4, eNB should be upgraded in terms of processing power as well as the buffer size such that it can support the MCG split bearer. While Option 3a in Figure 4 do not require any upgradation as it does not support the split carrier. In Option 3x, the splitting for the user data is configured at the SN enabling the utilization of both SN and MN carriers simultaneously, without any need for eNB modifications [25]. Figure 5 shows the user plane and control plane architecture.

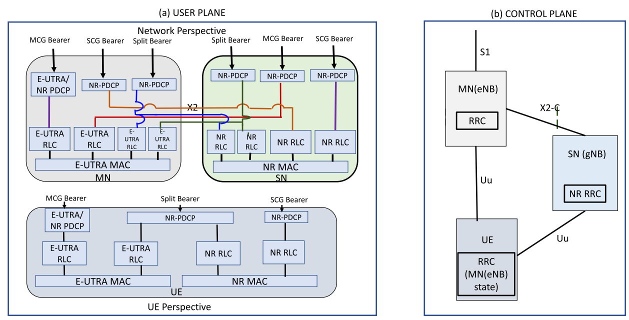


FIGURE 5. User Plane and Control Plane.

III. ANALYZING THE 4G-5G DUAL CONNECTIVITY PERFORMANCE

In this section we analyze the effect of 4G-5G interworking in terms of data rates, latency, mobility and reliability. These are key features of 5G [18] and understanding their performance with respect to the deployment options would facilitate efficient migration to standalone 5G networks.

A. DATA RATE

5G-NR targets to achieve 20 Gbps peak data rate for downlink [26]. Peak data rate can be defined as the highest theoretical data rate under error-free conditions such that all the assignable radio resources available for considered link direction are exclusively assigned to a single mobile station [26]. It can be analytically evaluated by multiplying peak spectral efficiency (η_{peak}) with the available bandwidth (BW). By aggregating a 5G cell with an LTE eNB at a radio access level, the peak data rates for the two can be summed up resulting is the higher values than that provided by 5G standalone [27]. The multi-link aggregation can enhance both,

TABLE 2. Data Rate Analysis.

Option	Key points	Peak data rate
LTE	Bandwidth = 20 MHz, Spectral Efficiency = 15 bits/s/Hz	300 Mbps
5G	Bandwidth = 100 MHz, Spectral Efficiency = 30 bits/s/Hz	3 Gbps
3	Supports throughput from both eNB and gNB Reasonable data rate when poor link from gNB	3.3 Gbps
3a	Supports throughput from gNB only Low data rate when gNB path becomes NLOS	3 Gbps
3x	Supports throughput from both eNB and gNB Reasonable data rate when poor link from gNB	3.3 Gbps

the peak data rates and the area spectral efficiency [28]. This is crucial for the early stages of 5G deployment since higher throughput can be achieved even at lower values of bandwidth considerations. The data traffic can be split at the transmitter's PDCP layer [29] and subsequently, data over the two radio paths can be aggregated at the receiver again at the PDCP layer. This would result in boost in throughput for the end user [29], [30].

It is to be noted that the definition of peak data rate considers an error free environment. However, the mmWave connection in 5G-NR are prone to impediments like, high path loss, severe shadowing, link limitations due to the height and placement of cells and frequent lack of line-of-sight [31]. When the signal to interference plus noise ratio (SINR) drops, the UE would experience loss in data rate. The simulation results in [31] point out that the throughput is very low when the path between UE and mmWave gNB becomes non-line of sight. Thus, it can be reasonably augured that the dual connectivity not only supports increase in throughput when SINR is good but can also be instrumental in maintaining reasonable data rate when link from the gNB is relatively low in quality. Thus, for maintaining reasonable throughput the 4G-5G dual connectivity is crucial. Table 2 provides understanding of throughput for different EN-DC deployment options. The Options 3 and 3x can aggregate data rates from both gNB and eNB. Thus, for these options the peak data rate can go up to 3.3 Gbps (20 MHz $\times$ 15 bits/s/Hz + 100 MHz $\times$ 30 bits/s/Hz). However, no aggregation of data rate is possible in Option 3a as it is supported by the gNB only and therefore, manifests a peak data rate of 3 Gbps (100 MHz $\times$ 30 bits/s/Hz). Moreover, the 4G- 5G inter working in Option 3 and 3x can address the loss of mmWave connection from the gNB by accomplishing the fast switching to the readily available eNB. According to authors in [32], fast switching can address more stable and higher throughput. One of the key consequence of the above procedure is the requirement of an efficient coordinated wireless transmission schemes for effective delay skew control [33]. Thus, it is needed that the analysis of 4G-5G inter working should also include the details of delay control between different transmission points [33] as discussed in the next subsection.

B. USER PLANE AND CONTROL PLANE LATENCY

3GPP defines user plane latency as the time for a successful delivery of an application layer packet over the radio such that the discontinuous reception (DRX) cycle do not restrict either the device or the base station. It is analyzed as the period from the time when the transmitter receives the packet till the time when the receiver successfully receives the packet [26]. Unlike legacy 4G LTE networks, 5G-NR manifests several delay evaluation possibilities as it supports multiple numerologies each with different subcarrier spacing (15 KHz, 30 KHz, 60 KHz, 120 KHz) [34]. In turn, the value of transmission time interval (TTI) depends upon the subcarrier spacing and number of symbols per TTI. 5G-NR supports variable TTIs and users can be scheduled according to their radio conditions and quality requirements [35], [36]. Another important fact to consider here is that 5G is expected to facilitate individual handling of user plane which is related to application payload transmission and control plane that is related to control functionality and signaling [37]. We first evaluate the user plane latency as in Table 3 while assuming subcarrier spacing as 60 KHz, two OFDM symbols per TTI for 5G, transmitter and receiver processing delay as one TTI each and HARQ round trip time as six to eight TTIs. The latency can be obtained as summation of transmitter and receiver processing delays (t_A), frame alignment time (t_B), transmission time (t_C) and re-transmission time (t_D). The HARQ RTT timer along with the Block Error Rate (BLER) failure rate (n_{BLER}) can be used to evaluate t_D. The HARQ RTT timer is set to 8 subframes for FDD LTE and thus, the HARQ RTT can be considered as 8 ms [38]. Assuming that the receiver successfully receives transmission at least 95% of times, then the HARQ BLER would be 5% resulting in the HARQ failure rate, n_{BLER}, of 0.05 as per the LTE protocol design [39]. Thus, we can obtain $t_D = n_{BLER} \times 8$ ms for LTE [27]. In EN-DC of the immediate 4G-5G dual connectivity, a packet flow is transmitted through either both, the LTE and the 5G-NR air interface or only 5G-NR interface depending on the deployment option.

Control plane latency is defined as the time taken for the transition from the inactive/idle state to the start of the continuous data transfer [26]. Similar to legacy network the transition to active state from the idle state is manifested by the random access procedure. Random access is a four-step procedure for the establishment of the initial link [40]. The 5G signalling procedure is similar to 4G. The difference between the 4G and 5G control plane occurs due to different system parameters, such as TTI and processing delay [27]. Table 3 shows the latency evaluations for both 4G and 5G networks. If we consider initial EN-DC deployment such that the control protocol is located in the legacy eNB, the control plane latency is still the same as in the legacy LTE networks. The user plan and control plane latency evaluations specifically for the EN-DC deployment options 3, 3a and 3x are presented in Tables 4. For Option 3, if the user receives data via eNB then it would experience a delay of 4.4 ms for 5% BLER. However, if the UE receives data from the gNB then the delay of 4.4 + 0.1536 ms is observed as the data traverses first through the eNB then the gBN as shown in Figure 4. Similarly, for Option 3x, the data can either experience user plane latency of 0.1536 ms or 4.4 + 0.1536 = 4.5536 ms

TABLE 3. Latency Evaluations.

User Plane Latency									
	Subcarrier Spacing	OFDM Symbols per TTI	OFDM Symbol = Symbol Duration +CP Duration (ms)	t_A=Transmitter + Receiver processing delay (ms)	t_B=Frame alignment Time (ms)	t_C= Transmission time (ms)	One way latency ($t_A + t_B + t_C$) (ms)	HARQ RTT (ms)	User Plane Latency with 5% BLER (ms)
LTE	15 KHz	14	0.07135	2.5	0.5	1	4	8	4.4
5G	60KHz	2	0.01784	0.3575	0.01784	0.01784	0.1429	0.2143	0.1536
Control Plane Latency									
	RACH scheduling (ms)	RACH Preamble (ms)	Preamble detection + transmission of RAR (ms)	UE Processing (ms)	RRC Connection Resume Request (ms)	Base station Processing + transmission of UL grant (ms)	L2 and RRC Processing at UE (ms)	Transmission of RRC Connection Resume complete (ms)	Total Delay (ms)
LTE	0.5	1	3 (2 + 1)	5	1	4+1	15	1	31.5
5G	0.07142	0.143	0.6667 + 0.143	5/3	0.143	4/3+0.143	15/3	0.143	9.5

TABLE 4. User Plane and Control Plane Latency.

Option	User Plane Latency	Control Plane Latency
3	4.4 and 4.5536 ms	31.5 and 41 ms
3a	0.1536 ms	31.5 and 41 ms
3x	0.1536 and 4.5536 ms	31.5 and 41 ms

when receiving data from gNB or eNB, respectively. The user plane latency for Option 3a (0.1536 ms) is better than that for Option 3x as in this deployment option all the user plan data is delivered only through the gNB. It is notable that the limits in the user or control plane latencies are the sum of limits of LTE and 5G and not the maximum between them. Since the control is always anchored as the eNB in ENDC, the control plane latency for all the three cases (3, 3a and 3x) can either be 31.5 ms or $31.5+9.5 = 41$ ms. In general, Option 3 performs worst in terms of user plan latency with the value greater than 4 ms for both the cases, whether data traverses through eNB or gNB.

C. MOBILITY AND RELIABILITY

The 5G requirements of low latency, high reliability and short interruption times require a tighter inter working with legacy LTE networks and their evolution [37]. Interruption time while a handover occurs can be used to understand the mobility performance. In EN-DC, the handover can manifest when (i) gNB changes while eNB remains unchanged (ii) eNB changes. When eNB remains unchanged in Option 3, the gNB to gNB handover can be performed with almost zero interruption since the UE can continue to receive data from the unchanged eNB. Such a scenario projects a multi-connectivity handover as shown in Figure 6. For Option 3a, we assume the interruption time to be MT_{gNb} that originated from a gNB to gNB handover through an EPC core. However, for Option 3x first gNB changes and then new gNB needs to be connected to the original eNB. If the time to establish link between new gNB and eNB is NT, then the total time for complete handover in Option 3x becomes equals to $M_{gNb} + NT$ as opposed to Option 3a where it is just MT_{gNb}. Option 3x however can also achieve low handover interruption if the data can be buffered at the eNB while gNB updates. The change in eNB, from one to another, can be assumed to be accomplished by change in gNB as well since the coverage of eNB is more than that of gNB. When eNB changes, Option 3x would require extra handover time than Option 3a to establish the new link for data communication between new eNB and new gNB.

Other than the interruption time, an important parameter to be considered when discussing mobility is the sudden dropping of the 5G-NR radio link to the device. Probability of sudden link deterioration is high for mmWave coverage and cell selection procedures are enhancing to incorporate mmWave related aspects like non-line-of-sight propagation and sensitivity to blockages [41], [42]. If the 5G-NR radio link is lost suddenly then the user service could be continued with LTE eNB after NR-to-LTE user path switching in Options 3 and 3x. Though, such possibility pertaining to user data is not feasible in Option 3a, it still offers advantage of low signaling overhead through control anchored at the eNB. It is notable that signaling overhead are considerable during handover execution phase [42]. In case when a robust connection and higher reliability is required, Option 3x has an advantage over Option 3a. The path for user traffic can immediately be changed to eNB in Options 3 and 3x to avoid impediments of the dropped 5G-NR radio link. In this case the NR to LTE path switching time determines the effectiveness of the deployment. Whenever the 5G-NR radio link drops the new/old gNB needs to added/reconnected which would require more signalling. The detection of new gNB and/or the new beam and the subsequent switching would also be time consuming. The details on beam recovery are discussed in the subsequent subsection. Table 5 gives dual connectivity performance based on the analysis of the handover/procedure.

D. BEAM FAILURE RECOVERY

Beamforming would have an essential role in 5G networks particularly in the context of frequency bands above 6 GHz [37]. The spotty coverage of high frequency mmWaves can be mitigated by directional beam communication [43].

TABLE 5. Handover Performance in ENDC.

	Option 3	Option 3a	Option 3x
Handover Interruption Time (LTE eNB Same gNB Changes)	No Interruption	Same as gNB to gNB	gNB interruption+ adding LTE eNB again
Handover Interruption Time (LTE eNB Changes gNB Changes)	eNB to eNB change+ adding new gNB	eNB to eNB change+ gNB to gNB change	eNB to eNB change+ gNB to gNB change+ adding new eNB
Fall to LTE from NR	Possible	Not Possible	Possible
Path Switching time in case of NR Link failure	Small since data aggregation at PDCP	High since more signalling for adding a new gNB	Small since data aggregation at PDCP
Advantage	Robustness against 5G link failure	Low interruption time when LTE or gNB changes	Robustness against 5G link failure

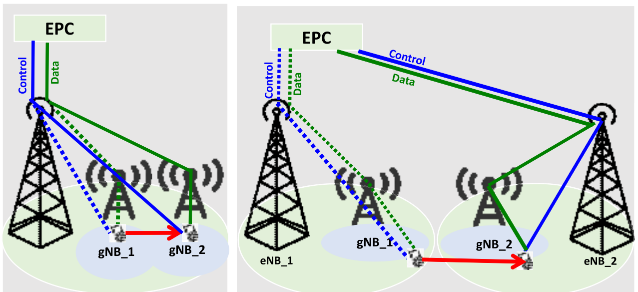


FIGURE 6. Handover: (a) gNB changes (b) eNB changes.

With highly directional beams, fast varying channels and several small cells, this directional tracking may be the main bottleneck in realizing reliable mmWave 5G networks. For communication between UE and gNB, the channel quality must be continuously scanned across multiple possible directions as each beam manifests limited spatial coverage. This can dramatically increase the time it takes to detect a reliable link. Moreover, when the beam is lost, the link fails and a path switch becomes necessary. To accommodate high frequencies, the number of beams would have to increase and both transmitter-side and receiver-side beamforming would be supported [44]. Thus, the beam searching time can be substantial and may cause interruption similar to when handover occurs. Moreover, the dynamic characteristics of mmWave channels imply that the links to any one cell can deteriorate rapidly, necessitating a faster link failure detection and re-routing [42]. Thus, fast switching for a beam failure recovery is an important consideration for reliability in 5G networks. It is notable that unlike LTE, networks that use only higher bands would require beamforming of not only data channels but also of the control channels [37]. Thus, the standalone 5G would require elaborate beam searching overheads. MR-DC on other hand can facilitate agile path switching when beam failure occurs as UE is simultaneously connected to gNB as well as the eNB. The eNB operating at sub GHz frequencies can enable more conducive propagation characteristics. In EN-DC, since the control plane connection is established by the more robust LTE link, it can assist the user plan to rapidly switch between the radio access technologies in case of beam failure. Thus, the subsequent beam recovery can then be performed without loss of connection. Thus, Options 3 and 3x of EN-DC deployment can manifest agile rerouting to LTE eNB in case the gNB link deteriorates and beam is lost. However, Option 3a lacks this advantage.

IV. RADIO RESOURCE MANAGEMENT

The major objective of the radio resource management (RRM) is the guarantee of network services at the desired quality optimized over the usage of system resources [45]. Since energy efficiency is an important issue in any wireless network, optimized energy expenditure is also a key feature that impacts the effective RRM [45]. Considering the 4G-5G dual connectivity scenario, the RRM operations deserve more attention so that the proper utilization of the multi connectivity solutions is achieved [22]. In MR-DC the efficient usage of resources over two inter working but different nodes would help increase the system performance. In this section we analyze the aforesaid RRM objectives in 4G-5G dual connectivity as we believe that the network operators main objective is to provide the promised quality of service to its subscribers. Thus, viable deployment should consider effective RRM along with other performance objectives.

A. SYSTEM INFORMATION AND MEASUREMENT

The UE in MR-DC can obtain the resources from either eNB or gNB. The resource allocation from two different nodes at different technologies offers to provide different types of services with different delay and throughput requirements. Moreover, the resource allocation to the UE can be based on the objective of the optimal power usage. The correct channel selection, power selection and frequency adjustment are important aspects that have attracted many researchers [45]–[47]. However, the basis of these correct selections is the correct measurements. Before the measurements are performed or noted, the system information is broadcast by the network and monitored by the UE. In MR-DC, the master node (MN) provides the UE with the system information for the initial configuration through a dedicated radio resource control (RRC) signalling. The secondary node (SN) needs to provides only the radio frame timing and the subframe number to the UE. In EN-DC, the control signalling is anchored at the eNB for all the deployment Options 3, 3a and 3x. Thus, the eNB performs all the major measurements and broadcasting in EN-DC. This has an advantage as the 4G is known to be more reliable over wider coverage as compared to standalone 5G systems.

B. SN ADDITION AND MODIFICATION

First MN establishes the connection with the UE through a random access procedure and then it can configure measurements to the UE so as to add an SN. To add an SN, UE performs random access procedure again but on the SN. Once the SN is connected, the measurements can be configured by MN or SN, independent of each other. In case of SN mobility, SN configures the measurements to UE. While SN addition is always initiated by the MN, the change in SN can be with or without the involvement of MN. For instance in EN-DC, eNB coordination is not needed when the security key is not required to be changed and the eNB need not involve in PDCP recovery [6]. In such a case the signalling for gNB mobility is substantially reduced. MR-DC also allows the the handling of combined messages from MN and SN in case both require reconfiguration. To accomplish this, the SN RRC message for reconfiguration can be encapsulated in the MN RRC message. It is to be noted that for Options 3, 3a and 3x the aforesaid control procedure remains the same even though the MN traverses no data in Option 3a.

C. SCHEDULING AND DATA FORWARDING

To exploit the potential of MR-DC, the traffic between the core and UE can be spread over two paths; (i) one over eNB and (ii) the other over gNB. While scheduling, the MN can not randomly transfers UE's data to the SN. If little data is transferred through the SN, then the benefits of the dual connectivity would be limited. On the other hand, more data may lead to congested back ground traffic. Moreover, as the channel conditions for the mmWave frequencies are susceptible to poor quality, it may not be feasible for the gNB to deliver data to the UE at all at some times. This raises two important issues, (i) Firstly, the scheduling in dual connectivity requires dedicated research effort such that neither the advantages of dual connectivity are compromised nor should it causes congestion. (ii) Secondly, the traffic traversing over two different paths would experience different delay due to different radio access techniques (RATs), parameters, background traffic, path loss, channel conditions, etc. Thus, the objective of scheduling in MR-DC could be to minimize the end-to-end latency while addressing the constraints of traversing the data over two different RATs. It can be defined as the maximum experienced delay by the UE over dual connectivity [17]. Authors in [17] have addressed the latency issue by adopting the deterministic network calculus theory. The work shows that the latency depends upon the traffic. When the traffic is less, the gNB can handle the entire assignment all by itself. However, for the burst traffic scenario, the optimal latency can be achieved by optimally dividing the traffic over the two paths.

D. RESOURCE ALLOCATION

For the allocation of resources, the bandwidth is divided into sections. The smallest resource section that can be assigned to a UE is a physical resource block or simply called as the resource block (RB). Several RBs can be allocated to a UE. The actual number of RBs that are allocated depends upon the QoS requirement of the running application(s) [45]. It is to be noted that in MR-DC, data traffic would have to be split over different paths leading to the receiving UE [33]. For instance, in EN-DC's Options 3 and 3x, the UE can simultaneously receive packets from both the eNB as well as the gNB. In order to guarantee the promised QoS, the number of RBs allocated from either of the nodes should be selected in a way such that the objectives of throughput and delay are guaranteed. For Options 3 and 3x, the packet schedulers at eNB and gNB, respectively, should work in coordination in order to provide the QoS requirements. It should also ensure fairness amongst the two traffic flows while leveraging the advantages of dual connectivity. However, the resource allocation in Option 3a is simpler since the all the data traffic is scheduled at the gNB only. As certain classes, for instance, voice over IP and video have very stringent QoS requirements, the rules for resource allocation should be clearly defined for the smooth operation [45]. However, there can not be one standard solution and network operators usually decide on resource allocation based on the traffic profiles and their RB utilization strategy. A few research works are also available that can act as guidelines to the operators [2]. Authors in [33], emphasize on keeping the delay difference between the two paths within the allowed threshold while designing the data flow control algorithms. The flow and performance routing problem is formulated as a mathematical programming model in [48]. Though, the problem is not specifically targeted at MR-DC but the problem formulation integrated various concepts like network model, delay model and revenue and penalty cost model. Through a centralized flow control method, the path selection is modelled as a maximal revenue problem subjected to the constraints of dropping penalties and tolerable delay [48]. While work in [48] is based on centralized flow control method, authors in [5] propose distributed framework to mitigate challenges like signaling overheads and increased processing costs. Mixed integer linear programming is used to maximize the minimum user throughput while considering the system constraint where every UE achieves throughput higher than its requirement [5]. The data flow algorithms discussed in recent works may not be very effective for supporting the URLLC services in MR-DC [49]. For instance, the non-ideal backhaul in Option 3x is not recommendable for services with strict requirement [50]. The mobile operators should consider data flow and scheduling algorithms that should be able to guarantee the stringent latency requirements and high throughput demands of various use cases in the emerging 5G network [49].

E. ENERGY EFFICIENCY

Power Handling at the UE is one of the issue that challenges dual connectivity. In MR-DC both gNB and eNB aggregate to serve the UE and thus both the LTE and 5G-NR air interfaces need to be activated at the same time to exchange UE data. This would increase the UE's power consumption. According to author's in [27], power consumption in NSA is more than double compared to standalone 5G when a video streaming application is considered [27]. Moreover, if the position of the UE intending to use dual connectivity is at a distance far from the serving cells then the received power may get strongly attenuated. This would result in unnecessary power impediments without any substantial dual connectivity gains like capacity and throughput [51]. The process of energy/power management and control is crucial due to limited power levels which may cause exhaustion and render the entire network inefficient. It is important to optimize idle mode energy, where a UE operates at low power even while being connected to more than one nodes. Since in MR-DC data traffic can be routed through different nodes, the energy required can be evaluated for transmission/reception by different nodes for selecting a particular route while considering idle mode operation. In this way, the dual connectivity may be exploited in an energy efficient manner. In terms of energy efficiency, Option 3a appears to have an advantage over Options 3 and 3x, since the UE is connected to only one gNB for data communication and can operate for DRX corresponding to the gNB. However, this advantage would dilute with increase in the number of beams in 5G networks to accommodate the increase in carrier frequency.

F. PACKET DUPLICATION FOR HIGHER RELIABILITY

In the standardization of 5G-NR upper layers, a duplicate transmission scheme on the PDCP layer has been discussed as a technology for improving the reliability of communications in radio access network (RAN) to achieve the 5G feature of URLLC. The fundamental principle includes independent transmission of same packet over redundant and uncorrelated links such that higher reliability is achieved at the receiver [29], [52]. As radio conditions can change dynamically due to radio quality, congestion in RAN, etc., it may not always be possible to achieve high-reliability communications via a single link. Fortunately, the duplication technique is directly applicable to the MR-DC architecture as it does not require excessive complexity at the radio level [6], [52]. It has therefore been discussed that frequency diversity be used to improve the communications reliability in RAN by applying the inherent multi-connectivity in MR-DC that uses multiple component carriers terminating to a single receiver. The radio protocol architecture for achieving this places multiple radio link control (RLC) layers below a single PDCP layer [22], [52]. Here, a packet processed and duplicated on the PDCP layer is transmitted via both MN and SN. The PDCP layer on the receiving side processes the packet that arrives earlier while discarding the delayed packet as a duplicate [22], [52]. Transmitting the same data over multiple radio links in this way enables data to be delivered over a good radio link in the event that the radio environment of the other radio link deteriorates. Thus, the MR-DC scheme makes up as a readily available solution for high-reliability communications which would be very effective for URLLC in 5G networks and beyond. However, EN-DC Option 3a lacks this advantage.

V. SPECTRUM SHARING

The frequency bands allocated to 5G are majorly in the mid and high band range [53]. Several users are expected to take benefits from these high frequency bands in near future as they offer to address dramatic increase in wireless traffic. It is expected that the traffic usage would increase to 49 EBs (Exa Bytes, 1 EB = 1,000,000 TB) per month in 2021 [54]. At the same time, it is also important to consider 5G operation in the lower frequency bands for the following advantages: (i) wide-area coverage such that the 5G users enjoy unprecedented access to the enhanced 5G use cases, (ii) improved mid and high frequency band spectrum utilization by enabling the most optimized scheduling while also addressing low latency use cases, (iii) cost-efficiency achieved by providing higher peak rates over larger coverage area [53]. However, most of the lower frequency bands are already in widespread use by the current LTE operators [53]. Given the high penetration of existing LTE users it is not viable to entirely re-frame the LTE carriers so that they can be used in NR. Thus, an efficient spectrum sharing is required that would not only allow the operation of 5G on existing low frequency bands but would also avoid impact on all end user either in LTE or 5G services.

International Telecommunication Union (ITU) identified spectrum below 6 GHz and above 24 GHz for 5G deployments [55]. 3GPP considers the 5G-NR interface according to the ITU and regional regulators guidelines. Several regional operators also included the C-band (3 GHz-5 GHz) for the first wave of 5G-NR deployments, along with the mmWave bands (26 GHz and 39 GHz) [56], [57]. Considering 5G operations only in a particular band would complicate the 5G deployments, especially the services that require good coverage, high data rate and low latency, simultaneously. Authors in [58] believe spectrum extension to be one of the main goals for 5G. Thus, for the initial 5G-NR deployments the 5G spectrum planning includes the C-band (4 GHz-8 GHz) along with the mmWave bands (spanning from 24 GHz to 29.5 GHz as well as spectrum in the 37 GHz to 43.5 GHz) [59].

The large available bandwidth in mmWaves would easily facilitate the time division duplexing (TDD) [56], [59]. However, the legacy LTE networks use frequency division duplexing (FDD) below 3 GHz [2]. If the coverage performance of the of the 1.8 GHz band (FDD) is compared to 3.5 GHz band (considering TDD and Massive MIMO), the UL coverage is observed to be much poorer compared to the DL as the UL power-spectral density at 3.5 GHz TDD is lower [56]. This creates a coverage gap as shown in Figure 7 and presents an opportunity such the extra resources in the existing LTE frequency band can be provided to 5G-NR operation as a complementary band in the TDD mode [56]. The uplink (UL) carrier from the sub GHz lower frequency LTE spectrum can be coupled to the carrier within the higher frequency mmWave band allocated for the 5G-NR UEs [59]. Such an arrangement manifests 2 UL carriers for 5G-NR UEs while in the legacy scenario only one UL carrier can be invoked. It is notable that in the above example (Figure 7) as well as in the legacy serving cell, there is only one DL carrier.

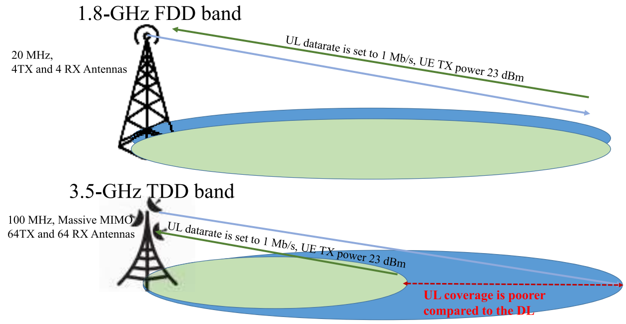


FIGURE 7. Frequency Sharing.

The utilization of LTE UL resources for 5G-NR UEs would guarantee the coverage performance improvement. The scenario is especially helpful for the cell edge 5G UEs for they can transmit the UL data using either of the two UL carriers. It is notable that the cell edge users suffer from poor service even after the application of coordinated signal processing [60]. As a result the coverage performance is substantially improved and UEs far from the base station can enjoy guaranteed data rates since the path loss at LTE frequency bands is much lower than at the 5G-NR mmWave frequencies [56]. The MR-DC scenario, inherently endorses this advantage as both the gNB and eNB work in coordination. We believe that the investigations of 4G-5G dual connectivity would assist to analyze not only the frequency saturation criterion but also performing the frequency relocation from 4G to 5G or even beyond (6G) when the frequency saturation criterion is exceeded.

A. REQUIREMENTS FOR 4G/5G FREQUENCY SHARING

To enable NR/LTE spectrum sharing, the relevant NR/LTE coexistence mechanisms needs to be specified. The key mechanisms include:

1) The management of the frequency sharing between LTE and NR should be efficient such that neither the LTE UL resources are underutilized nor the original LTE UEs suffer due to sharing impediments. Such a frequency division multiplexing between LTE and NR networks can be either static or dynamic in manner [56]. Static methodology can avoid interoperation coordination and thus, is more suitable when deployment for multiple vendors is considered. On the other hand, the dynamic sharing is more efficient and can be suitably used when same vendor is considered. Dynamic scheduling can achieve a higher spectral efficiency [56].
2) In order to fully utilize the spectral resources, the UEs from LTE and NR can be scheduled on orthogonal frequencies. Moreover, the UL scheduling at the supplementary frequency shared to NR by LTE networks should be at the granularity that facilitates the alignment of the physical resources to the LTE boundary to avoid any wastage.
3) The 5G-NR system that uses supplementary UL carrier from LTE while itself operating at TDD carrier would require frequency sensing for the selection of UL frequency and random access [61]. While the UEs at the cell edge can transmit random access preamble using supplementary LTE carrier, the cell-center UEs can be served better using higher-frequency TDD carriers. Thus, the UE is required to perform measurements with respect to TDD DL reference signal received power which can be subsequently compared to a threshold value to determine the UL carrier for initial access.
4) 5G is expected to support large number of services and applications requiring variable payload size, priority and quality [62]. 5G can be designed to perform either the quality based prioritized scheduling or service-oriented scheduling [56]. Such scheduling methodology can be integrated with the mechanism that allows the selection of the UL carrier from either of the two possibilities. For instance, as discussed before the URLLC service can select the LTE carrier for higher reliability while the applications requiring high data rate can use 5G carriers. This arrangement however would require quality awareness.

B. ADVANTAGES OF FREQUENCY SHARING

Frequency sharing in MR-DC can manifest the following advantages:

1) Extended coverage: The larger section of the teletraffic is comprised of the DL traffic which is expected to grow even further due to continuous rise in the popularity of video streaming [63]. Thus, for static TDD, the UL resources would be limited. In contrast, the UL spectrum would remain underutilized for LTE FDD bands since both DL and UL utilize the same amount of bandwidth [56]. Higher path loss at mmWave frequencies coupled with smaller portion of UL resources would limit the UL coverage in 5G-NR (expected to be working on TDD) [64]. The resulting small cells would not only increase cost impediments but would also create DL-UL coverage gap. According to authors in [64], the DL coverage can be 15.4 dB larger than the UL coverage. By employing LTE carriers for cell edge 5G UEs as shown in Figure 7, the network operators can reuse their existing LTE sites while at the same time they can address the challenge of coverage gap between DL and UL.
2) Mobility improvement: As explained aforesaid, the limited UL coverage would results in small cell size when 5G high frequencies are exclusively considered. Consequently, UE would experience several handovers in standalone 5G system. Moreover, for MR-DC but without spectrum sharing, several inter-RAT handovers would manifest as the coverage of 5G cell is substantially smaller than the LTE cell. By spectrum sharing, the coverage of 5G cell can be extended to become similar to that of the LTE cell [59]. Thus, both, the

inter-RAT and intra-RAT (for 5G) handovers can be substantially reduced and UE can enjoy improved mobility experience which is one of the key 5G requirements [18].

3) Efficient spectrum utilization: While the higher DL traffic can be scheduled on the higher mmWave frequencies, the UL traffic scheduling has two options. UL can either be scheduled on higher mmWave frequencies of 5G or low frequency bands that are borrowed from legacy LTE. The cell edge UEs can specifically benefit from the LTE bands as they portray reliable link quality over higher distance. The Internet-of-Things (IoT) devices can also gain from the supplementary LTE carriers as their UL data is usually small [65]. According to [56], the transmit power is low when small packet is transmitted over lower frequency LTE bands rather than the high frequency mmWaves. By scheduling cell edge and IoT devices on LTE bands, additional UE resources can be made available for other 5G users. Thus, the overall spectrum utilization is substantially improved in MR-DC.

4) Latency improvement: Under the concept of frequency sharing between LTE and NR, the URLLC devices for UL can be scheduled at the UL carrier at sub GHz frequencies that are shared from the legacy LTE networks. This would result in the availability of UL resources whenever a UL message arrives for URLLC device as legacy LTE mostly operates at FDD [2]. Thus, the latency that could have been caused because of the discontinuous UL resources of the TDD carrier in 5G-NR is substantially reduced [56]. Moreover, the overheads due to TDD switching can also be avoided.

VI. MIGRATION TO 5G AND BEYOND

On the road to 5G, the network operators would have to support several deployment options such that the service and monetary benefits are maximized in the transition period. According to [66], the key factors that affect the selection of the migration path are time to market, capital cost, operating cost, future compatibility, business trends and the current network conditions/ architecture. It would be reasonable for the LTE operators to deploy 5G NR at the existing LTE sites to simplify the deployment as well as to control the capital cost. However, SA NR is expected to operate at mmWaves frequencies that are much higher than current LTE frequencies. With higher path loss and spotty coverage it would be challenging for the operators to maintain the required quality (least equal to that of LTE) if the operation is restricted to the same site [67]. Moreover, NSA is not expected to have SA capabilities of finer QoS treatment and network slicing [68].

A. TRANSITION TO SA 5G

Many mobile operators have supported adoption of Options 3a and 3x for the initial 5G network deployment [69]. These options have gained attention due to their simplicity [69].

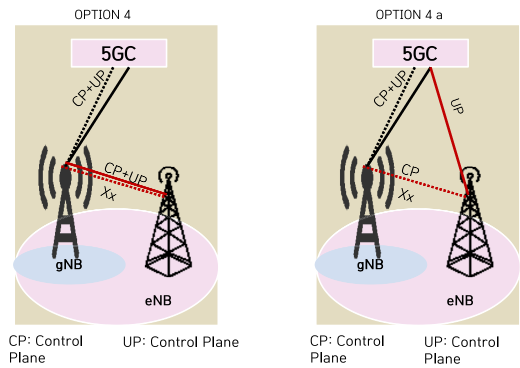


FIGURE 8. Deployment Options 4 and 4a.

Options 3a and 3x have encouraged rapid, straightforward and cost effective commercialization of 5G networks. Option 3x in particular offers gains due to throughput aggregation, packet duplication and mobility robustness while requiring the minimum investment in LTE. Moreover, Options 3x and 3a specifically can enable low latency user plane. Moving forward, as a natural step towards migration to 5G SA system, the networks operators are now considering Option 4 and 4a. These options also use tight integration between gNB and eNB which was fundamental to Option 3. However, in Options 4 and 4a, as shown in Figure 8, the gNB is considered as an MN and is connected to the 5G core (5GC). Option 4a is analogous to Option 3a and Option 4 is similar to Option 3 but with EPC changed to 5GC. Moreover, the gNB is responsible for anchoring all the control plane functionalities in Option 4 and 4a. While the main focus in Options 3/3a/3x was the enhancement of radio access, Options 4/4a are focused on enhancement of the core. The advantages of MR-DC (as discussed in this paper) in terms of throughput aggregation, packet duplication, beam failure fall back, frequency sharing and mobility management would be equally applicable to options 4/4a as they were to options 3/3a/3x. Similar to Option 3/3a/3x, transition to Options 4/4a would be subtle without immediate huge investments. Option 7 (Figure 3) also projects dual connectivity while upgrading the core, however, in this option the eNB acts as an MN and would require modifications as it connects to 5GC. Since, the network operators would ultimately like to move to 5G SA in future, investing in enhancement of LTE infrastructure may not be very effective approach. As the LTE frequencies are more conducive for wireless communications than mmWave frequencies, Option 7 may have some advantages from coverage and reliability point of view. However, it is too early to say what might prevail in deployment. While Options 3, 4 and 7 at the moment are very relevant NSA architectures, Option 2 (Figure 3) that presents a 5G SA system would be of much more relevance in future and has already received the industry consensus. Option 2 represents a 5G SA deployment with 5GC and 5G-NR. Transition directly to Option 2 would result in lower complexity and a direct evolution to the next generation mobile network. However, it would require a massive investment without guarantee of immediate returns. LTE and NR co-existence is expected to prevail at least for a while.

Option 3 has particularly presented the advantage of effective support for the eMBB services. The initial 5G network providers are focused on urban areas, hotspots, and high value areas that require rapid deployment to satisfy traffic and brand demands [59]. Options 3a and 3x are well suited for such scenarios requiring eMMB services. However, as 5G promises include the support for higher connectivity, reliable performance and enhanced user experience, the network operators would also be required to focus on these issues. It can be well augured that with popularity of vehicular automation, IoT and industrial IoT, the upcoming deployment options should emphasise more on URLLC and mMTC services along with eMMB. Thus, the network migrating from NSA to SA is expected to typically support several options at the same time [59] in order to encompass the variety in requirements.

From the perspective of business, while the current 4G networks are concentrated on business-to-customers (B2C) market, along with the development of 5G networks, 5G enabled business-to-business (B2B) services show rapid growth in terms of providing solutions for enterprise, industrial and government users. With this regard, the transition to 5G SA networks can have the following intermediate stages:

- Stage 1: by supporting dual connectivity, Options 3/3a/3x provide extended services for B2C market supported by 4G LTE networks.
- Stage 2: by smoothly combining Option 2 to the deployed Options 3/3a/3x, the network providers can support B2C and B2B market simultaneously.
- Stage 3: by replacing Options 3/3a/3x by Option 4, along with Option 2 the network providers can support 5GC based services for B2C and B2B markets.

Software Defined Network (SDN) [70] and Network Function Virtualization (NFV) [71] are introduced in 5GC to provide a service based architecture. NFV has been gaining popularity as one of the effective solutions for resource allocation and system scalability improvements in SA 5G networks [58]. While NSA focus on NR, the SA would be focusing on 5GC which is expected to provide the end to end network slicing using SDN and NFV. Network slicing would enable timely service and deployment for the diverse vertical industry [67]. According to authors in [72], the ongoing virtualization has potential to deliver cost ownership gains and energy savings. The virtualization in NSA is novel and has not yet delivered up to its potential, it is a very promising technology for SA 5G and beyond 5G due to ultra-dynamic slicing capabilities.

B. OPEN ISSUES AND RESEARCH DIRECTIONS

In this section we would like to bring about the research challenges that need to be addressed in order to derive the MR-DC advantages so that there full potential is achieved.

1) **Performance analysis of LTE and NR** In order to reveal a heterogeneous deployment of LTE- NR with effective performance in terms of different coverage, traffic and services, the following challenges can be addressed:
   - Overcoming path loss impediments: Despite the potential link budget gains of directional communication in 5G-NR, the reliability of a system purely relying on beamforming and operating in higher frequencies might be challenging, since the coverage is more sensitive to both time and space variations. Therefore, the analysis of optimum transmission method according to radio channel characteristics becomes important in the coexistence of 4G and 5G wireless networks.
   - Improving coverage: To improve network capacity, small cell NR base station using a high frequency can be installed in a hotspot area with high demands and with capacity shortage while at the same time the control can be supported by the existing LTE eNB, i.e., the macro base station. Contrary to this at some other sites, overlapping areas can be made available at the same time. The choice of hotspots or overlapping areas can be made based on traffic analysis such that the network is able to satisfactorily deliver quality requirements. Thus, the deployment solutions that incorporates real time traffic characteristics would be more helpful to the network operators.
   - Optimizing performance and power: NSA is a way to solve the capacity shortage problem by allowing a UE in the area to use the radio resources of both gNB and eNB. However, to enjoy higher capacity of communication, the UE might end up spending more power for maintaining dual connectivity. Thus, it is important to analyze criteria for UE's connectivity selection from not only performance improvement angle but power consumption perspective as well. It is to be noted that UE power saving has been identified as an important study item for 3GPP release 16 [73].
   - Dynamic switching: Spatial and temporal characterization of the wireless network show frequent unequal traffic load. To save network resources it would be discerning to explore the possibility of switching of gNB's that coexist with LTE eNBs in accordance with the dynamic adaptation to the traffic. Researchers can examine and explore criteria for dynamic traffic switching between LTE, MR-DC and NR based on artificial intelligence and learning algorithms.
   - Enhancing reliability: There is the problem of satisfying the extreme requirements related to URLLC in 5G systems [52]. It is thus, important to analyze the theoretic framework behind packet duplication and investigate recent enhancements made in the 4G-5G dual connectivity architecture for supporting packet duplication.

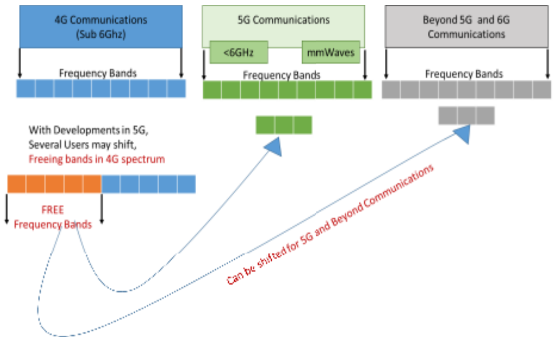


FIGURE 9. Frequency Reallocation.

- In order to derive the most benefit from MR-DC, the dynamic selection of suitable path can be performed based on the QoS requirements. The analysis may include the possibilities of loading/unloading of less/more congested paths in the real time. Artificial intelligence and machine learning could be instrumental in solving such problems.

2) **Spectrum resource usage and dual connectivity**
One of the techniques that encourages NR/LTE coexistence is the combination of low frequency carrier with the mmWave bands to improve both, the UL coverage as well as the mobility while simultaneously reducing the number of mmWave gNBs that are required for providing seamless coverage. To this regard the following problems can be addressed:
   - Synchronization for UL: There are several challenges when addressing spectrum usage by UEs in 4G-5G dual connectivity, such as the provision of UL synchronization, power control, UL access point switching, and so on. Thus, it is important to study methodology to obtain high level of synchronicity.
   - Strategy for spectrum selection: It is crucial to explore strategies for resource block allocation among sub GHz frequency and/or NR UL carriers so that the quality of service is maintained. Analysis of traffic use pattern can be performed to ascertain the UEs in coverage areas that would need 5G specific services.
   - Effect on 4G: To ensure that the already deployed LTE network performance is at its best while there is 4G-5G dual connectivity, it is important to investigate the changes in the amount of traffic received by 4G eNBs due to MR-DC. It would also be interesting challenge to predict the shift in traffic from LTE to 5G so that new gNBs can be accordingly added.

3) **Frequency usage and allocation** The operator should be able to provide required spectrum to 5G supporting devices and customers in compliance with the 3GPP standardization. The regional availability and technical feasibility of spectrum band should be investigated. To this regard following issues can be addressed:
   - Continuous bandwidth: It reduces device's power consumption and bandwidth wastage while increasing the bandwidth efficiency [66]. The study on re-farming of legacy spectrum and at what granularity would be helpful in avoiding the propagation of adversities of spectrum fragmentation in legacy networks to MR-DC.
   - Frequency usage efficiency: The study of frequency usage efficiency of 4G-5G dual connectivity is important for network operators to understand the advantages and limitations of either sharing spectrum with other operators or purchasing new bands from regulatory [66]. Mathematical modeling for the calculation of frequency efficiency using characteristics of actual traffic and comparative analysis with actual measurement results are needed to achieve the merits of dual connectivity in shared environment.
   - Frequency reallocation: The investigations of MR-DC can also assist to analyze the possibility of frequency relocation from 4G to 5G or even beyond (6G) when the frequency efficiency criterion is not achieved as shown in Figure 9. Thus, the frequency reallocation technique could be worked upon such that the optimum frequency efficiency is achieved over the legacy, existing and future frequency bands all considered together and not separately. This analysis can be based on current traffic trends along with the predicted future frequency demands.

C. FROM SA 5G TO 6G

On a high level, the technical points of difference between legacy LTE and SA 5G can be majorly considered as: (i) Scalable numerology and multiple sub carrier spacing (SCS) (ii) Adaptive bandwidth parts (iii) introduction of the SDAP layer and (iv) slot based scheduling.

SA 5G would support multiple numerologies depending on exponentially scalable SCS $\Delta f = 2 \times \mu \times 15$ KHz ($\mu = \{0, 1, 3, 4\}$). On the other hand, LTE supports SCS of only 15 KHz. From the radio perspective, 5G-NR has introduced a new access stratum sublayer in PDCP referred to as service data adaptation protocol (SDAP) sublayer [11] to support the radio bearer concept. The same is not available in LTE. Moreover, unlike LTE, the UE in 5G SA can be configured with a carrier bandwidth part (BWP). BWP defines the UE's operating bandwidth within the cell's operating bandwidth [74]. Several carrier BWPs can be configured for a UE, however, only one can be active on a given component carrier. SA 5G offers to support slot based scheduling as opposed to the subframe based scheduling in LTE. Basic transmission unit in SA 5G is a slot. While each LTE subframe is 0.5 ms, the slot scheduling would allow shorter possible scheduling unit based on the numerology used. These points make SA 5G much more flexible than existing LTE. We believe that these flexibilities would be very instrumental in moving to beyond 5G communications (B5G) and 6G communications. Researchers have already started exploring B5G and 6G communications as they would enable the expansion of wireless communications in the physical industries like, mining, transportation, electricity, utilities, shipping, etc. The Industrial IoT (IIoT) standardization has already been gaining momentum. The requirements from physical and industrial perspective would be much more different than the legacy human to human communications for instance, IIoT would require a reliability of 99.9999% which is much higher as opposed to current standards. Thus, B5G systems would require higher flexibility where the future of wireless communications is headed. One of the key technologies that offers to provide flexibility in B5G wireless networks is Artificial Intelligence (AI). With several small cells in 5G, distributing data and training sets for AI modeling would be an enormous overhead. The dual connectivity can be explored for B5G to address challenges of distributing the AI models and inference on those models across multiple locations.

VII. CONCLUSION

Though the exact use cases and advantages of 5G-NR are still being studied, the bandwidth growth and further development of the existing LTE use cases have clearly defined the trends of users' expectation. Several users are expected to take benefits from 5G-NR as it offers higher data rates. These users at the same time may be supported by existing LTE networks for better performance and high reliability. The 3GPP standard specifications have highlighted various deployment options that enable dual connectivity using both LTE and NR at the same time. In practice, if the network operator is using LTE system consisting of EPC and LTE eNBs nationwide, the dual connectivity that maximizes the role of the existing LTE system can be used for the initial commercialization of 5G networks. In this paper we have analyzed 4G-5G dual connectivity in detail while also exploring various possible research issues that need to be addressed.

REFERENCES

[1] *Road to 5G: Introduction and Migration*, GSMA, London, U.K., 2018.
[2] S. Park, M. Agiwal, H. Kwon, and H. Jin, "An evaluation methodology for spectrum usage in LTE—A networks: Traffic volume and resource utilization perspective," *IEEE Access*, vol. 7, pp. 67863–67873, 2019.
[3] S. C. Jha, K. Sivanesan, R. Vannithamby, and A. T. Koc, "Dual connectivity in LTE small cell networks," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Austin, TX, USA, Dec. 2014, pp. 1205–1210.
[4] *5G NR Requirements for Support of Radio Resource Management, (3GPP TS 38.133 Version 15.2.0 Release 15)*, document TS 38.133 V 15.2.0, 3GPP, 2018.
[5] V. F. Monteiro, D. A. Sousa, T. F. Maciel, F. R. P. Cavalcanti, C. F. M. E. Silva, and E. B. Rodrigues, "Distributed RRM for 5G multi-RAT multiconnectivity networks," *IEEE Syst. J.*, vol. 13, no. 1, pp. 192–203, Mar. 2019.
[6] *Radio Access Network; Evolved Universal Teestrial Radio Access (E-UTRA) and NR; Multi-Connectivity; Stage 2, (Release 15)*, document TS 37.340 V15.0.0.2, Mar. 2019.
[7] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
[8] J. Hoadley and P. Maveddat, "Enabling small cell deployment with HetNet," *IEEE Wireless Commun.*, vol. 19, no. 2, pp. 4–5, Apr. 2012.
[9] T. Nakamura, S. Nagata, A. Benjebbour, Y. Kishiyama, T. Hai, S. Xiaodong, Y. Ning, and L. Nan, "Trends in small cell enhancements in LTE advanced," *IEEE Commun. Mag.*, vol. 51, no. 2, pp. 98–105, Feb. 2013.
[10] I. Ashraf, F. Boccardi, and L. Ho, "SLEEP mode techniques for small cell deployments," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 72–79, Aug. 2011.
[11] M. G. Kibria, K. Nguyen, G. P. Villardi, K. Ishizu, and F. Kojima, "Next generation new radio small cell enhancement: Architectural options, functionality and performance aspects," *IEEE Wireless Commun.*, vol. 25, no. 4, pp. 120–128, Aug. 2018.
[12] R. Ratasuk and A. Ghosh, "Carrier aggregation and dual connectivity," Nokia Bell Labs, New Providence, NJ, USA, White Paper, 2017. [Online]. Available: https://www.its.bldrdoc.gov/media/66437/ratasuk_isart2017.pdf
[13] *Study on Small Cell Enhancements for E-UTRA and E-UTRAN: Higher Layer Aspects V12.0.0*, document TR 36.842, 3GPP, Dec. 2013.
[14] A. Mukherjee, "Optimal flow bifurcation in networks with dual base station connectivity and non-ideal backhaul," in *Proc. 48th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2014, pp. 521–524.
[15] *Scenarios and Requirements for Small Cell Enhancements for E-UTRA and E-UTRAN, V15.0.0 (Release 15)*, document TR 36.932, 3GPP, 2018.
[16] Working Group C, "Communication architectures and technologies," Wireless World Res. Forum, Zürich, Switzerland, Tech. Rep., 2014.
[17] K. Nguyen, M. G. Kibria, J. Hui, K. Ishizu, and F. Kojima, "Minimum latency and optimal traffic partition in 5G small cell networks," in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, Porto, Portugal, Jun. 2018, pp. 1–5.
[18] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1617–1655, 3rd Quart., 2016.
[19] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.
[20] Y. Liu, X. Fang, M. Xiao, and S. Mumtaz, "Decentralized beam pair selection in multi-beam millimeter-wave networks," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2722–2737, Jun. 2018.
[21] S. Sesia, I. Toufik, and M. Baker, *LTE—The UMTS Long Term Evolution: From Theory to Practice*. Hoboken, NJ, USA: Wiley, 2011.
[22] R. Antonioli, G. Parente, T. Maciel, F. Cavalcanti, C. Silva, and E. Rodrigues, "Dual connectivity for LTE-NR cellular networks," in *Proc. Anais de 35th Simpósio Brasileiro de Telecomunicaçes e Processamento de Sinais*, 2017, pp. 171–175.
[23] M.-S. Pan, T.-M. Lin, C.-Y. Chiu, and C.-Y. Wang, "Downlink traffic scheduling for LTE—A small cell networks with dual connectivity enhancement," *IEEE Commun. Lett.*, vol. 20, no. 4, pp. 796–799, Apr. 2016.
[24] *Radio Access Network: Study on New Radio Access Technology: Radio Access Architecture and Interfaces (Release 14)*, document TR 38.801 V14.0.0, 3GPP, Mar. 2017.
[25] A. Umesh, W. A. Hapsari, T. Uchino, T. Toeda, and H. Takahashi, "5G radio access network standardization trends," *NTT Docomo Tech. J.*, vol. 19, no. 3, Jan. 2018.
[26] *Study on Scenarios and Requirements for Next Generation Access Technologies*, document 38.913, 3GPP, 2018.
[27] *4G-5G Interworking RAN-level and CN-level Interworking*, Samsung, Seoul, South Korea, Jun. 2017.
[28] S. Singh, S.-P. Yeh, N. Himayat, and S. Talwar, "Optimal traffic aggregation in multi-RAT heterogeneous wireless networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, Kuala Lumpur, Malaysia, May 2016, pp. 626–631.
[29] N. H. Mahmood, M. Lopez, D. Laselva, K. Pedersen, and G. Berardinelli, "Reliability oriented dual connectivity for URLLC services in 5G new radio," in *Proc. 15th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Lisbon, Portugal, Aug. 2018, pp. 1–6.
[30] C. Rosa, K. Pedersen, H. Wang, E. Malkamaki, P.-H. Michaelsen, S. Barbera, T. Henttonen, and B. Sebire, "Dual connectivity for LTE small cell evolution: Functionality and performance aspects," *IEEE Commun. Mag.*, vol. 54, no. 6, pp. 137–143, Jun. 2016.
[31] S. Kang and S. Bahk, "Analysis of dual connectivity gain in terms of delay and throughput," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Jeju, South Korea, Oct. 2018, pp. 1218–1220.
[32] "Transport network support of IMT-2020/5G," Int. Telecommun. Union, Geneva, Switzerland, Tech. Rep. GST-RTN5G, 2018.

[33] T. Wigren, K. Lau, R. Delgado, and R. H. Middleton, "Delay skew packet flow control in wireless systems with dual connectivity," *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 5357–5371, Jun. 2018.
[34] *Sharetechnotes*. Accessed: Jul. 2020. [Online]. Available: https://www.shatetechnotes.com
[35] K. I. Pedersen, G. Pocovi, J. Steiner, and S. R. Khosravirad, "Punctured scheduling for critical low latency data on a shared channel with mobile broadband," in *Proc. IEEE 86th Veh. Technol. Conf. (VTC-Fall)*, Toronto, ON, Canada, Sep. 2017, pp. 1–6.
[36] K. I. Pedersen, G. Berardinelli, F. Frederiksen, P. Mogensen, and A. Szufarska, "A flexible 5G frame structure design for frequency-division duplex cases," *IEEE Commun. Mag.*, vol. 54, no. 3, pp. 53–59, Mar. 2016.
[37] P. Marsch, I. D. Silva, O. Bulakci, M. Tesanovic, S. E. El Ayoubi, T. Rosowski, A. Kaloxylos, and M. Boldi, "5G radio access network architecture: Design guidelines and key considerations," *IEEE Commun. Mag.*, vol. 54, no. 11, pp. 24–32, Nov. 2016.
[38] *Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) Protocol Specification, (Release 13)*, document TS 36.321(v 13.9.0), 3GPP, Jul. 2018.
[39] *Sharetechnotes*. Accessed: Jul. 2020. [Online]. Available: https://www.sharetechnote.com/html/Throughput_LTE_Research.html
[40] H. Jin, W. T. Toor, B. C. Jung, and J.-B. Seo, "Recursive pseudo-Bayesian access class barring for M2M communications in LTE systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 8595–8599, Sep. 2017.
[41] M. Polese, M. Giordani, M. Mezzavilla, S. Rangan, and M. Zorzi, "Improved handover through dual connectivity in 5G mmWave mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 2069–2084, Sep. 2017.
[42] T. Mumtaz, S. Muhammad, M. I. Aslam, and N. Mohammad, "Dual connectivity-based mobility management and data split mechanism in 4G/5G cellular networks," *IEEE Access*, vol. 8, pp. 86495–86509, 2020.
[43] M. K. Maheshwari, M. Agiwal, N. Saxena, and A. Roy, "Directional discontinuous reception (DDRX) for mmWave enabled 5G communications," *IEEE Trans. Mobile Comput.*, vol. 18, no. 10, pp. 2330–2343, Oct. 2019.
[44] M. Nekovee and R. Rudd, "5G spectrum sharing," 2017, *arXiv:1708.03772*. [Online]. Available: http://arxiv.org/abs/1708.03772
[45] A. H. Ali and M. Nazir, "Radio resource management: The vital subject for evolution to 5G," in *Proc. Int. Symp. Wireless Syst. Netw. (ISWSN)*, Lahore, Pakistan, Nov. 2017, pp. 1–7.
[46] R. Kausar, Y. Chen, K. K. Chai, L. Cuthbert, and J. Schormans, "QoS aware mixed traffic packet scheduling in OFDMA-based LTE-advanced networks," in *Proc. UBICOMM*, 2010, pp. 53–58.
[47] J. Branke, K. Deb, K. Miettinen, and R. Slowinski, *Multiobjective Optimization: Interactive and Evolutionary Approaches*. Berlin, Germany: Springer-Verlag, 2008.
[48] F. Y.-S. Lin, Y.-F. Wen, L.-W. Fang, and C.-H. Hsiao, "Resource allocation and multisession routing algorithms in coordinated multipoint wireless communication networks," *IEEE Syst. J.*, vol. 12, no. 3, pp. 2226–2237, Sep. 2018.
[49] M. Peng, C. Wang, V. Lau, and H. V. Poor, "Fronthaul-constrained cloud radio access networks: Insights and challenges," *IEEE Wireless Commun.*, vol. 22, no. 2, pp. 152–160, Apr. 2015.
[50] A. Maeder, A. Ali, D. Chandramouli, S. Chandrashekar, A. Bedekar, A. F. Cattoni, L. Du, M. Hesse, C. Sartori, and S. Turtinen, "A scalable and flexible radio access network architecture for fifth generation mobile networks," *IEEE Commun. Mag.*, vol. 54, no. 11, pp. 16–23, Nov. 2016.
[51] M. A. Lema, E. Pardo, O. Galinina, S. Andreev, and M. Dohler, "Flexible dual-connectivity spectrum aggregation for decoupled uplink and downlink access in 5G heterogeneous systems," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 11, pp. 2851–2865, Nov. 2016.
[52] J. Rao and S. Vrzic, "Packet duplication for URLLC in 5G dual connectivity architecture," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Barcelona, Spain, Apr. 2018, pp. 1–6.
[53] Ericsson. *Spectrum Sharing—A Better Way to Build 5G*. Accessed: Jul. 2018. [Online]. Available: https://www.ericsson.com/en/networks/offerings/5g/sharing-spectrum-with-ericsson-spectrum-sharing
[54] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016-021*, CISCO, San Jose, CA, USA, Feb. 2017.
[55] M. Nekovee, "Opportunities and enabling technologies for 5G spectrum sharing," 2019, *arXiv:1902.07573*. [Online]. Available: http://arxiv.org/abs/1902.07573
[56] L. Wan, Z. Guo, Y. Wu, W. Bi, J. Yuan, M. Elkashlan, and L. Hanzo, "4G/5G spectrum sharing: Efficient 5G deployment to serve enhanced mobile broadband and Internet of Things applications," *IEEE Veh. Technol. Mag.*, vol. 13, no. 4, pp. 28–39, Dec. 2018.
[57] *WF on Band Specific UE Channel Bandwidth*, document 3GPP TSG-RAN WG4-NR Meeting #2, NTT DOCOM, Jun. 2017. [Online]. Available: https://portal.3gpp.org/ngppapp/CreateTdoc.aspx?mode=view&contributionId=805939
[58] A. Al-Dulaimi, S. Mumtaz, S. Al-Rubaye, S. Zhang, and C.-L. I, "A framework of network connectivity management in multi-clouds infrastructure," *IEEE Wireless Commun.*, vol. 26, no. 3, pp. 104–110, Jun. 2019.
[59] *5G Implementation Guidelines*, GSMA, London, U.K., Jul. 2019.
[60] K. Alexandris, C.-Y. Chang, K. Katsalis, N. Nikaein, and T. Spyropoulos, "Utility-based resource allocation under multi-connectivity in evolved LTE," in *Proc. IEEE 86th Veh. Technol. Conf. (VTC-Fall)*, Toronto, ON, Canada, Sep. 2017, pp. 1–6.
[61] *Discussion on Scenario 2 for LTE NR DC With UL Sharing From UE Perspective*, document 3GPP R1-1712166, HSilicon and Huawei, Sophia Antipolis, France, Rep. TSG-RAN WG1 Meeting 90, 2017.
[62] M. S. Ali, E. Hossain, and D. I. Kim, "LTE/LTE—A random access for massive machine-type communications in smart cities," *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 76–83, Jan. 2017.
[63] *IMT Traffic Estimates for the Years 2020 to 2030*, Standard ITU-RM.2370-0, 2015.
[64] L. Wan, Z. Guo, and X. Chen, "Enabling efficient 5G NR and 4G LTE coexistence," *IEEE Wireless Commun.*, vol. 26, no. 1, pp. 6–8, Feb. 2019.
[65] M. Agiwal, M. K. Maheshwari, and H. Jin, "Power efficient random access for massive NB-IoT connectivity," *Sensors*, vol. 19, no. 22, p. 4944, Nov. 2019.
[66] A. Zakeri, N. Gholipoor, M. Tajallifar, S. Ebrahimi, M. R. Javan, N. Mokari, and A. R. Sharafat, "E2E migration strategies towards 5G: Long-term migration plan and evolution roadmap," 2020, *arXiv:2002.08984*. [Online]. Available: http://arxiv.org/abs/2002.08984
[67] G. Liu, Y. Huang, Z. Chen, L. Liu, Q. Wang, and N. Li, "5G deployment: Standalone vs. Non-standalone from the operator perspective," *IEEE Commun. Mag.*, vol. 58, no. 11, pp. 83–89, Nov. 2020.
[68] *System Architecture for the 5G System, Release 15*, document TR 23.501, 3GPP, Mar. 2020.
[69] *5G NR: A New Era for Enhanced Mobile Broadband*, Mediatek, Hsinchu, Taiwan, 2018.
[70] F. Hu, Q. Hao, and K. Bao, "A survey on software-defined network and OpenFlow: From concept to implementation," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 2181–2206, 4th Quart., 2014.
[71] Y. Li and M. Chen, "Software-defined network function virtualization: A survey," *IEEE Access*, vol. 3, pp. 2542–2553, 2015.
[72] M. Masoudi *et al.*, "Green mobile networks for 5G and beyond," *IEEE Access*, vol. 7, pp. 107270–107299, 2019.
[73] *NR Study on UE Power Saving (Release 16)*, document TR 38.840, 3GPP, Oct. 2018.
[74] *NR; NR and NG-RAN Overall Description*, document TS 38.300 V16.1.0, 3GPP, Dec. 2017. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3191

MAMTA AGIWAL received the Ph.D. degree from Sungkyunkwan University, South Korea, and the M.S. degree in digital electronics from Visvesvaraya Technological University, India. She worked as a Postdoctoral Researcher with Hanyang University, South Korea, and an Associate Professor with the CMR Institute of Technology, India. She is currently working as an Assistant Professor with Sejong University, South Korea. Her research interests include 5G and beyond 5G wireless communications, energy efficiency, and the Internet of Things.

HYEYEON KWON received the B.S. degrees in computer science and statistics and the M.S. and Ph.D. degrees in computer science from Chungnam National University, Daejeon, South Korea. She is currently a Principle member of Engineering Staff with the Electronics and Telecommunications Research Institute, South Korea. Her current research interests include mobile data mining and analysis for spectrum management.

SEUNGKEUN PARK (Member, IEEE) received the B.S. and M.S. degrees in applied statistics from Korea University, Seoul, South Korea, in 1991 and 1993, respectively, and the Ph.D. degree in information communication engineering from the University of Chungbuk, Cheongju, South Korea, in 2004. He is currently a Principal Member with the Electronics and Telecommunications Research Institute, Daejeon. His current research interests include communication theory and spectrum management.

HU JIN (Senior Member, IEEE) received the B.E. degree in electronic engineering and information science from the University of Science and Technology of China, Hefei, China, in 2004, and the M.S. and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2006 and 2011, respectively.

From 2011 to 2013, he was a Postdoctoral Fellow with The University of British Columbia, Vancouver, BC, Canada. From 2013 to 2014, he was a Research Professor with Gyeongsang National University, Tongyeong, South Korea. Since 2014, he has been with the Division of Electrical Engineering, Hanyang University, Ansan, South Korea, where he is currently an Associate Professor. His research interests include medium-access control and radio resource management for random access networks and scheduling systems considering advanced signal processing and queuing performance.