Fog-RAN Enabled Multi-Connectivity and Multi-Cell Scheduling Framework For Ultra-Reliable Low Latency Communication

BINOD KHAREL¹, (Student Member, IEEE), ONEL L. ALCARAZ LÓPEZ¹, (Member, IEEE), NURUL HUDA MAHMOOD¹, (Member, IEEE), HIRLEY ALVES¹, (Member, IEEE), MATTI LATVA-AHO¹, (Senior Member, IEEE)

¹Centre for Wireless Communications (CWC), University of Oulu, Finland. {binod.kharel, onel.alcarazlopez, nurulhuda.mahmood, hirley.alves, matti.latva-aho}@oulu.fi.

Corresponding author: Binod Kharel (e-mail: binod.kharel@oulu.fi).

This research has been financially supported by Academy of Finland 6Genesis Flagship (grant 318927).

ABSTRACT Ultra-Reliable Low Latency Communication (URLLC) is a newly introduced service class targeting emerging Internet-of-Things (IoT) application scenarios. This paper assumes an interference-limited Fog Radio Access Network (F-RAN) setup composed of multiple Remote Radio Heads (RRHs) equipped with multiple antennas serving single-antenna users. F-RAN facilitates collaborative solutions while reducing delay by pushing the network capabilities beyond the edge. By leveraging diversity, RRHs may cooperate through silencing, reducing interference, or joint transmission strategies such as maximal ratio transmission. We derive closed-form outage probability expressions and attain their diversity gain. We validate the derived analytical results through extensive numerical simulations. Furthermore, we propose a mini-slots-based scheduling framework to serve URLLC users within their fixed latency budget. In an interference-limited regime with the proposed scheduling framework, we show that a performance gain is superior when RRHs cooperate compared to when they do not. We briefly discuss the cost of reliability, i.e., the impact on the system’s average sum throughput under cooperation. Moreover, numerical results verify that cooperating transmission schemes boost transmission reliability with a significantly improved latency performance at the cost of reduced system’s average sum throughput.

INDEX TERMS diversity, F-RAN, maximal ratio transmission, machine-type communication, multi-connectivity, reliability, scheduling, silencing, ultra-reliable low latency communication.

I. INTRODUCTION

Many emerging applications in the domain of the Internet of Things (IoT) require efficient machine-type communications (MTC) to interconnect wirelessly without the need for human intervention [1]. Indeed, fifth-generation (5G) wireless communication systems have categorized MTC to address two main services: massive MTC (mMTC) and URLLC [2]. In this work, we focus on the URLLC, which is the service class aiming to meet stringent reliability and latency requirements in 5G New Radio (NR) [3]. In this context, several mission-critical applications require URLLC services, e.g., factory automation [4], process automation [4], intelligent transportation systems [4], automated guided vehicles (AGV) [5] and smart grids [6]. These applications require high reliability (e.g., $10^{-5}$ to $10^{-9}$ outage probability depending upon the application) and simultaneously latency budgets of few milliseconds [6].

We can study URLLC from two aspects: Ultra-Reliable Communication (URC) and Low-Latency Communication (LLC). URC technique is possible via diversity, including Multi-Connectivity (MC) and robust physical layer design. In contrast, LLC via flexible numerology, grant free instant uplink, and fast processing [7]. Some significant challenges and critical technology components related to ultra-reliability...
are enhanced control channel reliability, link adaptation, interference mitigation, and coexistence with other higher data rate services such as Enhanced Mobile Broadband (eMBB) [8]. However, the interplay between the diverse URLLC requirements makes the physical layer design of such systems highly complex [9].

MC is identified as a critical URLLC enabler in 5G systems. MC adopts spatial diversity, which can be enabled via centralized processing in the F-RAN architecture to ensure ultra-reliability through collaborative solutions [10]. In the Fog Radio Access Networks (F-RAN), a large amount of signal processing and computing is performed in a distributed manner. At the same time, Access Points (APs), e.g., RRHs or fog APs, integrate radio frequency and signal processing functionalities which are beneficial for interference management and radio resource allocations [11]. APs are connected to a Centralized Unit (CU) in the fronthaul via high-speed optical fiber links that can support low latency and high capacity communication, both of which lead to improving network performance [12]. F-RAN has four modes of operations: global centralized mode, locally distributed mode, high power node mode, and device-to-device mode. We focus on the F-RAN enabled global centralized mode where collaborative radio signal processing and radio resource management functions are implemented centrally at the Baseband Unit (BBU) pool [13].

The objective of this paper is to establish an MC framework through RRH cooperation strategies enabled by F-RAN. The proposed F-RAN model presented in this paper represents indoor Industrial IoT (IIoT) scenarios and is similar to the one considered by 3GPP Release 16 [14]. For example, the factory setting scenario consists of CU with BBU as a controller that enables the F-RAN with storage and computation capabilities at the edge and the User Equipment (UE) as an actuator. We exploit MC through the Maximum Ratio Transmission (MRT) scheme or reducing interference through silencing in interference-limited downlink cellular networks.

A. RELATED LITERATURE

Much literature related to URLLC appeared after the seminal work by Popvski [15]. Furthermore, Popvski et al., discuss the communication theoretic principles for supporting URLLC in [16], [17], such as the use of various diversity resources, design of packets, and access protocols. In [18], the authors investigate different diversity sources, e.g., time, frequency, and space, to meet the challenging requirements of URLLC. However, due to critical latency and bandwidth constraints of IoT applications, the diversity gain from the frequency and time domains are very limited to overcome the possible deep fading caused by shadowing [19]. Hence, spatial diversity may often be more attractive.

The authors in [20] detail the comprehensive study and importance of spatial diversity in wireless communication. In spatial diversity, multiple antennas are physically separated from each other transmit to the user. MC adopts spatial diversity where more than one connection jointly serves the UE. The basic idea is to send replicas of the same message through more than one link. If one of them is decoded successfully, then the packet is received [21]. For instance, [22] studies MC as one of the sources to improve reliability, enabling the transmission of redundant data through multiple links using standard diversity schemes like joint decoding, selection combining, and maximum ratio combining. In [23], [24], the authors study the potential of diversity and interference management techniques to achieve ultra-reliability operation. Furthermore, [25] proposes to interface diversity where each interface is based on different technologies to offer URLLC without intervention in baseband/physical layer design. Also, cooperative diversity emerges as a workable alternative to direct communication [26]. On the other hand, such reliability gain comes at the cost of transmission of redundant packets, leading to an increase in radio-resource consumption [27]. In [28], the authors propose an MC concept in edge RAN to reduce mobility-related link failures and cell-edge degradation. Likewise, [29] discusses centralized RAN technology to support URLLC, lowering the traffic latency by a functional split of the central and radio units. The primary key difference between cloud and fog networking is that data is processed at the edge nodes in the latter. In contrast, in the former, the data is processed at CU.

Furthermore, the flexible 5G frame structure and numerology in 5G NR provides a scalable and configurable air interface design to support low-latency transmissions with the mini-slots down to two symbols in duration as defined by 3GPP Release 15 [30]. Different works [31]–[34] adopt the concept of mini-slot transmission to satisfy latency constraints for URLLC while proposing different scheduling and resource utilization techniques in multi-cell 5G networks.

In this study, we also adopt the concept of mini slots based URLLC user scheduling framework to reduce the data Transmission Time Interval (TTI) and schedule the URLLC users within the target latency budget. The main difference with the works above is that we analyze the system performance in an interference-limited scenario with different modes of cooperative transmission schemes enabled by F-RAN, and discuss the gain of cooperation. Although this paper exhibits some similarities in the system model assumptions concerning [22], herein we pay close attention to spatial diversity methods considering interference and cooperation of spatially distributed RRHs enabled by the F-RAN.

B. OUR CONTRIBUTIONS

From physical layer URLLC system design perspective there are three major technical challenges. First, minimizing the system overhead in terms of channel access, user scheduling and allocation of resources. Second, outage probability should be reduced in first transmission as retransmission can affect the latency. Third, the transmission of URLLC packets as soon as possible to reduce latency. Therefore, this paper aims to design an F-RAN framework to serve a low-latency UE with an ultra-reliable transmission scheme in an
interference-limited downlink scenario. The main contributions of this article are summarized as follows:

- We identify and analyze diversity techniques like silencing and MRT schemes for URC in an interference-limited downlink communications scenario. A network with fully connected RRHs, i.e., a global centralized mode, is assumed where all the RRHs are connected to CU where RRHs might cooperate or may not their transmission through BBU pool. We analytically formulate the system reliability in terms of outage probability when RRHs cooperate through BBU via silencing or MRT. Exact closed-form outage probability expressions are attained, and we verify them through numerical simulations. We show the superiority of MRT over silencing, mainly supported by the fact that a smaller number of cooperating RRHs are required to achieve the desired reliability.

- In addition, we study the asymptotic behaviour of outage probability for MRT and silencing scheme in terms of diversity gain. The analysis shows that MRT scheme provides \( k + 1 \) times higher diversity gain compared to silencing scheme under cooperation.

- We propose a mini slots-based scheduling framework to serve URLLC users within their latency budget. We discuss performance gains of different transmission schemes in terms of percentage of user served, mean latency, and the average number of transmission/cooperation through extensive computer-level simulations. We show that adopted MRT and silencing cooperation schemes not only reduce the interference on F-RAN but also enhance the network performance by successfully serving a higher number of URLLC users active at random positions within the given area.

- Finally, we discuss the impact of cooperation on overall network performance by analyzing the trade-off between average system sum throughput and reliability through computer-level simulations for MRT and silencing schemes. We show that with cooperation, both schemes achieve URC operation while the system’s average sum throughput gradually decreases.

Section II introduces the system model. Section III details the investigated transmission schemes, and provides signal-to-interference ratio (SIR) derivations. Section IV provides the outage probability formulation and throughput reliability trade-off analysis, while Section V presents the scheduling framework. Section VI shows performance evaluation and results. Finally, Section VII concludes the paper.

Table 1 provide the list of acronyms used throughout in this paper in an alphabetical order. Throughout this paper, the superscript \( H \) denotes the complex conjugate transpose, and a boldface lowercase letter denotes a column vector. \( \mathbb{C} \) denotes complex domain. \(|a|\) denotes the norm of complex valued vector \( a \), i.e., \(|a| = \sqrt{a^H a} \), while \( E[\cdot] \) denotes the expectation operation and \( \Gamma(\cdot) \) denotes the gamma function. Uppercase and lowercase letters denote random variables (RVs) and their realizations. The probability density function (PDF) and cumulative distribution function (CDF) of RV \( X \) is denoted by \( f_X(x) \) and \( F_X(x) \), respectively. For ease of reference, Table 2 summarizes all the important notations used throughout this paper.

| Symbol | Description |
|--------|-------------|
| AP     | Access Point |
| BBU    | Baseband Unit |
| CC     | Chase Combining |
| CSI    | Channel State Information |
| CU     | Centralized Unit |
| eMBB   | Enhanced Mobile Broadband |
| F-RAN  | Fog Radio Access Network |
| HARQ   | Hybrid Automatic Request |
| IloT   | Industrial Internet of Things |
| IoT    | Internet of Things |
| LLC    | Low Latency Communication |
| mMTC   | Massive MTC |
| MTC    | Machine Type Communication |
| NR     | New Radio |
| RTT    | Round Trip Time |
| SIR    | Signal-to-Interference Ratio |
| TTI    | Transmission Time Interval |
| UE     | User Equipment |
| URC    | Ultra-Reliable Communication |
| URLLC  | Ultra-Reliable Low Latency Communication |
| 3GPP   | 3rd Generation Partnership Project |
| 5G     | Fifth Generation Networks |

Table 2. Summary of notations

| Notation | Description |
|----------|-------------|
| \( N \)  | Total number of RRHs |
| \( k \)  | Total number of cooperating RRHs |
| \( e \)  | Set of cooperating RRHs defined as \( 0 \ldots k \) |
| \( i \)  | Set of interfering RRHs defined as \( k + 1 \ldots N \) |
| \( M_t \) | Number of transmit antennas |
| \( d_{i,j} \) | Distance from RRH \( i \) to UE \( j \) |
| \( w_{i,j} \) | Transmit beamformer vector from RRH \( i \) to UE \( j \) |
| \( h_{i,j} \) | Channel gain vector from RRH \( i \) to UE \( j \) |
| \( s_j \)  | Information symbol of UE \( j \) |
| \( \mathbb{C} \) | Circular symmetric complex Gaussian noise vector |
| \( \alpha \) | Path loss exponent |
| \( \gamma \) | Signal-to-interference ratio |
| \( \theta \) | Signal-to-interference ratio threshold |
| \( D \)  | Diversity gain |
| TP      | Average system sum throughput |
| \( R \)  | Average sum rate of all active users in the system |
| \( K_{\text{max}} \) | Maximum number of available RRHs |
| \( k_{\text{min}} \) | The minimum number of RRHs in cooperation |
| \( N_{\text{serv}} \) | Total number of URLLC users to be served |
| \( N_{\text{re-ct}} \) | Total number of URLLC users in retransmission |
| \( \tau \) | Critical latency threshold |
II. SYSTEM MODEL

We consider an F-RAN system model as illustrated in Fig. 1, where all RRHs are connected to an edge cloud consisting of a baseband unit (BBU) pool via fronthaul links with the high-bandwidth and low latency communication. We assume that all RRHs are using the same spectrum resources (i.e., time and frequency) when transmitting to their corresponding UEs. In the setup, we assume that \( N + 1 \) RRHs, i.e., \( RRH_0 \), \( RRH_1 \), \ldots, \( RRH_N \) are spatially distributed in a given area \( \mathcal{A} \subset \mathbb{R}^2 \). The link between \( RRH_0 \) and \( UE_0 \) is considered as a typical link. Meanwhile, the link from other \( RRH_i \) to \( UE_0 \) for all \( i \in \{ 1, \ldots, N \} \) can be interfering or cooperating links to the typical link. We consider cooperating links when several \( RRHs \) cooperate to the typical link in transmission to the \( UE_0 \) while the rest of the \( RRHs \) are interfering to the typical link.

We assume that \( RRHs \) are equipped with \( M_t \) transmit antennas and UE with single antenna. Let \( d_{i,j} \) and \( w_{i,j} \in \mathbb{C}^{M_t \times 1} \) denote the distance and transmit beamformer from \( RRH_i \) to \( UE_j \), respectively. Furthermore, let \( h_{i,j} \in \mathbb{C}^{M_t \times 1} \) denote the channel vector (small-scale fading) between \( RRH_i \) to \( UE_j \). Then, the received signal vector at \( j^{th} \) UE is expressed as

\[
\mathbf{r}_j = d_{0,j}^{-\alpha} \mathbf{h}_{0,j}^H \mathbf{w}_{0,j} s_j + \sum_{c=1}^{k} d_{c,j}^{-\alpha} \mathbf{h}_{c,j}^H \mathbf{w}_{c,j} s_j + \sum_{i=k+1}^{N} d_{i,j}^{-\alpha} \mathbf{h}_{i,j}^H \mathbf{w}_{i,j} s_i + \mathbf{z}_j, \tag{1}
\]

where \( \alpha \) is a path loss exponent, \( s_j \) is the information symbol, and \( \mathbf{z}_j \) is circular symmetric complex Gaussian noise vector. Also (1) assumes that \( k \) closest \( RRHs \) are cooperating with the transmission of the typical link, and rest of the the \( RRHs \) are interferers. Note that \( k \) is the number of cooperating \( RRHs \). Furthermore, assuming a dense \( RRHs \) deployment, we can neglect the noise impact since interference is dominant [35]. Hence, in the subsequent analysis of this paper we neglect the noise.

III. TRANSMISSION SCHEMES AND SIR DERIVATION

We analyze the typical link performance when F-RAN operates under the following two modes when serving the UEs:

- **without cooperation**: As shown in Fig 2(a), the interfering \( RRHs \) do not cooperate with the typical link through F-RAN BBU, and each \( RRH \) transmits to its own UE. Under this condition, each UE experiences full interference from neighboring \( RRHs \).

- **cooperation**: In this mode, neighboring \( RRHs \) cooperate with the typical link through the F-RAN BBU pool to serve the desired UE to fulfill user-centric objectives like high reliability and low latency. The main benefit of the F-RAN framework enables multi-point transmission and cooperative solutions which coordinate data transmission to the typical UE pushing network capabilities beyond the edge, reducing the delay, leveraging less burden on the BBU pool.

In the cooperation mode, we consider one of the following transmission strategies to serve the desired UE:

- **Silencing**: F-RAN silences some of the strongest interfering \( RRHs \) to mitigate the interference at the reference user as shown in Fig. 2(b). At a silenced \( RRH \), the transmission of the data channel, control channel and reference signals are completely turned off. This enhances the reliability of typical link as silencing or muting a strong interferer helps to boost the SIR in the victim cell, and it has been proposed for 5G [36], [37].

- **MRT**: Several \( RRHs \) jointly coordinate the transmission to the reference UE, as shown in Fig. 2(c). The same packets are sent from multiple \( RRHs \) independently, which gives redundancy against fading, blocking, or radio link failure [38]. The joint transmission through MRT gives the best reception reliability enabling multi-point MC. Also, sending the same message through independent transmissions from multiple \( RRHs \) saves resources, which are very limited due to the stringent latency constraints in URLLC.

Similar to the work in [39] for the above-mentioned transmission schemes, we assume that each data packet is trans-
mitted once to enable low latency transmission (i.e., one-shot transmission) is considered from RRH to the desired UE.). The optimal beamformers are computed by conjugate beamforming, assuming local Channel State Information (CSI) at each RRH which avoids the use of explicit control and extra signalling information exchange among cooperating RRHs. Thus, the aforementioned schemes do not incur signalling and hence are feasible for upcoming 5G systems and services [40]. We focus our attention on the above transmission schemes to enhance the reliability of the considered F-RAN system. Next, we derive the SIR received at the desired UE for the above-mentioned transmission schemes.

Silencing
Here, the RRH₀ is the only one serving the UE₀. At the same time, k cooperating RRHs remain silent during the corresponding transmission slots, and the remaining non-cooperating RRHs are interfering. Then from (1) the received signal at UE₀ can be expressed as

\[
\mathbf{r}_0 = d_{0,0}^{\alpha} \mathbf{h}_{0,0}^H \mathbf{w}_{0,0} s_0 + \sum_{i=k+1}^{N} d_{i,0}^{\alpha} \mathbf{h}_{i,0}^H \mathbf{w}_{i} s_i. \tag{2}
\]

Thus, the SIR received at desired UE₀ using the silencing scheme is given by

\[
\gamma_s = \frac{E_{s_0} \left[ |\mathbf{h}_{0,0}^H \mathbf{w}_{0,0} s_0 d_{0,0}^{\alpha}|^2 \right]^2}{\sum_{i=k+1}^{N} d_{i,0}^{\alpha} \mathbf{h}_{i,0}^H \mathbf{w}_{i} s_i d_{i,0}^{\alpha} \mathbf{h}_{i,0}^H \mathbf{w}_{i} s_i d_{i,0}^{\alpha}} = \frac{\left| \mathbf{h}_{0,0}^H \mathbf{w}_{0,0} s_0 \right|^2}{\sum_{i=k+1}^{N} \left| \mathbf{h}_{i,0}^H \mathbf{w}_{i} s_i \right|^2} \tag{3}
\]

where (a) comes after algebraic simplification and assuming \(E[|s_0|^2] = E[|s_i|^2] = 1\), we obtain (3).

MRT
Here the RRH₀ and k cooperating RRHs transmit simultaneously the same signal to the UE₀ while the remaining non-cooperating RRHs are interfering. Similar to [41], we assume that full channel state information (CSI) is available at centralized BBU. Then from (1) the received signal vector at UE₀ can be written as

\[
\mathbf{r}_0 = \sum_{c=0}^{k} d_{c,0}^{\alpha} \mathbf{h}_{c,0}^H \mathbf{w}_{c,0} s_c + \sum_{i=k+1}^{N} d_{i,0}^{\alpha} \mathbf{h}_{i,0}^H \mathbf{w}_{i} s_i. \tag{4}
\]

Therefore, following the same procedure to attain (3), SIR perceived at UE₀ using MRT is given by

\[
\gamma_{\text{MRT}} = \frac{\sum_{c=0}^{k} \left| \mathbf{h}_{c,0}^H \mathbf{w}_{c,0} s_c d_{c,0}^{\alpha} \right|^2}{\sum_{i=k+1}^{N} \left| \mathbf{h}_{i,0}^H \mathbf{w}_{i} s_i d_{i,0}^{\alpha} \right|^2} \tag{5}
\]

IV. PERFORMANCE ANALYSIS
In this section, we derive an analytical expression of the outage probability for the considered transmission schemes. Similar to the works in [16], [22], [42], we analyze the network performance in terms of outage probability as the critical performance metric evaluating the CDF of SIR derived earlier in Section III and defined as the probability that UE is in outage if

\[
F_{\gamma_o}(\theta) = \mathbb{P}\{\gamma_o < \theta\} \tag{6}
\]

where \(\rho \in \{S, \text{MRT}\}\) and \(\theta\) is an SIR threshold. In this paper, we focus on the outage model defined in (6) because this definition meets the definition of reliability in the context of URLLC, which states that a system can assure the URLLC requirements only if it can satisfy the required reliability level within the target latency budget [43]. For this reason, we analyze the system performance in terms of reliability, i.e., 1 − outage probability) for both considered transmission schemes. We also study the impact on system performance with throughput-reliability trade-off when RRHs cooperate to attain the URLLC target reliability level for the desired user. Instead of the end-to-end latency, in this work, we consider data transmission latency which is the time duration of generated URLLC packet delivered successfully to the intended UE.

A. OUTAGE PROBABILITY
We proceed to calculate the outage probability by evaluating the CDF of SIR from (6) as

\[
F_{\gamma_o}(\theta) = \mathbb{P}\{\gamma_o < \theta\} = \mathbb{P}\{X < \frac{Y}{\theta}\} = \mathbb{P}\{Y > \frac{X}{\theta}\} = \int_0^\infty f_X(x) \int_{x/\theta}^{\infty} f_Y(y) dy dx, \tag{7}
\]

where \(X \triangleq \sum_{c=0}^{k} d_{c,0}^{\alpha} \mathbf{h}_{c,0}^H \mathbf{w}_{c,0} \) for all \(c \in \{0, \ldots, k\}\), and \(Y \triangleq \sum_{i=k+1}^{N} \left| \mathbf{h}_{i,0}^H \mathbf{w}_{i} s_i \right|^2 d_{i,0}^{\alpha} \) are the corresponding SIR for desired links and interfering links obtained in (3) and (5). We assume that the channels coefficients of \(\mathbf{h}_{i,j}\) are independent, complex Gaussian normally distributed RVs with zero mean and variance, which includes the effects of path loss i.e., \(\mathbf{h}_{i,j} \sim CN(0, d_{i,j}^{-\alpha})\). Therefore, it is easy to show that the numerator \(X\) are independent RVs, which follows a gamma distribution with PDF given by

\[
f_X(x) = \frac{x^{M_t-1} \exp \left( -x d_{c,0}^{\alpha} \right)}{\Gamma(M_t) (d_{c,0}^{\alpha})^{M_t}}, \tag{8}
\]
where $\Gamma(\cdot)$ denotes gamma function [44]. Similarly, we can show that for the interference power from any $i^{th}$ RRH to the UE0, i.e., $y_i \triangleq |h_i^H b_{i,0}|^2 d_{i,0}^{-\alpha}$, is an exponential distributed RV with PDF

$$f_{y_i}(y_i) = d_{i,0}^\alpha \exp (-y_i d_{i,0}^\alpha).$$  \hspace{1cm} (9)

**Remark 1.** In the scenario where $N > 1$ and $d_{i,0} \neq d_{n,0}$ for $i \neq n$, the distribution of $Y$ can be obtained as [45]

$$f_Y(y) = \sum_{i=k+1}^{N} \left[ - \frac{y}{d_{i,0}^\alpha} \prod_{n=k+1}^{N} \frac{d_{n,0}^{-\alpha}}{d_{i,0}^{-\alpha} - d_{n,0}^{-\alpha}} \right].$$  \hspace{1cm} (10)

The validity of (10) is illustrated in Fig. 3. We see that analytical expression matches the corresponding simulation results perfectly. Note that for $N = 1$, $f_Y(y)$ is computed with (9). As shown, the total interfering signal power $Y$ does not change much with increasing $N$ because multi-antenna beamforming to serve the desired UE benefits the desired link while the additional interferers are so far away that their contribution to $Y$ is almost negligible in the considered setup. After obtaining the corresponding PDF for desired and interference links, outage probabilities for silencing and MRT schemes are derived.

**Theorem 1.** The outage probability when UE0 is served through silencing scheme is given by

$$F_{\gamma_{\text{Sil}}}(\theta) = \sum_{i=k+1}^{N} \Psi_i \left[ 1 + \frac{1}{\theta} \left( \frac{d_{i,0}}{d_{c,0}} \right)^\alpha \right]^{-M_t},$$  \hspace{1cm} (11)

with $\Psi_i = \prod_{n=k+1}^{N} \frac{d_{n,0}^{-\alpha}}{d_{i,0}^{-\alpha} - d_{n,0}^{-\alpha}}$.

**Proof.** Please refer to Appendix A. □

**Theorem 2.** Under the F-RAN cooperating mode, when the RRHs cooperate to serve the UE0 through MRT, the outage probability is given as

$$F_{\gamma_{\text{MRT}}}(\theta) = \sum_{i=k+1}^{N} \prod_{c=0}^{k} \frac{\Psi_i}{\left( 1 + \frac{1}{\theta} \left( \frac{d_{i,0}}{d_{c,0}} \right)^\alpha \right)^{M_t}},$$  \hspace{1cm} (12)

with $\Psi_i = \prod_{n=k+1}^{N} \frac{d_{n,0}^{-\alpha}}{d_{i,0}^{-\alpha} - d_{n,0}^{-\alpha}}$.

**Proof.** Please refer to Appendix A. □

1) Diversity gain

The main goal of URLLC systems is to enhance reliability, therefore outage probability curve provides benchmark for performance evaluation. However, to further investigate the behaviour of outage probability expression at infinite SIR, we assess the diversity gain that can be achieved from Silencing and MRT schemes. Similar to [46], we define diversity gain $D$ as

$$D_{\rho} = - \lim_{\theta \to \infty} \frac{\log F_{\gamma_{\rho}}(\theta)}{\log \theta},$$  \hspace{1cm} (13)

where $F_{\gamma_{\rho}}(\theta)$ is the outage probability obtained in (11) and (12) and $\rho \in \{\text{S}, \text{MRT}\}$ for silencing and MRT schemes.

Next, the diversity gain for silencing and MRT schemes at infinite SIR regime is investigated as follows:

- **Silencing:** Using (13) the diversity gain for silencing scheme is expressed as

$$D_S = - \lim_{\theta \to \infty} \frac{\log F_{\gamma_{\text{Sil}}}(\theta)}{\log \theta},$$

\begin{align*}
&\overset{(a)}{=} - \lim_{\theta \to \infty} \frac{\log \sum_{i=k+1}^{N} \Psi_i \left( \frac{d_{i,0}}{d_{c,0}} \right)^{-\alpha M_t} \left( \frac{1}{\theta} \right)^{-M_t}}{\log \theta}, \\
&\overset{(b)}{=} - \lim_{\theta \to \infty} \left( \frac{\log \left( \frac{1}{\theta} \right)^{-M_t}}{\log \theta} + \log \sum_{i=k+1}^{N} \Psi_i \left( \frac{d_{i,0}}{d_{c,0}} \right)^{-\alpha M_t} \right), \\
&\overset{(c)}{=} M_t. \hspace{1cm} (14)
\end{align*}

where (a) follows from CDF of silencing scheme (11) and taking the binomial approximation of $(1 + a b) M_t$ and (b) follows immediately after taking the limit.

- **MRT:** Following the similar procedure for attaining (14), the diversity gain for MRT is obtained as

$$D_{\text{MRT}} = - \lim_{\theta \to \infty} \frac{\log F_{\gamma_{\text{MRT}}}(\theta)}{\log \theta},$$

\begin{align*}
&\overset{(a)}{=} - \lim_{\theta \to \infty} \frac{\log \sum_{i=k+1}^{N} \left( \frac{1}{\theta} \right)^{-(k+1)M_t} \prod_{c=0}^{k} \Psi_i \left( \frac{d_{i,0}}{d_{c,0}} \right)^{-\alpha M_t}}{\log \theta}, \\
&\overset{(b)}{=} M_t(k+1). \hspace{1cm} (15)
\end{align*}

Note that comparing (14) and (15), MRT asymptotic diversity gain is $k + 1$ times higher than silencing scheme.
2) Complexity analysis
The derived closed form expression in (11) and (12) include simple algebraic scalar operation of product, addition and division terms. Note that the complexity of MRT increases with the number of cooperating RRHs $k$, while the complexity of silencing does not scale up. However, the diversity gain of MRT is also scale up $k+1$ compared to silencing as shown in (14) and (15). All in all, the derived closed-form solutions do not involve heavy computational efforts at BBU of proposed F-RAN network with respect to number of RRHs and users.

B. THROUGHPUT-RELIABILITY TRADE-OFF
We define throughput-reliability trade-off as the cost the considered reliability-oriented system model has to bear on its average system sum throughput when increasing the number of cooperating RRHs to meet URLLC service requirements. We implement a system-level simulation to study the fundamental trade-offs between reliability and average system sum throughput for the considered transmission schemes in the F-RAN cooperating mode. We assess the reliability of the typical link, which is given by

$$\text{Rel}_{p}= 1 - F_{\gamma_{p}}(\theta), \quad (16)$$

where $\gamma_{p}$ is the SIR evaluated from (3) and (5) with $\rho \in \{S, MRT\}$ for silencing and MRT schemes, respectively. Meanwhile, the average system sum throughput is given by

$$TP_{p} = R_{p}^{ref} + R_{p}^{A}, \quad (17)$$

where $R_{p}^{ref}$ is the achievable reliable rate at the typical link, and $R_{p}^{A}$ is the corresponding average sum rate of non-cooperating RRHs active users. Then,

$$R_{p}^{ref} = \text{Rel}_{p}^{ref} \log_{2}(1 + \theta),$$

$$R_{p}^{A} = \mathbb{E} \left[ \sum_{i=k+1}^{N} \log_{2}(1 + \gamma_{p}^{i}) \right]. \quad (18)$$

Finally, from (17) and (18) we have that

$$TP_{p} = \text{Rel}_{p}^{ref} \log_{2}(1 + \theta) + \mathbb{E} \left[ \sum_{i=k+1}^{N} \log_{2}(1 + \gamma_{p}^{i}) \right].$$

(19)

Note that when the number of cooperating RRHs $k$ increases, the average system sum throughput decreases.

V. SCHEDULING FRAMEWORK FOR URLLC USERS
This section proposes a scheduling framework suitable for URLLC users. 3GPP introduces a flexible frame structure for 5G NR with different options to shorten TTI as compared to LTE [30], we consider mini slots of duration 0.125 ms for transmission in short TTI to meet low latency requirements as in [47]. We use different transmission modes for serving the typical URLLC users. Fig. 4 shows the time frame structure of downlink URLLC with the different transmission modes. As shown, the initial latency budget of 6 TTIs for the URLLC user leaves a sufficient time budget for a maximum of two re-transmission since a hybrid automatic request (HARQ) round trip time consumes 2 TTIs [48]. In the case of re-transmission(s), multiple re-transmitted packets are combined using chase combining (CC), boosting the desired signal power [49].

Note that there are multiple active URLLC users simultaneously competing for the resources. Hence, the overall objective of the given scheduling framework is to serve the maximum number of active URLLC users within the latency budget. Three different performance metrics are considered to evaluate the performance of the scheduling framework:

1) Percentage of users served: Accounts for all the successfully delivered packet.
2) Mean latency: An average latency (difference between start time and end time) of all successful users.
3) Average number of transmissions/cooperations: Average total number of transmissions for all successful users in case of re-transmissions with HARQ. While in the case of cooperating schemes, we average the total number of cooperating RRHs.

To elaborate more on the concept of the proposed scheduling framework and serving strategies for URLLC users. Fig. 5 shows a high-level flow diagram of the scheduling framework at the CU and typical URLLC users, respectively. The scheduling framework works as follows: at each time slot, CU checks if there are any active URLLC users in the network. The data corresponding to the active URLLC users is added to the transmission buffer. If the maximum available RRHs $K_{\text{max}}$ is smaller than the number of active URLLC users to be served $N_{\text{serve}}$ it reschedules newly active URLLC users ($N_{r,s}$) from the current slot by updating the corresponding latency budget to the next scheduling time. CU has information regarding the availability of RRHs, which are currently not serving any URLLC user, including pending
Silencing is achieved without re-transmission and cooperation. The CU may serve the URLLC user using one out of four transmission modes.

1) Without cooperation and re-transmission (Baseline):
   - Only the typical RRH transmits to the UE. CU serves desired UE if SIR is greater than or equal to the given SIR threshold. Otherwise, drops the corresponding UE.

2) Re-transmission with CC-HARQ: In this non-cooperating mode, CU checks for any pending HARQ re-transmission. If any pending HARQ re-transmission, then it updates the chase combined SIR ($\gamma_{cc}$) of intended UE. In this framework, chase combined SIR denoted as $\gamma_{cc}$ is computed combining the SIR from the previous transmission as $\gamma_{cc} = \sum_{f=1}^{T_x} \gamma_f$, where $\gamma_f$ is the SIR at current time slot for re-transmission users, and $T_x$ is the maximum allowable re-transmission. CU serves intended UE if $\gamma_{cc} \geq \theta$, otherwise it drops the respective users if only if the available latency budget, i.e., $t_{ava}$ is less than the critical latency threshold ($\tau$). If $t_{ava}$ $\geq \tau$, CU re-schedules the intended UE by updating the latency budget to the next scheduling time.

3) Silencing: In this cooperating mode, CU forces some of the available cooperating RRHs to remain silent during the transmission slots and computes the minimum RRH ($k_{min}$) required to meet the given SIR threshold from (3) as $\gamma_S \geq \theta$. CU serves respective URLLC users if $\gamma_S \geq \theta$ and $k_{min} \leq K_{max} - N_{serv}$, where $K_{max}$ is the maximum number of available RRHs and $N_{serv}$ is the number of RRH assigned to URLLC users during that transmission slots. Otherwise, it drops the respective URLLC users.

4) MRT: In this cooperating mode, some of the cooperating RRHs jointly cooperate to transmit to the UE during that transmission slot. CU computes the minimum RRH ($k_{min}$) required to satisfy the given SIR threshold from (5) as $\gamma_{MRT} \geq \theta$. CU serve respective URLLC users if $\gamma_{MRT} \geq \theta$ and $k_{min} \leq K_{max} - N_{serv}$ during that transmission slots, otherwise it drops the respective URLLC users.

In this scheduling framework, there are two practical constraints for the RRHs cooperation:

- C1: In a given time instant, RRHs are available for cooperation only if not serving any active URLLC users.
- C2: Cooperating RRHs are selected based on their distance to intended active URLLC users. If available, the closest RRH is given priority, and so on.

The detailed performance evaluation for the above mentioned metrics are presented and analyzed in Section VI-D.

**VI. NUMERICAL RESULTS**

This section presents numerical and simulation results related to the system performance under the discussed transmission schemes. In the analysis, we set $\alpha = 3.5$ based on practical radio propagation measurement in the industrial setups [50]. The plots presented in the Section VI-C and VI-D are generated using a system-level simulation, where we adopt the parameters summarized in Table 3. We generate Rayleigh fading channel realizations and random users locations over run time. In particular, our Monte Carlo simulations comprise $10^7$ runs such that the performance results are accurate for targeted reliability of up to $1 - 10^{-5} = 0.99999$ (five 9's) corroborating our analytical expressions provided in Section IV.

**A. IMPACT OF THE NUMBER OF TRANSMIT ANTENNAS**

Fig. 6 shows the outage probability of both MRT and silencing schemes as a function of the SIR threshold for the
In this section, we present the throughput-reliability trade-off results discussed in Section IV-B. As shown in Fig. 8, there is a clear trade-off between average system sum throughput and reliability when we increase cooperating RRHs to achieve the target reliability. For example, as the number of cooperating RRHs increases, there is substantial improvement in reliability for both schemes. However, the average system sum throughput is reduced. Meanwhile, the average system sum throughput comparatively reduces in the case of MRT as compared to silencing. For example, with MRT, the average system sum throughput for \( k = 1 \) and \( M_t = 2 \) is around 67 bps/Hz, and silencing is around 69 bps/Hz. This is because cooperating RRHs in silent mode reduces interference factors for reference UE and other active UEs, boosting their corresponding SIRs and increasing average system sum throughput. With the MRT schemes, interference with other active users persists. However, the average system throughput does not increase with \( k \) under silencing because ongoing transmission to the other users is interrupted, impacting overall system sum throughput. Furthermore, we observe that with an increase in the transmit antennas at the RRHs, average system sum throughput and reliability increase for the same number of cooperating RRHs. This suggests that the number of cooperating RRHs can be reduced by increasing the number of transmit antennas per RRH. This reduction in the number of cooperating RRHs improves throughput, thereby reducing overhead consumption. Furthermore, MRT operates at the URLLC target regime with smaller \( k \) for the same \( M_t \) compared to silencing, thus allowing better utilization of the network resources.

**D. PERFORMANCE ANALYSIS WITH THE SCHEDULING FRAMEWORK**

This section shows the performance analysis of the scheduling framework as discussed in Section V. We consider a 0.25 km² communication area where RRHs are randomly deployed, and URLLC users get activated with a specific activation rate. We set the number of URLLC users to be equal to 10. We test the performance metrics of the proposed scheduling framework as a function of transmit antennas \( M_t \), activation rate, and SIR threshold \( \theta \) for the different
transmission schemes. Note that we only consider the time duration of successfully delivered packets in mean latency analysis, i.e., dropped packets are not considered.

Fig. 9 shows the performance of the scheduling framework in terms of the percentage of users served metric as a function of the number transmit antennas \( M_t \), activation rate, and \( \theta \). Note that the percentage of users served increases with \( M_t \) while decreasing with the activation rate and \( \theta \). More transmit antennas per RRH increases the desired received power at intended UE, boosting SIR and more URLLC users satisfying the target SIR. However, when the activation rate grows, more URLLC users are activated in the system; thus, URLLC users demand more resources to satisfy the target, SIR. HARQ scheme performs poorly when activation rate is more than 0.1 since more users in re-transmission hold resources during the re-transmission, which results in unavailability of resources to the other new users; hence few numbers of the user is served. We observe that at low, e.g., \( \theta < 0 \) dB, cooperating and re-transmission with HARQ schemes perform similarly regarding the percentage of users served. Cooperating transmission schemes with MRT and silencing performance seems better compared to schemes without cooperation and re-transmission. However, we observe that the silencing scheme performs better than MRT in all three cases because in multi-user URLLC scenarios, silencing some RRHs can reduce overall interference in the system, which favors the SIR at all the active URLLC users. However, MRT benefits intended users with MC while interference continues for other active users in the system.

In Fig. 10, it is interesting to observe that the mean latency of baseline and the cooperating scheme is approximately 0.125 ms since there is a single-shot transmission for the intended active URLLC user. However, the mean latency of re-transmission with HARQ decreases when increasing \( M_t \) because most users satisfy the target \( \theta \) at first transmission by exploiting antenna array gains. Meanwhile, the mean latency for re-transmission with the HARQ scheme increases with increasing activation rate and \( \theta \). Due to more active URLLC users, the demand for re-transmission is more to satisfy target \( \theta \), thus latency increases.

Fig. 11 evaluates the performance of the scheduling framework in terms of a number of transmission/cooperation required for HARQ and cooperation schemes of all successful users as a function of transmit antennas \( M_t \), activation rate, and \( \theta \). The results reveal that the number of transmissions with the HARQ scheme slightly decreases with increasing \( M_t \). A large \( M_t \) promotes high diversity gain, so the target \( \theta \) is satisfied in a first transmission. Meanwhile, we see that number of cooperating RRHs is more in silencing than MRT because the diversity gained from silencing is less than MRT. Thus, more cooperating RRHs need to be silenced to fulfill the target, SIR. As shown, when \( \theta \) increases, the average number of cooperation decreases, but when \( \theta > 8 \) [dB], the number of cooperation rises because users demand more cooperating RRHs to attain the target, SIR. These results suggest that with the proposed framework, one-shot transmission with cooperation seems more appropriate than re-transmission and without cooperation to support more URLLC users and overcome the hard latency deadline.

### VII. CONCLUSION

We studied the performance of the F-RAN-enabled framework for URLLC with MRT and silencing diversity schemes in the interference-limited downlink scenarios. We attained accurate closed-form expressions for the outage probability for MRT and silencing schemes. The analysis presented herein demonstrates that the outage probability performance improves with the number of cooperating RRHs, transmitting antennas at the RRHs, and diversity schemes. We studied the asymptotic behaviour of outage probability expressions for MRT and silencing scheme in terms of diversity gain. The result showed that MRT provides \( k+1 \) times higher diversity gain compared with silencing. Furthermore, we proposed a mini-slots-based scheduling framework to serve...
URLLC users under hard latency deadlines. The analysis showed that cooperating schemes like MRT and silencing serve more URLLC users under the hard latency deadline than with no cooperation and re-transmission with HARQ schemes. Besides, we evaluated the impact on average system sum throughput when increasing the number of cooperating RRHs to ensure URLLC. The results showed that MRT and silencing schemes enhance the system performance in terms of reliability but at the cost of reduced average system sum throughput. Overall, the extensive numerical investigations showed that with the RRHs cooperation, diversity schemes like MRT and silencing could achieve URLLC in an F-RAN
network. As future work, we intend to extend our analysis for efficient multiplexing of users with heterogeneous quality-of-service requirements.

**APPENDIX A: PROOF OF THEOREM 1**

In silencing scheme, $X$ is distributed as (8). Then from the definition of moment generating function (MGF), we get the MGF of $X$ as

$$M_X(t) = \mathbb{E}[\exp(tx)] = \frac{1}{1 - td_{d_0}} M_t.$$  \hspace{1cm} (20)

In fact, $\int_0^\infty f_Y(y)dy = 1$, thus, we can obtain $\sum_{i=0}^N \Psi_i = 1$. Using (7) and (10) we have that

$$F_{\gamma_S}(\theta) = \int_0^\infty f_X(x) \int_0^\infty \sum_{i=1}^N \Psi_i \exp \left( -\frac{x}{\theta d_{d_0}} \right) dydx$$

$$= \int_0^\infty f_X(x) \sum_{i=1}^N \Psi_i \left[ \exp \left( -\frac{x}{\theta d_{d_0}} \right) \right] dx$$

$$= \sum_{i=1}^N \Psi_i \mathbb{E} \left[ \exp \left( -\frac{1}{\theta d_{d_0}} x \right) \right].$$  \hspace{1cm} (21)

Substituting (20) with $t = \frac{-1}{\theta d_{d_0}}$ into (21), we attain (11), thus, concluding the proof.

**APPENDIX B: PROOF OF THEOREM 2**

In MRT, we have several RRHs jointly transmitting to the desired UE. The MGF of the sum of independent RVs $X_c$ can be expressed through a product of MGFs of each RVs $X_c$ as

$$M_{X_c}(t) = \prod_{c=0}^k M_{X_c}(t) = \prod_{c=0}^k \mathbb{E}[\exp(tx_c)]$$

$$= \prod_{c=0}^k \left( \frac{1}{1 - td_{d_0}} M_t \right).$$  \hspace{1cm} (22)

Then, using (7) and (10) we have that

$$F_{\gamma_{MRT}}(\theta) = \int_0^\infty f_{X_c}(x_c) \int_0^\infty \sum_{i=1}^N \Psi_i \exp \left( -\frac{y}{\theta d_{d_0}} \right) dydx_c$$

$$= \int_0^\infty f_{X_c}(x_c) \sum_{i=1}^N \Psi_i \left[ \exp \left( -\frac{x_c}{\theta d_{d_0}} \right) \right] dx_c$$

$$= \sum_{i=1}^N \Psi_i \mathbb{E} \left[ \exp \left( -\frac{1}{\theta d_{d_0}} x_c \right) \right].$$  \hspace{1cm} (23)

Substituting (22) with $t = \frac{-1}{\theta d_{d_0}}$ into (23), we reached (12), thus, concluding the proof.

**References**

[1] N. H. Mahmood et al., “White paper on critical and massive machine type communication towards 6G,” 6G Research Visons, no. 11, 2020, http://jultika.tulu.fi/files/sbnl9789526226781.pdf.

[2] M. Series, “IMT vision-framework and overall objectives of the future development of IMT for 2020 and beyond,” Recommendation ITU, vol. 2083, 2015.

[3] H. Tullberg et al., “The METIS 5G system concept: Meeting the 5G requirements,” IEEE Commun. Mag., vol. 54, no. 12, pp. 132–139, December 2016.

[4] B. Holfeld et al., “Wireless communication for factory automation: an opportunity for LTE and 5G systems,” IEEE Commun. Mag., vol. 54, no. 6, pp. 36–43, June 2016.

[5] M. Gharba et al., “5G enabled cooperative collision avoidance: System design and field test,” in IEEE 18th Int. Symp. on a World of Wireless, Mobile and Multimedia Netw. (WoWMoM), June 2017, pp. 1–6.

[6] P. Schulz et al., “Latency critical IoT applications in 5G: Perspective on the design of radio interface and network architecture,” IEEE Commun. Mag., vol. 55, no. 2, pp. 70–78, February 2017.

[7] M. Bennis, M. Debbah, and H. V. Poor, “Ultra-reliable and low-latency wireless communication: Tail, Risk, and Scale,” Proc. IEEE, vol. 106, no. 10, pp. 1834–1853, Oct 2018.

[8] G. Pocovi et al., “Achieving ultra-reliable low-latency communications: Challenges and envisioned system enhancements,” IEEE Netw., vol. 32, no. 2, pp. 8–15, March 2018.

[9] H. Ji et al., “Ultra-reliable and low-latency communications in 5G downstream: Physical layer aspects,” IEEE Wireless Commun., vol. 25, no. 3, pp. 124–130, June 2018.

[10] G. J. Sutton et al., “Enabling technologies for ultra-reliable and low latency communications: From PHY and MAC layer perspectives,” IEEE Commun. Surveys. Tuts., pp. 1–1, 2019.

[11] H. Zhang et al., “Fog radio access networks: Mobility management, interference mitigation, and resource optimization,” IEEE Wireless Commun., vol. 24, no. 6, pp. 120–127, 2017.

[12] N. Alliance, “Further study on critical C-RAN technologies,” Next Generation Mobile Netw., 2015.

[13] M. Peng et al., “Fog-computing-based radio access networks: issues and challenges,” IEEE Netw., vol. 30, no. 4, pp. 46–53, 2016.

[14] 3GPP, “3GPP specification group radio access network: Study on channel model for frequencies from 0.5 to 100 GHz (release 16),” 3GPP TR 38.901 v15.2.0, Tech. Rep., 2019.

[15] P. Popovski, “Ultra-reliable communication in 5G wireless systems,” in 1st Int. Conf. 5G Ubiquitous Connectivity, Nov 2014, pp. 146–151.

[16] P. Popovski et al., “Wireless access in ultra-reliable low-latency communication (URLLC),” IEEE Trans. Commun., vol. 67, no. 8, pp. 5783–5801, Aug 2019.

[17] P. Popovski et al., “Wireless access for ultra-reliable-low-latency communication: Principles and building blocks,” IEEE Netw., vol. 32, no. 2, pp. 16–23, March 2018.

[18] D. Ohmann et al., “Diversity trade-offs and joint coding schemes for highly reliable wireless transmissions,” in IEEE 84th Veh. Technol. Conf. (VTC-Fall), 2016, pp. 1–6.

[19] J. Zeng et al., “Enabling ultrareliable and low-latency communications under shadow fading by massive MU-MIMO,” IEEE Internet Things J., vol. 7, no. 1, pp. 234–246, 2020.

[20] S. N. DIGGAVI et al., “Great expectations: the value of spatial diversity in wireless networks,” Proc. IEEE, vol. 92, no. 2, pp. 219–270, 2004.

[21] M. Suer et al., “Multi-connectivity as an enabler for reliable low latency communications—an overview,” IEEE Commun. Surveys. Tuts., vol. 22, no. 1, pp. 156–169, 2020.

[22] A. Wolf et al., “How reliable and capable is multi-connectivity?” IEEE Trans. Commun., vol. 67, no. 2, pp. 1506–1520, Feb 2019.

[23] B. Soret et al., “Interference coordination for dense wireless networks,” IEEE Commun. Mag., vol. 53, no. 1, pp. 102–109, 2015.

[24] G. Pocovi et al., “Signal quality outage analysis for ultra-reliable communications in cellular networks,” in IEEE Globecom Workshops (GC Wkshps), 2015, pp. 1–6.

[25] J. J. Nielsen, R. Liu, and P. Popovski, “Ultra-reliable low latency communication using interface diversity,” IEEE Trans. Commun., vol. 66, no. 3, pp. 1322–1334, March 2018.

[26] O. L. A. López et al., “Ultra reliable short message relaying with wireless power transfer,” in IEEE Int. Conf. Commun. (ICC), 2017, pp. 1–6.

[27] M. Centenaro et al., “System-level study of data duplication enhancements for 5G downlink URLLC,” IEEE Access, vol. 8, pp. 565–578, 2020.

[28] F. B. Tesema et al., “Mobility modeling and performance evaluation of multi-connectivity in 5G intra-frequency networks,” in IEEE Globecom Workshops, Dec 2015, pp. 1–6.

[29] G. Montasser, S. Malei, and M. Reisslein, “Cloud-RAN in support of URLLC,” in IEEE Globecom Workshops, Dec 2017, pp. 1–6.

[30] 3GPP, “5G NR: physical channels and modulation,” TS 38.211, ver. 15.2.0, Jun 2018.

[31] A. Karimi et al., “5G centralized multi-cell scheduling for URLLC: Algorithms and system-level performance,” IEEE Access, vol. 6, pp. 72253–72262, 2018.
[32] N. H. Mahmood et al., “On the resource utilization of multi-connectivity transmission for URLLC services in 5G new radio,” in *IEEE Wireless Commun. Netw. Conf. Workshop (WCNWC)*, 2019, pp. 1–6.

[33] G. Pocovi, K. I. Pedersen, and P. Mogensen, “Joint link adaptation and scheduling for 5G ultra-reliable low-latency communications,” *IEEE Access*, vol. 6, pp. 28 912–28 922, 2018.

[34] A. A. Esswie and K. I. Pedersen, “Opportunistic spatial preempive scheduling for URLLC and eMBB coexistence in multiuser 5G networks,” *IEEE Access*, vol. 6, pp. 38 451–38 463, 2018.

[35] X. Zhang and J. G. Andrews, “Downlink cellular network analysis with Handbook of mathematical functions,” *Proceedings of the IEEE*, vol. 106, no. 11, pp. 2074–2096, 2018.

[36] B. Kharel et al., “Achieving ultra-reliable communication via CRAN-enabled diversity schemes,” in *Eur. Conf. Netw. Commun. (EuCNC)*, June 2019, pp. 320–324.

[37] 3GPP. “Service requirements for the 5G system,” 3rd Generation Partnership Project, no. TS 22.261, V.16.0.0, June 2017.

[38] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, US Government printing office, 1948, vol. 55.

[39] Y. Li et al., “Statistical analysis of MIMO beamforming with co-channel unequal-power MIMO interferers under path-loss and rayleigh fading,” *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3738–3748, 2011.

[40] D. Tse and P. Viswanath, *Fundamentals of wireless communication*, Cambridge university press, 2005.

[41] Y. Liu et al., “Analyzing grant-free access for urllc service,” *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 741–755, 2021.

[42] N. H. Mahmood and H. Alves, “Dynamic multi-connectivity activation for ultra-reliable and low-latency communication,” in *16th Int. Symp. Wireless Commun. Syst. (ISWCS)*, 2019, pp. 112–116.

[43] P. Frenger, S. Parkvall, and E. Dahlman, “Performance comparison of HARQ with chase combining and incremental redundancy for HS-PDSCH,” in *Proc. IEEE 54th Veh. Technol. Conf. (VTC Fall)*, vol. 3, 2001, pp. 1829–1833 vol.3.

[44] D. A. Wasse et al., “Radio propagation analysis of industrial scenarios within the context of ultra-reliable communication,” in *IEEE 87th Veh. Technol. Conf. (VTC Spring)*, 2018, pp. 1–6.

ONEL L. A. LÓPEZ was born in Sancti-Spíritus, Cuba, in 1989. He received the B.Sc. (1st class honors) degree in Electrical Engineering from the Central University of Las Villas (UCLV), Cuba, in 2013. From 2013-2015 he served as a specialist in telematics at the Cuban telecommunication companies (ETECSA). In 2017, he received the M.Sc. degree in Electrical Engineering from the Federal University of Paraná (UFPR), Brazil, with a grant from CAPES/CNPQ and Phd in 2020 from University of Oulu. He is currently Assistant Professor at University of Oulu. He is a co-recipient of the 2019 IEEE European Conference on Networks and Communications Best Student Paper Award and a collaborator to the 2016 Research Award given by the Cuban Academy of Sciences. His research interests are on wireless communications, specifically in ultra-reliable, low-latency communications, energy harvesting setups, and efficient access techniques for massive machine-type communications.

NURUL HUDA MAHMOOD was born in Bangladesh. Nurul is currently a Senior Research Fellow with 6G Flagship at Center for Wireless Communication, University of Oulu, where he joined in December 2018. Prior to that, he was an Assoc. Professor with Department of Electronics Systems at Aalborg University, Denmark. His current research interests include resource optimization techniques with focus on Ultra Reliable Low Latency Communication (URLLC), Industrial Internet of Things communication and modeling and performance analysis of wireless communication systems.

HIRLE ALVES received the B.Sc. and M.Sc. degrees from the Federal University of Technology-Paraná (UTFPR), Brazil, in 2010 and 2011, respectively, both in electrical engineering, and the dual D.Sc. degree from the University of Oulu and UTFPR, in 2015. In 2017, he was an Adjunct Professor in machine-type wireless communications with the Centre for Wireless Communications (CWC), University of Oulu, Oulu, Finland. In 2019, he joined CWC as an Assistant Professor and is currently the Head of the Machine-type Wireless Communications Group. He is actively working on massive connectivity and ultra-reliable low latency communications for future wireless networks, 5G and 6G, full-duplex communications, and physical-layer security. He leads the URLLC activities for the 6G Flagship Program. He is a co-recipient of the 2017 IEEE International Symposium on Wireless Communications and Systems (ISWCS) Best Student Paper Award and a co-recipient of the 2016 Research Award from the Cuban Academy of Sciences. He was the General Chair of the ISWCS’2019 and the Co-Chair of the 1st 6G Summit, Levi 2019.

MATTI LATVA-AHO received the M.Sc., Lic.Tech. and Dr. Tech (Hons.) degrees in Electrical Engineering from the University of Oulu, Finland in 1992, 1996 and 1998, respectively. From 1992 to 1993, he was a Research Engineer at Nokia Mobile Phones, Oulu, Finland after which he joined CWC. Prof. Latva-aho was Director of CWC during the years 1998-2006 and Head of Department for Communication Engineering until August 2014. Currently he serves as Academy of Finland Professor in 2017 2022 and is Director for 6Genesis-Finnish Wireless Flagship for 2018-2026. His research interests are related to mobile broadband communication systems and currently his group focuses on 5G and beyond systems research. Prof. Latva-aho has published 350+ conference or journal papers in the field of wireless communications. He received Nokia Foundation Award in 2015 for his achievements in mobile communications research.

BINOD KHAREL was born in Kapilvastu, Nepal. Binod received the M.Sc. degree from University of Oulu in 2018, and the B.Sc. degree from the Himalaya College of Engineering, Tribhuvan University, Nepal, in 2012. Currently, he is working toward the doctoral degree at the Centre for Wireless Communications (CWC) at the University of Oulu. He is a recipient of the 2019 IEEE European Conference on Networks and Communications (EuCNC) Best Student Paper Award. His research interests are on machine type communications, specifically in Ultra Reliable Low Latency Communication (URLLC).