

Fronthaul Network Planning for Hierarchical and Radio-Stripes-Enabled CF-mMIMO in O-RAN

Anas S. Mohammed, *Student Member, IEEE*, Krishnendu S. Tharakan¹, *Member, IEEE*,
 Hussein A. Ammar², *Member, IEEE*, Hesham ElSawy³, *Senior Member, IEEE*,
 and Hossam S. Hassanein⁴, *Fellow, IEEE*

Abstract—The deployment of ultra-dense networks (UDNs), particularly cell-free massive MIMO (CF-mMIMO), is mainly hindered by costly and capacity-limited fronthaul links. This work proposes a two-tiered optimization framework for cost-effective hybrid fronthaul planning, comprising a Near-Optimal Fronthaul Association and Configuration (NOFAC) algorithm in the first tier and an Integer Linear Program (ILP) in the second, integrating fiber optics, millimeter-wave (mmWave), and free-space optics (FSO) technologies. The proposed framework accommodates various functional split (FS) options (7.2x and 8), decentralized processing levels, and network configurations. We introduce the hierarchical scheme (HS) as a resilient, cost-effective fronthaul solution for CF-mMIMO and compare its performance with radio-stripes (RS)-enabled CF-mMIMO, validating both across diverse dense topologies within the open radio access network (O-RAN) architecture. Results show that the proposed framework achieves better cost-efficiency and higher capacity compared to traditional benchmark schemes such as all-fiber fronthaul network. Our key findings reveal fiber dominance in highly decentralized deployments, mmWave suitability in moderately centralized scenarios, and FSO complements both by bridging deployment gaps. Additionally, FS7.2x consistently outperforms FS8, offering greater capacity at lower cost, affirming its role as the preferred O-RAN functional split. Most importantly, our study underscores the importance of hybrid fronthaul effective planning for UDNs in minimizing infrastructural redundancy, and ensuring scalability to meet current and future traffic demands.

Index Terms—Fronthaul, planning, optimization, ultra-dense network (UDN), cell-free massive MIMO (CF-mMIMO), radio-stripes, mmWave, fiber optics, free space optics (FSO).

I. INTRODUCTION

AS MOBILE networks evolve toward beyond-5G (B5G) and 6G, demand for connectivity, high throughput, and ubiquitous coverage continues to rise [1]. This trend has driven

Received 19 October 2025; revised 26 January 2026; accepted 21 March 2026. Date of current version 7 April 2026. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) under Grant RGPIN-2023-03743 and Grant RGPIN-2025-05001. The associate editor coordinating the review of this article and approving it for publication was Z. Guan. (*Corresponding author: Hesham ElSawy.*)

Anas S. Mohammed is with the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: anas.m@queensu.ca).

Krishnendu S. Tharakan is with the School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, 114 28 Stockholm, Sweden (e-mail: tharakan@kth.se).

Hussein A. Ammar is with the Department of Electrical and Computer Engineering, Royal Military College of Canada, Kingston, ON K7K 7B4, Canada (e-mail: hussein.ammar@rmc.ca).

Hesham ElSawy and Hossam S. Hassanein are with the School of Computing, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: hesham.elsawy@queensu.ca; hossam.hassanein@queensu.ca).

Digital Object Identifier 10.1109/TWC.2026.3678907

the adoption of ultra-dense networks (UDNs), such as cell-free massive MIMO (CF-mMIMO) and small cells, which rely on dense deployment of Access Points (APs) and Central Processing Units (CPUs) to extend coverage and improve capacity [2], [3]. Unlike small cells, CF-mMIMO employs many simple cooperative APs that simultaneously serve users, removing cell boundaries and ensuring seamless service [4]. However, the *fronthaul links* connecting APs to CPUs remain a major deployment bottleneck, making the design of cost-effective and scalable fronthaul a critical priority for 6G networks [5].

The transition from Distributed RAN (D-RAN) to Cloud/Open RAN (C/O-RAN) highlights virtualization and functional split options (e.g., FS7.2x), which aim to balance fronthaul capacity and latency requirements [6], [7]. Still, UDNs, particularly CF-mMIMO, face challenges in scalable and cost-efficient fronthaul design [2], [3]. Conventional wired solutions, such as fiber, provide reliability and capacity but entail high costs and poor scalability [8]. Alternatives like Ericsson's *radio stripes (RS)* offer cost-effective wired serial connections, while wireless options such as mmWave and Free-Space Optics (FSO) promise flexibility and fast deployment but remain limited by environmental factors and capacity constraints [9], [10], [11].

In this work, we propose the hierarchical scheme (HS) as an alternative fronthaul solution for CF-mMIMO, leveraging a hierarchical topology. Similar to the RS scheme, HS significantly reduces the number of required fronthaul links. However, unlike RS, HS enhances network resilience by eliminating single points of failure inherent in the RS serial connections, while effectively addressing architectural limitations in UDNs.

Recent studies have explored optimizing several operational aspects of fronthaul networks for UDNs, specifically for RS-enabled CF-mMIMO systems. For instance, [12] proposed an optimized sequential processing algorithm for RS-enabled CF-mMIMO, enhancing signal-to-interference-plus-noise ratio (SINR) and reducing latency under limited fronthaul capacity. In [13], a geometric programming approach was proposed for the strategic placement and grouping of APs in RS deployments, emphasizing the importance of effective network planning. In contrast, [10] explored mmWave fronthaul for RS-enabled CF-mMIMO, demonstrating its potential capacity for UDN deployments. Nevertheless, these studies predominantly focus on isolated performance dimensions, neglecting fronthaul deployment cost, scalability, and infrastructural constraints, which are critical considerations towards realizing cost-effective and resilient UDNs [2], [14].

Motivated by these limitations, we aim to answer the following question: *How can we architect a scalable, cost-efficient, resilient, and high-capacity fronthaul infrastructure that supports practical CF-mMIMO deployments in future UDNs?* To address this question, we investigate the feasibility and economic viability of an optimized HS and RS-enabled hybrid fronthaul network that integrates both wired and wireless technologies, namely, fiber optics, mmWave and FSO. We show that exclusive reliance on single-technology deployment does not meet the *cost-effectiveness* and *performance* needed for UDNs.

Our primary objective is to minimize the Total Cost of Ownership (TCO) of the fronthaul network while ensuring compliance with critical metrics. Our framework supports diverse UDN scenarios, including small cells [14], and CF-mMIMO systems with RS or HS topologies, adhering to contemporary RAN architectures. It accommodates various decentralized processing levels, FS options, and AP groupings, offering actionable insights for Service Providers (SPs) on cost-efficient, high-capacity and resilient UDNs deployment. The main contributions of this paper are outlined as follows:

- We propose a two-tiered hybrid fronthaul design for UDNs through **a)** developing a Near-Optimal Fronthaul Association and Configuration (NOFAC) algorithm for the first tier, and **b)** formulating an Integer Linear Program (ILP) for the second tier. The proposed framework accommodates various fronthaul network connection schemes such as P2P small cells, along with HS and RS-enabled CF-mMIMO within the O-RAN architecture.
- We formulate and solve an ILP optimization framework that minimizes fronthaul TCO, while satisfying Quality of Service (QoS) metrics, including reliability, individual link and overall network capacities, along with fronthaul technology-specific component requirements.
- We analyze key network parameters, including varying the number of deployed CPUs, APs groups, homogeneous and non-homogeneous FS fronthaul rates for FS option 7.2x (FS7.2x) and FS option 8 (FS8), capturing practical factors affecting cost and performance. These include association distances, FS capacity thresholds, and trade-offs between group sizes and TCO.
- We evaluate the proposed framework against multiple planning benchmarks, including traditional all-fiber fronthaul. We also assess the deployment resilience of both HS and RS-enabled CF-mMIMO connection topologies.

This paper thus addresses a critical research gap by presenting a *robust fronthaul planning* framework for UDN and CF-mMIMO deployments. The remainder of the paper is organized as follows: Section II describes the system model, FS options fronthaul rates, and channel models for the candidate fronthaul technologies. Section III details the fronthaul network design to construct a NOFAC HS and RS-enabled CF-mMIMO. Section IV presents the proposed two-tier fronthaul TCO optimization. Section V discusses the numerical results, highlighting the cost and performance effectiveness of

the proposed framework. Finally, Section VI concludes the paper.

II. SYSTEM MODEL FOR HYBRID FRONTHAUL NETWORKS

This section presents the system model and key network components involved in building a hybrid fronthaul network for RS- and HS-enabled CF-mMIMO systems. We determine fronthaul capacity requirements and the channel models used to evaluate the achievable capacities of the candidate fronthaul technologies. To align with modern RAN architectures, we adopt the O-RAN and C-RAN terminologies, where the CPU is referred to as the Distributed Unit (DU) henceforth [6], [7].

A. Hybrid Fronthaul Network Architecture

We consider a hybrid fronthaul network comprising APs, DUs, and fronthaul links that leverage either fiber, mmWave or FSO. Without loss of generality, we model the deployment area as a two-dimensional square region of size $R \times R$ m², containing L APs that are randomly distributed following a uniform spatial distribution, given by the coordinates $(x_\ell, y_\ell) \sim \mathcal{U}(0, R)$ for $\ell \in \mathcal{L} = \{1, 2, \dots, L\}$. To emulate actual deployment perturbations, we consider multiple spatial realizations of APs. Initially, we employ the K-Means Clustering (KMC) algorithm to group nearby APs, constructing G preliminary groups. Grouping of APs is a critical step for both HS and RS-enabled CF-mMIMO topologies, where each group is denoted by $\mathcal{G}_i \subseteq \mathcal{L}$, for $i = 1, 2, \dots, G$, and the number of all APs in a group i is denoted by $L_{\mathcal{G}_i}$. Furthermore, each group \mathcal{G}_i is associated with one of W distributed DUs based on proximity. These DUs, indexed by $w \in \mathcal{W} = \{1, 2, \dots, W\}$, are responsible for serving several APs groups.

In RS, and as described in the early RS antenna arrangement patent [9], APs within a group are connected serially via a shared wired fronthaul infrastructure. Specifically in this work, all APs within the same group \mathcal{G}_i utilize a common intra-group fiber link to receive data streams from their serving DU, rather than each AP requiring a dedicated P2P link. This interpretation is strictly at the group level and does not imply that all APs in the network are connected through a single global fiber link. To reduce signaling overhead, a single AP is designated as the **leading AP** in each group. This AP, selected as one of the terminal points of the stripe, establishes a direct communication link with its serving DU via either a wired or wireless fronthaul connection [10], [15], as illustrated in Figure 1b. The leading AP processes and forwards fronthaul signals to other APs in the stripe, referred to as **non-leading APs**, through the shared fronthaul infrastructure [10], [12], [15]. This setup diverges from conventional small-cell architectures, which rely on dedicated P2P fronthaul links as depicted in Figure 1a. We extend the concept of RS to the HS configuration shown in Figure 1c, where the serial connection is replaced by a hierarchical topology. In this configuration, the leading AP is the one having the highest number of connection degrees, i.e., it has the largest number of dependent non-leading APs.

Efficient deployment of fronthaul technologies requires careful consideration of their unique components. For

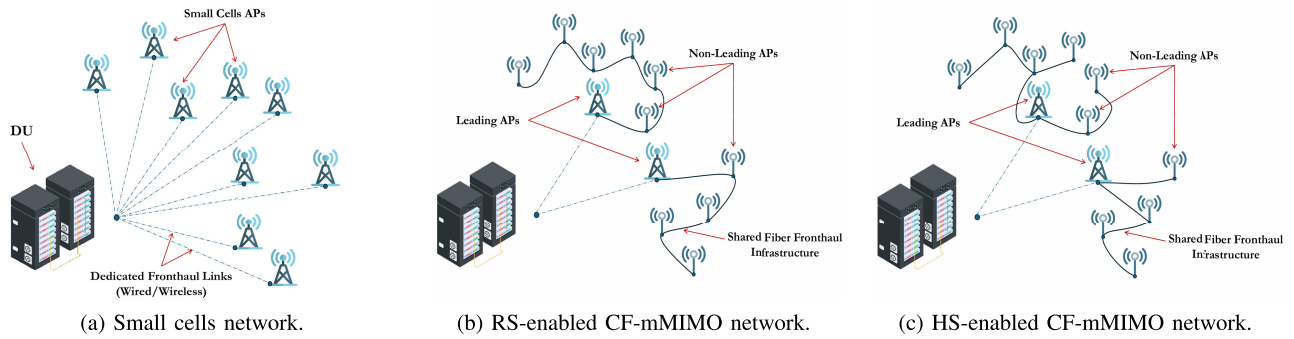


Fig. 1. Illustration of the different fronthaul topologies employed for UDNs schemes; (a) Conventional small-cell architecture with dedicated P2P fronthaul links, (b) RS topology and (c) HS topology.

fiber-based fronthaul, the hardware needed for a typical Wavelength-Division Multiplexing Passive Optical Network (WDM-PON) includes fiber cables, Optical Add-Drop Multiplexers (OADMs) and Optical Network Units (ONUs) integrated with each AP ℓ , along with the Optical Transport Network (OTN) collocated at each DU w . OTNs manage optical signals from multiple ONUs and employ components such as splitters, multiplexers (MUXs), and Optical Line Terminals (OLTs) to aggregate signals [16]. On the other hand, the antenna configuration in mmWave-based fronthaul influences the network model. For simplicity, we assume that all APs utilizing mmWave fronthaul have a single antenna (i.e., $N_{AP} = 1$), while DUs are equipped with N_{DU} antennas. In contrast, FSO-based fronthaul uses P2P links, with each AP having a dedicated transceiver paired with its associated DU, ensuring single transmission and reception points.

We also assume that all APs within a group cooperate to provide spatial diversity for jointly-served users. That is for every group of APs \mathcal{G}_i , all APs need to receive a copy of the same message from their serving DU, hence, eliminating the need for transmitting user-specific data to each AP individually. Instead, the same message received by leading APs is shared among all APs in the group, thereby simplifying fronthaul processing. Given the ultra-dense nature of the network components deployment, line-of-sight (LoS) connectivity between APs and DUs in wireless fronthauling is assumed. Consequently, wireless relays and repeaters are excluded from consideration, as unobstructed connections are deemed achievable.

We assume uncompressed fronthaul to isolate cost-performance tradeoffs between transmission technologies and avoid introducing codec-specific variability, FS options tradeoffs, technology-specific or transport medium limitations, as these could be additional metrics for comparison affecting the fronthaul rates (e.g., μ -law and Block Floating Point (BFP) compression techniques, free space and fiber transport mediums, etc). Fronthaul compression is typically considered in specific scenarios rather than general infrastructure-level planning, and our analysis ensures a conservative estimate of fronthaul demands and preserves generality without biasing toward a specific compression method.

B. FS-Options Capacity Requirements

FS7.2x and FS8 are regarded as key candidates in B5G networks, due to their alignment with O-RAN architecture,

TABLE I
OFDM-BASED STANDARD PARAMETERS VALUES FOR 5G NR NUMEROLOGY 0, AND THE REQUIRED FRONTHAUL CAPACITY FOR FS8 AND FS7.2X

Configuration Parameters for 5G NR Numerology 0	
Bandwidth (B)	20 MHz
Sampling Frequency (f_s)	30.72 MHz
Subcarrier Spacing (Δf)	15 kHz
Symbol Duration (T_{symbol})	66.67 μ s
Total Number of Subcarriers (N_{DFT})	2048
Effective/Used Number of Subcarriers (N_{used})	1200
Quantization Bits (N_{bits})	12
Number of Access Antennas per AP ($N_{\text{AP}}^{\text{ac}}$)	4
FS8 Required Capacity (ψ^{FS8})	2.95 Gbps
FS7.2x Required Capacity ($\psi^{\text{FS7.2x}}$)	1.73 Gbps

[7], [17]. We introduce a fronthaul data rate threshold, denoted as ψ , to specify the minimum fronthaul capacity required for each AP, which will guide the optimization process and selection of appropriate fronthaul technologies. TABLE I lists the capacity requirements for FS8 and FS7.2x under standard 5G NR numerology 0 system configuration. This configuration is one of several possible numerologies defined by 5G NR, where numerology 0 and 1 correspond to a subcarrier spacing of 15 kHz and 30 kHz, respectively, with both commonly implemented in 5G systems. While these values may resemble LTE parameters, they are fully compliant with 5G NR as per 3GPP TS 38.211 (Release 18) [18]. We emphasize that our choice of numerology does not affect the solution design for fronthaul deployment, as subcarrier spacing Δf only changes the frame structure and symbol timing, but not the overall data rate or required fronthaul capacity, which remains dependent on bandwidth and antenna configuration. Specifically, the system employs Orthogonal Frequency-Division Multiplexing (OFDM), a widely used modulation scheme in LTE and 5G systems, expected to be present in B5G networks [17]. Key parameters include OFDM symbol duration T_{symbol} , sampling frequency f_s , and the quantization bit-width N_{bits} , which denotes the number of bits per in-phase (I) and quadrature (Q) components. We assume that the access channel between users and APs is modeled as a block-fading channel, with each coherence block comprising τ_c time-frequency OFDM samples [17]. The available bandwidth B is divided into N_{DFT}

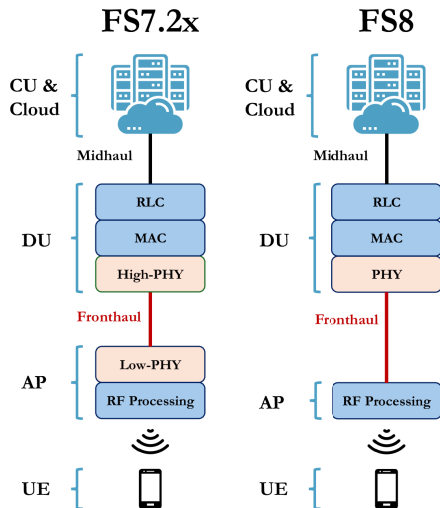


Fig. 2. Distribution of processing tasks in FS7.2x and FS8.

subcarriers using the Discrete Fourier Transform (DFT), with N_{used} effective subcarriers for data transmission and N_{null} reserved for guard bands. The number of AP antennas on access channel, $N_{\text{AP}}^{\text{ac}}$, is distinct from those used for mmWave fronthaul, N_{AP} .

1) *Functional Split 8 (FS8)*: The lowest layer split is defined as FS8, which is the PHY-RF split option aimed at fully benefiting from the efficient processing capabilities at the DUs, Centralized Units (CUs) and Cloud, while reducing APs complexity. As shown in Figure 2, APs in FS8 only perform RF processing and receive raw, time-domain, quantized baseband signals from DUs through the fronthaul, leading to high fronthaul capacity requirements. For FS8, we assign each AP ℓ with the minimum required fronthaul capacity as follows [17]:

$$\psi^{\text{FS8}} = 2 \times N_{\text{bits}} f_s N_{\text{AP}}^{\text{ac}}. \quad (1)$$

It is important to note that N_{bits} in equation (1) refers to the bit-width per component (I or Q), and the multiplication factor of 2 accounts for both. Thus, the total bits per complex I/Q sample in our formulation equals $2 \times N_{\text{bits}}$.

2) *Functional Split 7.2x (FS7.2x)*: Similarly, FS7.2x assigns part of the intra-PHY functions to be processed at the APs, while the remaining high-PHY functions are shifted to the DUs, as shown in Figure 2. FS7.2x strikes a balance between AP complexity and fronthaul bandwidth, making it ideal for scalable O-RAN deployments in different UDNs schemes. Specifically, APs only receive the effective subcarriers (N_{used}) from DUs, leading to APs performing additional low-PHY functions, hence, lowering the fronthaul capacity requirements compared to FS8. The required fronthaul capacity in FS7.2x is [17]:

$$\psi^{\text{FS7.2x}} = \frac{2 \times N_{\text{bits}} N_{\text{used}} N_{\text{AP}}^{\text{ac}}}{T_{\text{symbol}}}. \quad (2)$$

Note that (1) and (2) provide a baseline estimation of fronthaul data rate requirements under FS8 and FS7.2x options, following the widely used formulations in [17]. These expressions primarily capture the user-plane data transfer. In practical systems, additional overhead from control signaling, CSI exchange, synchronization, and protocol encapsulation may

further increase the fronthaul load, which we will account for next. The values listed in Table I assume each AP operates at full load, serving the maximum expected user throughput. This worst-case planning approach, widely adopted in network design, ensures robustness under peak demand conditions. Consequently, our fronthaul deployment decisions are shaped by user traffic indirectly, through strict capacity constraints applied in the optimization framework. This approach is consistent with UDN/CF-mMIMO literature, where dense and uniform user demand is a standard modeling assumption [2], [17].

3) *Non-Homogeneous Traffic and the Control Plan*: Equations (1) and (2) define a minimum required fronthaul capacity for any AP assuming a homogeneous spatial traffic scenario, in which mobile operators assign this value for every AP in the network. In contrast, mobile operators can use a non-homogeneous spatial traffic scenario, where the expected traffic is determined based on a data traffic survey. In this scenario, each leading-AP $\ell \in \mathcal{M}_w$ has a different minimum required fronthaul capacity that is defined as [14]:

$$\psi_{\ell}^{\text{FS}[8,7.2x]} = f_{\text{traf.}}(x_{\ell}, y_{\ell}), \quad (3)$$

where $f_{\text{traf.}}(\cdot)$ is the traffic-aware fronthaul capacity needed in the area centered on (x_{ℓ}, y_{ℓ}) which is the location of the leading-AP $\ell \in \mathcal{M}_w$.

Adding control plane traffic for the equations in (1) and (2) can be a very involved task. Interestingly, we can add an overhead term that accounts for the control plane. According to ORAN technical report [19], there is a defined typical control plane that is sent on the fronthaul which includes PRACH channel I-Q data, scheduling and beamforming commands, configuration parameters and request, ACK/NACK message, and other signals. We account for these messages that represent the control plane as a factor proportional to the data plane by defining the following:

$$\psi^{\text{FS8, CP}} = (1 + \alpha^{\text{FS8, ovh}}) \psi^{\text{FS8}}, \quad (4)$$

$$\psi^{\text{FS7.2x, CP}} = (1 + \alpha^{\text{FS7.2x, ovh}}) \psi^{\text{FS7.2x}}, \quad (5)$$

where $\alpha^{\text{FS8, ovh}}, \alpha^{\text{FS7.2x, ovh}} \in [0, 1]$, and represent the amount of overhead introduced by the control plane. In our results, we will account for this control plane traffic using a safety margin that is calculated through surplus. We will illustrate this point in the results section.

C. Fiber-Based Fronthaul Link Capacity

Due to the short distances between network elements in CF-mMIMO and the high efficiency of fiber optics, we assume lossless fronthaul links with a constant capacity denoted by $R_{w\ell}^{\text{Fiber}}$. We consider employing a 10 Gbps-capable symmetrical (XGS) WDM-PON, providing equal UL and DL data rates while accommodating evolving capacity demands. WDM-PON supports multiple wavelengths and is expected to remain a key solution for fronthaul/backhaul networks [16]. Consequently, each AP with a fiber-based fronthaul link has a constant capacity $R_{w\ell}^{\text{Fiber}}$ of 10 Gbps.

D. mmWave-Based Fronthaul Link Capacity

For mmWave, we adopt 3GPP 38.901 Urban Microcell street canyon (UMi-SC) model to characterize both LoS and Non-Line-of-Sight (NLoS) propagation, modeling the down-link (DL) performance from DUs to APs [20]. SPs typically optimize terrestrial network fronthaul by leveraging the static positioning of APs and DUs to maintain LoS conditions during wireless fronthaul planning. Thus, LoS parameters are distance-based, while NLoS parameters, including the number of paths $P \sim \mathcal{U}[1, 6]$ and angle-of-departure $\theta_p \sim \mathcal{U}[-\frac{\pi}{2}, \frac{\pi}{2}]$, are randomly selected based on discrete uniform distribution to account for possible reflections [10]. Accordingly, the path loss for both LoS and NLoS scenarios is expressed respectively as follows:

$$PL_{w\ell, \text{dB}}^{\text{LoS}} = 32.4 + 21 \log_{10}(d_{w\ell}) + 20 \log_{10}(f_c) + \mathcal{S}_{\text{LoS}}, \quad (6)$$

$$PL_{w\ell, \text{dB}}^{\text{NLoS}} = 32.4 + 31.9 \log_{10}(d_{w\ell}) + 20 \log_{10}(f_c) + \mathcal{S}_{\text{NLoS}}, \quad (7)$$

where f_c is the carrier frequency in GHz and $d_{w\ell}$ denotes the distance between AP ℓ and its serving DU w in meters. While $\mathcal{S}_{\text{LoS}} \in \mathcal{N}(0, \sigma_{\text{LoS}}^2)$ and $\mathcal{S}_{\text{NLoS}} \in \mathcal{N}(0, \sigma_{\text{NLoS}}^2)$ are the shadowing terms modeled as Gaussian random variables with zero mean and standard deviation of 4 and 8.2 dB, respectively [20]. Moreover, let $\mathbf{h}_{w\ell} \in \mathbb{C}^{N_{\text{DU}}}$ be the frequency-domain channel between the ℓ -th AP and its serving w -th DU, and is expressed as the sum of LoS and NLoS components:

$$\mathbf{h}_{w\ell} = \mathbf{h}_{w\ell}^{\text{LoS}} + \mathbf{h}_{w\ell}^{\text{NLoS}}. \quad (8)$$

Before transmission to APs, DUs adopt analog beamforming with a network of quantized phase shifters to compute the beamforming vectors, approximating ZF beamforming to mitigate interference and focus the transmitted signal toward the intended APs [10]. The assumption of quantized phase shifters is for satisfying the practical constraints of mmWave, which means that the beamforming vector $\mathbf{f}_{w\ell}$ could only be selected from a certain set of vectors \mathcal{F} , thus, $\mathbf{f}_{w\ell} \in \mathcal{F}$. That is if every phase shifter has q quantization bits, then we will have 2^q phase shift values defined by the discrete set $\mathcal{Q} = \left\{0, \frac{\pi}{2^q}, \dots, \frac{(2^q-1)\pi}{2^q}\right\}$. Hence, the set of all possible beamforming vectors for any DU is expressed as [10]:

$$\mathcal{F} = \left\{ \frac{[e^{j\phi_1} \dots e^{j\phi_{N_{\text{DU}}}}]^T}{N_{\text{DU}}}; \phi_i \in \mathcal{Q}, \forall i \in \{1, \dots, N_{\text{DU}}\} \right\}, \quad (9)$$

and the fronthaul signal received at AP ℓ is given as:

$$\begin{aligned} \mathbf{r}_{\ell}^{\text{mmW}} &= \sqrt{p_t^{\text{mmW}}} \mathbf{h}_{w\ell}^T \mathbf{f}_{w\ell} \varsigma_{w\ell} + \underbrace{\sum_{\substack{w'=1, i=1 \\ w' \neq w, i \neq \ell}}^W \sum_{\substack{L}}^L \sqrt{p_t^{\text{mmW}}} \mathbf{h}_{w'i}^T \mathbf{f}_{w'i} \varsigma_{w'i}}_{\text{interference}} + \mathbf{n}_{\ell}, \end{aligned} \quad (10)$$

where p_t^{mmW} is the normalized fronthaul transmission power, $\mathbf{n}_{\ell} \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2)$ is the receiver (RX) noise at the ℓ -th AP, and $\varsigma_{w\ell} \in \mathbb{C}$ is the signal transmitted from the w -th DU to AP ℓ with $\mathbb{E}\{|\varsigma_{w\ell}|^2\} = 1$. Given the relatively large distances

between APs and non-serving DUs, combined with our use of beamforming, inter-DU interference is assumed negligible. Consequently, we focus on the Signal-to-Noise Ratio (SNR) to determine the actual rates of the mmWave link. Therefore, the SNR at AP ℓ and the corresponding mmWave fronthaul link capacity can be expressed as follows:

$$\text{SNR}_{\ell} = \frac{p_t^{\text{mmW}} |\mathbf{h}_{w\ell}^T \mathbf{f}_{w\ell}|^2}{\sigma^2}, \quad (11)$$

$$R_{w\ell}^{\text{mmW}} = \text{BW}^{\text{mmW}} \log_2(1 + \text{SNR}_{\ell}). \quad (12)$$

Remark 1: Although analog beamforming with quantized phase shifters may not exactly replicate ZF performance leading to an imperfection in canceling interference, it remains effective in focusing energy toward serving APs and attenuating unintended directions, particularly under LoS or dominant-path channel conditions, as assumed in our fronthaul model. In practice, DUs and APs are statically positioned as part of a planned deployment, enabling precomputed beam directions and stable LoS-dominant channels. Additionally, the usage of mmWave frequencies limits the range of communication and hence interference. These factors make the assumption of negligible inter-DU interference stronger. Motivated by this, we use SNR for our performance evaluation for the mathematical tractability advantage. In an actual deployment with static fronthaul scenario, mobile operators can relax our assumption by setting a target network surplus capacity (Figure 8).

E. FSO-Based Fronthaul Link Capacity

FSO has been proposed as a viable communication solution for fronthaul and backhaul, demonstrating sufficient availability and data rates over long distances [11], [21], [22]. Nevertheless, FSO technology remains severely affected by environmental conditions, where losses such as scattering, turbulence, scintillation, geometrical spreading, and optical inefficiencies can jeopardize its reliability. Scattering occurs when the FSO beam interacts with atmospheric particles, and expressed as [21]:

$$L_{\text{sca}} = 4.34 \left(\frac{3.91}{V} \left(\frac{\lambda}{550} \right)^{-\delta} \right) \frac{d_{w\ell}}{1000}, \quad (13)$$

where L_{sca} is the scattering loss in dB, V is the visibility range in km, λ is the signal wavelength in nm, and δ is a visibility-dependent constant given as $\delta = 0.585 V^{1/3}$, for $V < 6$ km [21]. Additionally, atmospheric turbulence losses are caused by variations in the refractive index structure parameter $C_n^2(h)$ of the atmosphere at altitude h , measured in $\text{m}^{-2/3}$, and transiently affected by wind speeds [22]. During the planning stage, an average value for the refractive index structure parameter for moderate turbulence is $[10^{-15} \text{m}^{-2/3}]$, as considered in [22]. Scintillation loss, denoted as L_{sci} , is caused by the rapid fluctuations in light intensity when it propagates through a turbulent atmosphere, and the resultant attenuation in dB is given by:

$$L_{\text{sci}} = 2 \sqrt{23.17 \left(\frac{2\pi}{\lambda} 10^9 \right)^{\frac{7}{6}} C_n^2(h) d_{w\ell}^{\frac{11}{6}}}. \quad (14)$$

Finally, geometrical losses L_{geo} result from light spreading over a larger area. Assuming circular mirrors at both TX and RX, the geometrical loss in linear scale is [21]:

$$L_{\text{geo}} \approx \left(\frac{r_r}{\left(\frac{\theta_t d_{w\ell}}{2}\right)} \right)^2, \quad (15)$$

where r_r is the RX aperture radii in meters, and θ_t is the divergence angle of the transmitter (TX) in radians. Several assumptions are made before deriving the achievable data rate for FSO links ($R_{w\ell}^{\text{FSO}}$). Due to the short APs-DUs distances, an average visibility range of $V = 0.4$ km is assumed. A study in Berlin, Germany [23] shows that visibility drops below 0.4 km only 0.25% of the year, resulting in an average FSO link availability of 99.75% ($\zeta^{\text{FSO}} = 0.9975$). Consistent with the assumptions made in [21], we consider LoS FSO transmission. Under moderate fog conditions, fog loss is calculated as $L_{\text{fog}} = 20.99 d_{w\ell}$ dB, where $d_{w\ell}$ is the distance in km between the AP and its associated DU [21]. Lastly, a 10 dB loss (L_{rain}) is included to account for rain and other worst-case scenarios. Therefore, total atmospheric losses are expressed as:

$$L_{\text{dB}}^{\text{atm}} = (L_{\text{sca}} + L_{\text{rain}} + L_{\text{fog}} + L_{\text{sci}})_{\text{dB}}. \quad (16)$$

The rate for the FSO link between DU w and AP ℓ is [21]:

$$R_{w\ell}^{\text{FSO}} = \frac{p_t^{\text{FSO}} \eta_t \eta_r r_r^2}{L_{\text{atm}} E_p N_b \left(\frac{\theta_t d_{w\ell}}{2}\right)^2}, \quad (17)$$

where p_t^{FSO} is the transmit power, L_{atm} is the atmospheric losses in linear scale, η_t and η_r are the optical losses caused by imperfections in optical TX and RX, respectively. E_p is the photon energy expressed as $E_p = \frac{h_p c}{\lambda}$, where $h_p = 6.625 \times 10^{-34}$ J-s is Max Planck's constant, and N_b is the RX sensitivity in photons/bit.

III. HIERARCHICAL AND RADIO-STRIPES-ENABLED CF-mMIMO NETWORK CONFIGURATION

To simplify the problem formulation for selecting optimal fronthaul technologies, and without any loss of practicality, a *brownfield deployment scenario* is assumed. In this scenario, the locations of APs are assumed to be predetermined by SPs according to the specific coverage and economic requirements. In this section, we introduce NOFAC, an iterative algorithm that minimizes inter-APs distances, balances cluster formation, efficiently deploys DUs, and achieves a **Near-Optimal Fronthaul Association and Configuration (NOFAC)**.

A. Preliminary Network Construction Methodology

The spatial density and geographic positioning of APs influence the grouping process, ensuring that APs in close proximity are grouped together to minimize inter-AP distances, fronthaul costs, and latency [8], [10]. To mitigate limitations of basic KMC, which generates G preliminary AP groups of irregular sizes, initial clusters are refined by applying Split-Merge Rules (SMR) based on predefined system parameters, namely: split (g_s) and merge (g_m) thresholds, defining the maximum and minimum permissible number of APs per group, respectively.

Following the initial clustering, groups exceeding the maximum permissible size ($L_{G_i} > g_s$) are split into smaller clusters. Conversely, groups smaller than the minimum allowable size ($L_{G_i} < g_m$) are merged with adjacent groups, provided the resulting group size does not exceed g_s (i.e., $g_s < L_{G_i} < g_m$). These refinement steps ensure that all groups maintain a balanced number of APs, balanced groups sizes, eliminating oversized and undersized clusters of both HS and RS-enabled CF-mMIMO networks construction. The sets of all APs and groups associated with DU w are represented by $\mathbf{L}_w \subseteq \mathcal{L}$, and $\mathbf{G}_w \subseteq \mathbf{L}_w$, respectively.

B. Constructing Radio-Stripes-Based CF-mMIMO Network

After refining the AP groups, APs within each group \mathcal{G}_i for in RS are connected in a serial topology to minimize the total inter-AP distances. This involves identifying a sequence of APs that forms the lowest total traversal path, which is analogous to the Traveling Salesmen Problem (TSP). For smaller groups with $L_{G_i} \leq L^{\text{max}}$, where L^{max} is a system parameter ensuring that the factorial complexity ($L_{G_i}!$) remains computationally tractable, the TSP is solved optimally using brute-force search across all APs permutations. Nevertheless, when $L_{G_i} > L^{\text{max}}$, we use the alternative Nearest Neighbor (NN) algorithm as a computationally efficient approximation. Thus, a practical approach to minimize the overall serial connection distance between the same-group APs is done through the mix use of TSP and NN algorithms. The process for groups with $L_{G_i} \leq L^{\text{max}}$ begins by computing the pairwise distances between all APs in a group \mathcal{G}_i . The path with the minimal total distance is then selected as follows

$$d_{G_i}^{\text{RS}} = \arg \min_{\pi \in \mathcal{P}} L_{\pi}, \quad \forall i = 1, 2, \dots, G, \quad (18)$$

where L_{π} represents the total path length of a permutation π , and \mathcal{P} is the set of all possible paths. If the number of APs in a cluster exceeds L^{max} , NN algorithm is applied with multiple starting points to improve the chances of approximating the optimal solution. Following the construction of RS, we deploy W DUs with their placement initially optimized using KMC, ensuring proper association planning between DUs $w = 1, 2, \dots, W$ and APs groups $\mathcal{G}_i = 1, 2, \dots, G$ based on proximity to groups centroids. This approach ensures that each group is fully assigned to a single DU, as shown in Figure 3a. From each group \mathcal{G}_i , a leading AP ℓ_i is selected from the stripe endpoints, prioritizing the AP closest to the DU to minimize the total association distance. Since DU positions and leading AP selections are interdependent, Algorithm 1 implements NOFAC, which attempts to minimize the total association distance while maintaining effective AP-DU associations through updating DU positions, re-assigning leading APs, and re-iterating over the previous steps. Convergence is achieved when the maximum movement of any DU w between iterations falls below a predefined threshold ϵ :

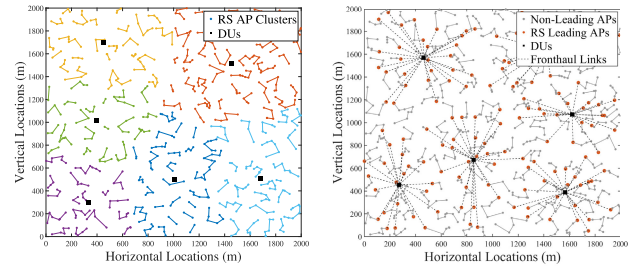
$$\max_w \|\mu_w^{\text{new}} - \mu_w^{\text{old}}\| < \epsilon, \quad \forall w \in \mathcal{W}, \quad (19)$$

where μ_w^{new} and μ_w^{old} represent the updated and previous positions of DU w , and ϵ is a very small positive number. If the convergence condition is not met, the process restarts. The final positions of the DUs and the selection of leading APs

Algorithm 1 NOFAC: For a Near-Optimal HS and RS-Based CF-mMIMO Fronthaul Network Association and Configuration

- 1: **Input:** W , $-\text{DEL } L$, $-\text{DEL } G$, $-\text{DEL } \epsilon$, $-\text{DEL } L^{\max}$, $-\text{DEL } g_s$ and g_m .
- 2: **Initialization:**
- 3: Uniformly distribute L APs.
- 4: Construct G initial AP groups ($\mathcal{G}_{\text{initial}} = \{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_G\}$) using KMC.
- 5: Deploy W DUs using KMC (μ^{initial}).
- 6: $\mu_w^{\text{old}} \leftarrow \mu_w^{\text{initial}} \quad \forall w \in \mathcal{W}$;
- 7: **while** true **do**
- 8: **for** $\forall w \in \mathcal{W}$ **do**
- 9: **for** $\forall i \in \mathcal{G}_w$ **do**
- 10: **if** $L_{\mathcal{G}_i} > g_s$ **then**
- 11: Split group \mathcal{G}_i .
- 12: **else if** $L_{\mathcal{G}_i} < g_m$ **then**
- 13: Merge group \mathcal{G}_i with the nearest group.
- 14: **else**
- 15: Skip \mathcal{G}_i .
- 16: **end if**
- 17: **if** RS-enabled CF-mMIMO network, **then**
- 18: **if** $L_{\mathcal{G}_i} \leq L^{\max}$ **then**
- 19: Apply TSP.
- 20: **else**
- 21: Apply NN.
- 22: **end if**
- 23: Compute $d_{\mathcal{G}_i}^{\text{RS}} = \arg \min_{\pi \in \mathcal{P}} L_{\pi}$.
- 24: **else** \triangleright HS-enabled CF-mMIMO network
- 25: Construct MST for \mathcal{G}_i .
- 26: Compute $d_{\mathcal{G}_i}^{\text{MST}} = \min_{\mathcal{T}_i \subseteq \mathcal{G}_i} \sum_{e_{jk} \in \mathcal{T}_i} w_{jk}$.
- 27: **end if**
- 28: **end for**
- 29: **end for**
- 30: Assign Leading APs (\mathcal{M}_w).
- 31: Refine DUs positions using KMC (μ_w^{new}).
- 32: **if** $\max_w \|\mu_w^{\text{new}} - \mu_w^{\text{old}}\| < \epsilon$, $\forall w \in \mathcal{W}$ **then**
- 33: **Break**
- 34: **else**
- 35: $\mu_w^{\text{old}} \leftarrow \mu_w^{\text{new}} \quad \forall w \in \mathcal{W}$;
- 36: **end if**
- 37: **end while**
- 38: **Output:** μ_w^{new} , \mathbf{L}_w , \mathbf{G}_w , \mathcal{M}_w , $\mathcal{A}_w \quad \forall w \in \mathcal{W}$, and \mathcal{G}_i , $L_{\mathcal{G}_i}$, $-\text{DEL } d_{\mathcal{G}_i}^{\text{RS}}$ or $d_{\mathcal{G}_i}^{\text{MST}} \quad \forall i = 1, 2, \dots, G$.

forming the final RS-based CF-mMIMO network are depicted in Figure 3b. After convergence, each DU w is assigned a set of leading APs \mathcal{M}_w and non-leading APs \mathcal{A}_w , with corresponding sizes M_w and A_w , such that $\mathcal{M}_w, \mathcal{A}_w \subseteq \mathbf{L}_w$. Although Algorithm 1 is a heuristic procedure, it includes a well-defined convergence condition, i.e. (19) based on the stabilization of DU placements. Given the finite number of APs and the deterministic refinement rules applied to the groupings, the algorithm consistently converges within a small number of iterations.



(a) Initial RS network deployment. (b) Final optimized associations.

Fig. 3. Sample realization of the optimized RS network.

C. Constructing Hierarchical-Based CF-mMIMO Network

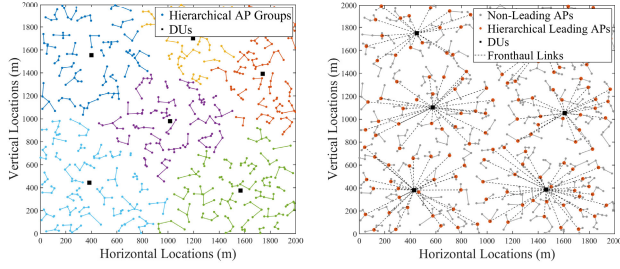
The HS-enabled CF-mMIMO network construction builds on the principles employed in the RS network described in Section III-B. However, instead of using serial connections, the APs within each group \mathcal{G}_i are connected in a hierarchical topology to minimize inter-AP distances using Minimum Spanning Tree (MST) algorithm. This graph-theoretic approach ensures cost-effective connectivity between APs [24]. Each group \mathcal{G}_i is represented as a complete weighted graph $\mathcal{G}_i = (\mathcal{V}_i, \mathcal{E}_i)$, where \mathcal{V}_i is the set of APs (vertices) in group \mathcal{G}_i , and \mathcal{E}_i represents the set of links (edges) connecting every pair of APs in a group. Thus, the weight of each link $e_{jk} \in \mathcal{E}_i$ is calculated as the Euclidean distance between APs j and k . The MST for each group \mathcal{G}_i is constructed by minimizing the total edge weight of the tree as follows:

$$d_{\mathcal{G}_i}^{\text{MST}} = \min_{\mathcal{T}_i \subseteq \mathcal{G}_i} \sum_{e_{jk} \in \mathcal{T}_i} w_{jk}, \quad \forall i = 1, 2, \dots, G, \quad (20)$$

where $d_{\mathcal{G}_i}^{\text{MST}}$ is the total weight (i.e., total distance) of the MST for group \mathcal{G}_i , and \mathcal{T}_i is the MST of group \mathcal{G}_i , which is a subgraph of \mathcal{G}_i that connects all APs with the minimum possible total link distances. The MST problem in (20) can be optimally computed using Prim's or Kruskal's algorithm [24], with the resulting HS illustrated in Figure 4a. Then, a similar process to the RS construction steps in Section III-B is followed to deploy W DUs. The leading AP from each group \mathcal{G}_i is chosen as the node with the highest number of edges connected to it, which reflects the number of neighboring APs in the tree. In cases of tie, the AP nearest to the DU is selected to minimize association distance, with remaining APs classified as non-leading. The positions of DUs and selection of leading APs are iteratively optimized following the NOFAC approach in Algorithm 1, and the process repeats until convergence is achieved as per equation (19).

IV. FRONTHAUL TCO OPTIMIZATION FORMULATION

This section presents the planning and cost optimization of the fronthaul network for HS and RS-based CF-mMIMO schemes. The objective is to minimize the TCO while ensuring optimal fronthaul technology selection. Our formulation starts by incorporating critical information, including the locations of APs and DUs, APs-DUs clusters (\mathbf{L}_w) and distances ($d_{w\ell}$), the effective capacities of leading APs for all fronthaul technologies ($R_{w\ell}^{\text{Fiber}}$, $R_{w\ell}^{\text{mmW}}$, and $R_{w\ell}^{\text{FSO}}$), alongside the outputs of NOFAC in Algorithm 1. We consider LTE-based functional



(a) Initial HS network deployment. (b) Final optimized associations.

Fig. 4. Sample realization of the optimized HS network.

split options 8 and 7.2x for capacity requirements as defined in Section II-B, collectively represented as ψ^{FSX} , where X denotes either 7.2x or 8. This optimization model employs a two-tiered approach to minimize TCO, encompassing both operational (OPEX) and capital expenditures (CAPEX), with OPEX averaged over a fixed deployment period N_{period} .

A. Tier 1 Optimization

The first tier of optimization focuses on minimizing the deployment cost the shared fiber fronthaul infrastructure that consists of ONUs/OADM and fiber cables per meter for the set of non-leading APs (\mathcal{A}_w), while masking leading APs (\mathcal{M}_w) in every group. This has been addressed by NOFAC in Algorithm 1, by optimizing the grouping of APs and minimizing the pairwise connection distances $d_{\mathcal{G}_i}^{\text{RS}}$ and $d_{\mathcal{G}_i}^{\text{MST}}$ between all APs in a group \mathcal{G}_i . Following the traditional construction of RS [9], [10], we assume that all APs in a group are connected via fiber cables, with each AP equipped with an ONU and an OADM to facilitate signal transmission and conversion from electrical to optical over a stripe. At each AP in a group, one signal is dropped using OADMs and inserted to a receiving ONU. For all the set of non-leading APs (\mathcal{A}_w) associated with DU w , the incurred costs comprise the ONU cost (C^{ONU}), which incorporates the OADM cost, installation cost, O&M costs over N_{period} years, collectively represented as C_{ℓ}^{Fiber} . Considering that both the HS and RS systems were optimized in Sections III-A, III-B and III-C, the TCO for Tier 1 remains constant at this stage, where X refers to either RS or MST, and it can be expressed as:

$$C_{T_1}^A = \sum_{w=1}^W A_w C^{\text{ONU}} + \sum_{i=1}^G \eta^{\text{Fiber}} d_{\mathcal{G}_i}^X. \quad (21)$$

B. Tier 2 Optimization

The second tier of optimization focuses exclusively on optimizing the fronthaul links between leading APs (\mathcal{M}_w) and their associated DUs. We formulate a unique objective function for each candidate fronthaul technology to minimize the TCO of the fronthaul network by accounting for both OPEX and CAPEX of each technology, where CAPEX is further divided into deployment costs per AP (C_{ℓ}) and DU (C_w).

1) *Fiber-Based Fronthaul*: For every leading AP ℓ utilizing fiber, the incurred costs comprise the ONU cost (C^{ONU}), which incorporates the OADM cost, installation cost, O&M costs over N_{period} years, collectively represented as C_{ℓ}^{Fiber} . Additionally, the cost of trenching and burial of fiber cables

per meter is given by η^{Fiber} . At the DU side, the OTN cost associated with DU w serving fiber-connected leading APs is denoted by C_w^{Fiber} . This includes the necessary collocated infrastructure at the DU to facilitate fiber-based fronthaul, such as MUXs, OLTs, etc. Therefore, the TCO for fiber-based fronthaul is expressed as:

$$C^{\text{Fiber}} = \sum_{w=1}^W \sum_{\ell=1}^{M_w} \left(\underbrace{C^{\text{ONU}} + N_{\text{period}} C_{\text{O\&M}}^{\text{Fiber}}}_{C_{\ell}^{\text{Fiber}}} + \eta^{\text{Fiber}} d_{w\ell} \right) + \sum_{w=1}^W \left(\underbrace{C^{\text{OLT}} + C^{\text{OTN}} + C^{\text{Other}}}_{C_w^{\text{Fiber}}} \right). \quad (22)$$

2) *mmWave-Based Fronthaul*: For mmWave-based fronthaul, the TCO per leading AP connection, including annual power consumption, installation, and O&M costs, is denoted by C_{ℓ}^{mmW} . On the other hand, each DU that serves mmWave-based leading APs will have a massive MIMO antenna device with N_{DU} elements, with its cost denoted by C_w^{mmW} . Hence, the TCO for the mmWave-based fronthaul is given by:

$$C^{\text{mmW}} = \sum_{w=1}^W \sum_{\ell=1}^{M_w} \left(\underbrace{C_{\ell}^{\text{mmW-RX}} + N_{\text{period}} C_{\text{O\&M}}^{\text{mmW}}}_{C_{\ell}^{\text{mmW}}} \right) + \sum_{w=1}^W C_w^{\text{mmW}}.$$

Remark 2: The DU cost $C_{\text{DU}}^{\text{pool}}$ used in this study reflects a pooled compute infrastructure model rather than standalone edge units. This cost is fixed at 91,035 based on FCC estimates [25], and includes shared resources such as cooling, power, and rack space. While our model does not explicitly scale cost with the number of connected APs or implement load-aware pooling, it captures the trade-off between centralized and distributed computing by simulating scenarios with different numbers of DU pools (ranging from 2 to 12). This setup enables analysis of centralization impact on cost and fronthaul deployment. Future work may incorporate dynamic, load-dependent pooling models to more precisely capture cloudification efficiency.

3) *FSO-based Fronthaul Network TCO*: In FSO-based fronthaul, we focus on P2P FSO transmission with dedicated TXs and RXs. Simply, for every leading AP that uses FSO, we have a reserved exclusive link and equipment, and we could aggregate these costs into a single term C^{FSO} , representing the average cost of deployment for every group \mathcal{G}_i with its leading AP utilizing FSO for fronthauling. Hence, the TCO for FSO-based fronthaul links can be expressed as follows:

$$C^{\text{FSO}} = \sum_{w=1}^W \sum_{\ell=1}^{M_w} \left(\underbrace{C_{\ell}^{\text{FSO}} + C_{\text{install}}^{\text{FSO}} + N_{\text{period}} C_{\text{O\&M}}^{\text{FSO}} + C_w^{\text{FSO}}}_{C_{\ell}^{\text{FSO}}} \right). \quad (23)$$

4) *Final Objective Function*: By combining the TCO of all fronthaul technologies (Fiber, mmWave, and FSO), the final joint cost objective function can be expressed as follows:

$$g_0(\mathbf{x}, \mathbf{z}, \mathbf{u}, \mathbf{v}, \boldsymbol{\kappa}) = \sum_{w=1}^W \sum_{\ell=1}^{M_w} x_{w\ell} (C_{\ell}^{\text{Fiber}} + \eta^{\text{Fiber}} d_{w\ell})$$

$$\begin{aligned}
 & + \sum_{w=1}^W \sum_{\ell=1}^{M_w} [z_{w\ell} C_{\ell}^{\text{mmW}} + u_{w\ell} C^{\text{FSO}}] \\
 & + \sum_{w=1}^W [v_w C_w^{\text{mmW}} + \kappa_w C_w^{\text{Fiber}}], \quad (24)
 \end{aligned}$$

where $x_{w\ell}, z_{w\ell}, u_{w\ell} \in \{0, 1\}$ are the binary controlling variables for leading APs selecting fiber, FSO or mmWave, respectively. These vectors will have their ℓ -th entry being 1 to indicate if the ℓ -th AP is using the corresponding technology for fronthauling, and 0 otherwise. The integer $\kappa_w \in \mathbb{Z}$ is a variable indicating the number of fiber OTNs required at the DU side, based on the number of leading APs using fiber associated with the w -th DU, and the capacity of OTNs (Θ) in terms of how many fiber-based leading APs they can support. Lastly, $v_w \in \{0, 1\}$ is a binary variable indicating if a DU w has any connected mmWave leading APs if 1, and 0 otherwise.

Effective fronthaul planning and joint optimization of multiple technologies necessitate the incorporation of practical and realistic constraints guiding the selection process of the most cost-effective technologies. These constraints can be divided into three main categories, namely: General architectural constraints, technology-specific constraints, QoS metrics constraints, and are as follows:

5) *General Architectural Constraints:* This category ensures robust formulation and network components association, and the constraints in this category include the following:

$$\sum_{w=1}^W (x_{w\ell} + z_{w\ell} + u_{w\ell}) = 1, \quad \forall \ell \in \mathcal{M}_w, \quad \forall w \in \mathcal{W}, \quad (25)$$

$$x_{w\ell}, z_{w\ell}, u_{w\ell} \in \{0, 1\}, \quad \forall \ell \in \mathcal{M}_w, \quad (26)$$

$$v_w \in \{0, 1\}, \quad \forall w \in \mathcal{W}, \quad (27)$$

$$\kappa_w \in \mathbb{Z}, \quad \forall w \in \mathcal{W}, \quad (28)$$

where (25) ensures that each group \mathcal{G}_i is only associated with a single DU w and selecting one fronthaul technology for every leading AP $\ell \in \mathcal{M}_w$, while both (26) and (27) define the binary controlling variables, and (28) defines the fiber-associated controlling variable κ_w as an integer.

6) *QoS Metrics Constraints:* This category ensures that individual leading APs and the entire network meet performance standards, and it includes:

$$\begin{aligned}
 & x_{w\ell} R_{w\ell}^{\text{Fiber}} + z_{w\ell} R_{w\ell}^{\text{mmW}} + u_{w\ell} R_{w\ell}^{\text{FSO}} \\
 & \geq \psi^{\text{FSX}}, \quad \forall \ell \in \mathcal{M}_w, \quad \forall w \in \mathcal{W}, \quad (29)
 \end{aligned}$$

$$\begin{aligned}
 & \sum_{\ell=1}^{M_w} (x_{w\ell} \zeta^{\text{Fiber}} + z_{w\ell} \zeta^{\text{mmW}} + u_{w\ell} \zeta^{\text{FSO}}) \\
 & \geq M_w \zeta^{\text{SLA}}, \quad \forall w \in \mathcal{W}, \quad (30)
 \end{aligned}$$

where constraint (29) guarantees that the selected fronthaul technology minimizing the TCO for each leading AP $\ell \in \mathcal{M}_w$ is also meeting a fixed fronthaul capacity threshold imposed by the FS option requirements ψ^{FSX} , where FSX refers to options FS7.2x or FS8. While the availability constraint (30) ensures that for the entire N_{period} , the Service Level Agreement (SLA) of SPs is not breached, by capturing the average uptime for a network. Availability means that all leading APs in the network are ready for immediate use, by ensuring that the number of

fully operational leading APs associated with each DU w are always larger than or equal to $M_w \zeta^{\text{SLA}}$.

7) *Technology-Specific Constraints:* This category ensures that the unique characteristics of each technology are accurately represented, and it includes:

Fiber Constraint: In fiber-based fronthaul, the maximum number of groups with their leading APs choosing fiber for fronthauling connected to a single DU depends on the capacity (Θ) of the OTN deployed at the DU side. For simplicity, we are taking the ratio option of splitters to determine the bottleneck capacity of OTNs. Assuming a single splitter supports Θ number of links (i.e., 1: Θ PON), then if more than Θ APs associated with w -th DU are using fiber, and additional OTN must be deployed, effectively doubling the cost of C_w^{Fiber} .

$$\sum_{\ell=1}^{M_w} \frac{x_{w\ell}}{\Theta} \leq \kappa_w \leq \sum_{\ell=1}^{M_w} \frac{x_{w\ell}}{\Theta} + 1 - \epsilon, \quad \forall w \in \mathcal{W}. \quad (31)$$

mmWave Constraint: Similar to fiber, this constraint guarantees that the w -th DU will not be equipped with a mmWave antenna device unless there is at least one leading AP employs mmWave technology for fronthauling. As a result, since v_w is a binary variable, v_w will be equal to 1 if and only if $\sum_{\ell=1}^{M_w} z_{w\ell} \neq 0$.

$$\sum_{\ell=1}^{M_w} \frac{z_{w\ell}}{M_w} \leq v_w \leq \sum_{\ell=1}^{M_w} z_{w\ell}, \quad \forall w \in \mathcal{W}. \quad (32)$$

Constraints (31) and (32), combined with the definition of their decision variables in constraints (27) and (28) resemble a linearized form of a ceiling function applied to the decision variables κ_w and v_w .

C. Final Formulation and Proposed Algorithm

By combining the TCO of both tiers, the final optimization problem for both the studied HS and RS-enabled CF-mMIMO networks that aims to minimize the fronthaul network TCO and ensures effective performance through the selection of fronthaul technologies is presented as follows:

$$\min_{\substack{x_{w\ell}, z_{w\ell}, u_{w\ell} \\ v_w, \kappa_w}} g_0(\mathbf{x}, \mathbf{z}, \mathbf{u}, \mathbf{v}, \boldsymbol{\kappa}) + C_{T_1}^A, \quad (33a)$$

$$\text{subject to (25) - (32)}. \quad (33b)$$

The above formulation is classified as an Integer Linear Program (ILP), which is a combinatorial optimization problem characterized by the combination of binary ($x_{w\ell}, z_{w\ell}, u_{w\ell}, v_w$) and integer (κ_w) variables, alongside the linear nature of the objective function and constraints. To achieve the optimal solution of equation (33), specifically for the second tier, we employ the branch-and-bound method [26], and the steps are outlined in Algorithm 2. The algorithm returns the fronthaul technologies selection ($\mathbf{x}, \mathbf{z}, \mathbf{u}, \mathbf{v}$, and $\boldsymbol{\kappa}$), cost values of each technology (C^{Fiber} , C^{mmW} , and C^{FSO}), in addition to the TCO of both tiers of the optimized network. Furthermore, while Algorithm 1 provides the topology-level input, the core optimization problem presented in Section IV is an ILP, which is solved using a branch-and-bound method. This guarantees convergence to a globally optimal solution due to the discrete and bounded nature of the solution space. Since this

TABLE II
INPUT SYSTEM PARAMETERS VALUES USED IN SIMULATION FOR THE HS AND RS-BASED CF-mMIMO SYSTEM

Fiber		mmWave		FSO		General	
η^{Fiber}	\$26	C_w^{mmW}	\$34,500	C^{FSO}	\$15,000 [21]	Coverage area (A)	2 km \times 2 km
C^{ONU}	\$6,502	$C_\ell^{\text{mmW-RX}}$	\$6,000	$C_{\text{O\&M}}^{\text{FSO}}$	\$13,000 [21]	C_{DU}^{pool}	\$91,035
C^{OTN}	\$61,727	$C_{\text{O\&M}}^{\text{mmW}}$	\$13,000	r_r, V	0.05 m, 400 m	$\psi^{\text{FS7.2x}}$	1.73 Gbps
C^{OLT}	\$20,100	$N_{\text{AP}}, N_{\text{DU}}$	1, 256	η_t, η_r	0.5, 0.5	ψ^{FS8}	2.95 Gbps
C_w^{Fiber}	\$81,827	p_t^{mmW}	120 W	$p_t^{\text{FSO}}, \theta_t$	0.5 W, 10 mrad	L	1000
$C_{\text{O\&M}}^{\text{Fiber}}$	\$2,285	q	6	$C_n^2(h)$	$10^{-15} \text{m}^{-2/3}$ [22]	L^{max}	9
R^{Fiber}	10 Gbps	BW^{mmW}	2.5 GHz	E_p	1.2823×10^{-19}	N_{period}	1 year
ζ^{Fiber}	1.00	ζ^{mmW}	0.99999	ζ^{FSO}	0.9975	ζ^{SLA}	0.9999
Θ	16	f_c	80 GHz	λ, N_b	1550 nm, 100	$\mathbf{g}_s, \mathbf{g}_m$	15, 3
Variables:		No. of groups (G)		No. of DUs (W)		No. of APs per group (L_{g_i})	

Algorithm 2 Proposed Algorithm for Solving the Hierarchical and Radio-Stripes-Based CF-mMIMO Fronthaul Planning and Cost Minimization ILP in Eq. (33)

```

1: Input: Set up the given technology-specific and general
   parameters values from TABLE II as input to the system.
2: for realization  $r = 0$  do
3:    $r \leftarrow r + 1$ ;
4:   Initialization:
5:   Obtain NOFAC Algorithm 1 outputs as input.
6:   for  $\forall w \in \mathcal{W}$  do
7:     Initialize  $\mathbf{v}^{(r)}$  and  $\kappa^{(r)} \leftarrow 0$ ;
8:     for  $\forall \ell \in M_w$  do
9:       Compute  $d_{w\ell}$ ,  $R_{w\ell}^{\text{Fiber}}$ ,  $R_{w\ell}^{\text{mmW}}$ , and  $R_{w\ell}^{\text{FSO}}$ .
10:      Initialize  $\mathbf{x}^{(r)}$ ,  $\mathbf{z}^{(r)}$  and  $\mathbf{u}^{(r)} \leftarrow 0$ ;
11:     end for
12:   end for
13:   Compute  $C_{T_1}^A$  from eq. (21).
14:   Solve eq. (33) using branch and bound to obtain:
15:    $\mathbf{x}_*^{(r)}$ ,  $\mathbf{z}_*^{(r)}$ ,  $\mathbf{u}_*^{(r)}$ ,  $\mathbf{v}_*^{(r)}$ , and  $\kappa_*^{(r)}$ .
16: end for
17: Output:  $\mathbf{x}, \mathbf{z}, \mathbf{u}, \mathbf{v}, \kappa, C^{\text{Fiber}}, C^{\text{mmW}}, C^{\text{FSO}}, C_{T_1}^A$ , and
    $g_o(\mathbf{x}, \mathbf{z}, \mathbf{u}, \mathbf{v}, \kappa)$ .

```

framework is intended for offline planning rather than real-time deployment, convergence speed is not a limiting factor. Nonetheless, we observe stable and efficient convergence behavior in all tested configurations.

V. NUMERICAL RESULTS

We evaluate the effectiveness of the proposed framework to understand the critical role of fronthaul network planning in CF-mMIMO and, more broadly, in UDNs. In particular, we focus on the use of either HS or RS connection schemes within the O-RAN paradigm. The selection of fronthaul technologies, optimization of TCO and network capacity are examined with respect to several key factors, mainly: (1) Varying the number of DUs (W), impacting distances between DUs and APs. (2) Varying the number of AP groups (G) for the same total number of deployed APs L , which affects the number of leading APs and non-leading APs, eventually impacting the fronthaul infrastructural units needed. (3) Fronthaul capacity

thresholds, namely $\psi^{\text{FS7.2x}}$ and ψ^{FS8} , where ψ^{FS8} makes constraint (29) more stringent. Unless otherwise specified, the cost estimates for all equipment in TABLE II are primarily derived from average values reported by the US FCC [25].

It is important to note that the simulation results presented in this section are averaged over hundreds of spatially randomized network realizations, reflecting diverse AP distributions, group topologies, and DU placements. This extensive sampling implicitly captures the performance variability induced by different deployment topologies, including effects stemming from radio-stripe interconnections. Although our framework does not explicitly model physical layer processing strategies such as sequential decoding or cooperative fusion, the averaged performance over varied topologies ensures robustness and generalizability of the results. Moreover, the proposed planning strategy is modular and can be adapted to operator-specified topology constraints or processing architectures as needed.

The performance of the optimization framework is evaluated by comparing it against three benchmarks, each focusing on different fronthaul technology strategies for leading APs in the second tier of the network, and these are: (a) **Standard all-Fiber fronthaul network**, serving as a benchmark for the most reliable performance, satisfying all constraints (25) - (32). (b) **Suboptimal all-mmWave fronthaul network**, representing an economically appealing benchmark, but falls short of meeting the capacity ψ^{FSX} constraint (29) for all leading APs. Lastly, (c) **Heuristic method for hybrid deployment**, balancing cost and capacity without necessarily achieving optimal TCO. In this method, mmWave is initially assigned to all APs, and if the capacity threshold (ψ^{FSX}) for a leading AP ℓ exceeds its mmWave capacity $R_{w\ell}^{\text{mmW}}$, it is switched to fiber, while ensuring compliance with all the remaining constraints.

In our solution, we do not focus on enhancing the bandwidth capabilities of fiber optics or any other fronthaul technology. Instead, we adopt a planning perspective that works within the constraints of existing technologies. While our study considers standard fronthaul capacity requirements based on uncompressed data for FS7.2x and FS8, it is agnostic to specific fronthaul enhancements (e.g., compression, protocol efficiency improvements). Our optimization framework remains applicable and flexible to adapt to future technology

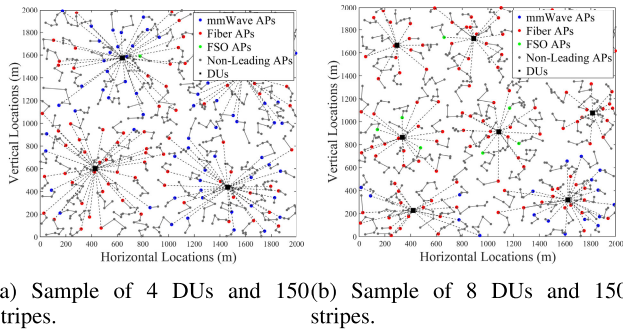


Fig. 5. Optimized fronthaul mixed-technology selection in RS with FS7.2x under different levels of decentralized processing.

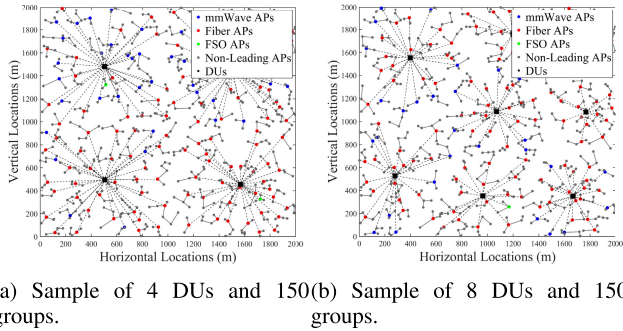


Fig. 6. Optimized fronthaul mixed-technology selection in HS with FS7.2x under different levels of decentralized processing.

upgrades, as it fundamentally operates on cost, capacity, and reliability parameters, regardless of underlying physical layer innovations.

A. Fronthaul Technologies Selection

Figures 5 and 6 illustrate fronthaul technology choices for RS and HS topologies under FS7.2x, considering different numbers of deployed DUs (W). Although the results reflect specific realizations, the insights extend to other configurations and traffic scenarios [14]. For both HS- and RS-enabled CF-mMIMO, APs near DUs typically use fiber due to lower cost, while distant APs rely on mmWave when capacity and reliability allow. FSO adoption remains limited, given its higher cost and lower reliability. Interestingly, and contrary to widely held assumptions, as processing becomes more decentralized and AP-DU distances shrink, fiber deployment increases. Additionally, the optimization framework prioritizes efficient utilization of DU-associated infrastructure, by maximizing connections using the same technology, hence minimizing infrastructural redundancy across technologies.

B. Network Optimized TCO and Number of Groups

Figure 7 shows the average TCO per AP for different numbers of DUs (W) and groups (G). Although an all-mmWave deployment often appears to have the lowest cost, it consistently violates the capacity constraint (29) and is therefore infeasible. Our optimized design achieves the most cost-efficient deployment across both connection schemes and FS options. Increasing group sizes (L_{G_i}) lowers fronthaul costs but adds processing burden on each DU, while deploying more DUs distributes the load more evenly and reduces overhead.

Moreover, at lower decentralization levels (e.g., Figures 7a and 7b), fiber-only deployments incur the highest TCO due to the long distances between APs and DUs, whereas the suboptimal all-mmWave case typically shows the lowest cost despite constraint violations. On the other hand, at higher decentralization levels (e.g., Figure 7c), the cost-effectiveness of the suboptimal all-mmWave scheme diminishes. Unlike common belief, this shift occurs due to the reduced AP-DU distances, which sometimes render fiber to be a more economically appealing option. The average TCO per AP is nearly identical for RS and HS-enabled CF-mMIMO under the same FS option. However, FS8 incurs slightly higher costs than FS7.2x due to stricter fronthaul capacity requirements. This cost difference becomes negligible at DU densities (e.g., $W = 8$), where fiber dominates for both FS options. Nevertheless, the selection of fronthaul technologies is contingent on the network configuration, FS option, connection scheme employed, and the number of deployed groups.

To generalize these findings, Figure 9 highlights the average network TCO, summing both tiers, in Millions of US dollars (MM), along with a breakdown of Tier 2 cost contributions of each technology and their respective percentages for RS-enabled CF-mMIMO. The results are averaged over numerous Monte Carlo simulations with group counts G ranging from 100 to 200, for both FS7.2x and FS8 options. Figure 9 results are also applicable to the HS topology, providing valuable insights into the economic considerations associated with different deployment strategies and O-RAN FS options. With sparse DU deployment ($W = 2$), long AP-DU distances result in higher attenuation for wireless links, making fiber the preferred choice despite its high TCO. As W increases (e.g., $W = 4-8$), AP-DU distances become more manageable, and wireless fronthaul technologies become more viable options. At higher levels of decentralized processing ($W > 8$), fiber deployment costs drop significantly, further solidifying its role as an effective fronthaul solution in these scenarios. Finally, Figure 9b reveals that the stringent fronthaul capacity requirement of FS8 (ψ^{FS8}) leads to a higher percentage of fiber deployment compared to FS7.2x (Figure 9a), across all DU densities and connection schemes.

C. Network Surplus Capacity and Number of Groups

The efficiency of deployment strategies can be assessed through fronthaul surplus capacity, which measures the difference between the total capacity provided by the deployed fronthaul network and the actual demand imposed by various FS options. A positive surplus indicates that the network has excess capacity beyond the current demand, offering room for future traffic growth or accommodating unexpected surges. Conversely, a deficiency signifies that the network is operating below the traffic demand, compromising its ability to maintain QoS. Figure 8 shows that while the all-fiber benchmark offers the highest capacity, our hybrid optimized schemes consistently achieve substantial surplus at comparatively lower cost, demonstrating the effectiveness of the proposed framework. Additionally, the suboptimal all-mmWave frequently fails to meet the minimum required rates (ψ^{FSX}), resulting in the lowest surplus capacity, or even capacity deficiencies. While

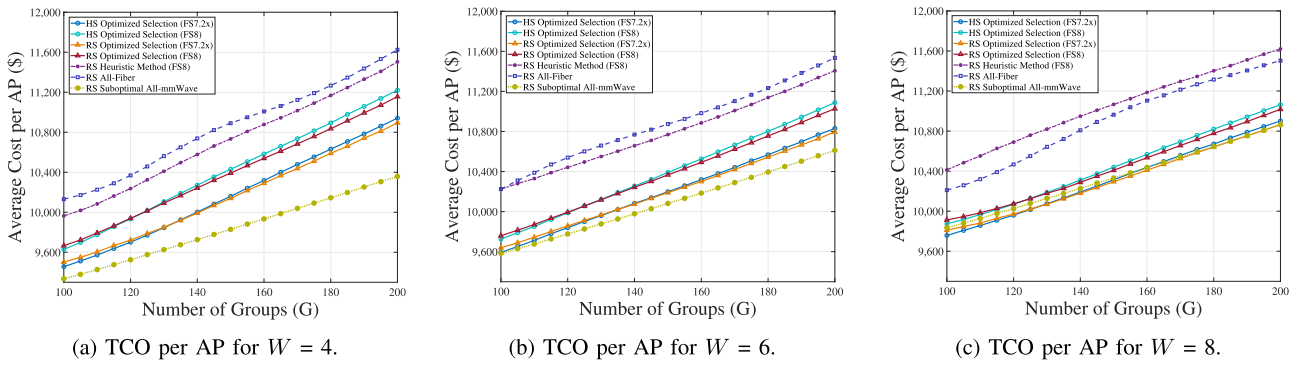


Fig. 7. Average TCO per AP vs. number of groups (G) across different levels of decentralized processing and benchmarks.

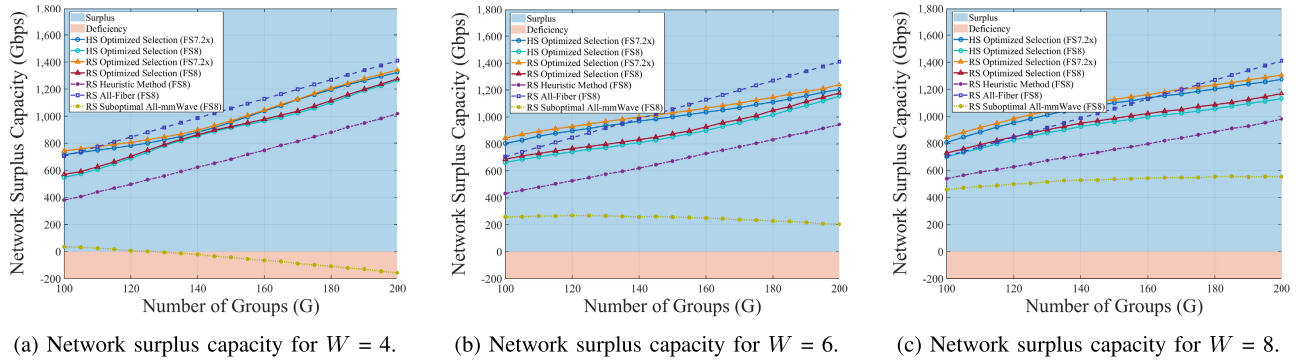
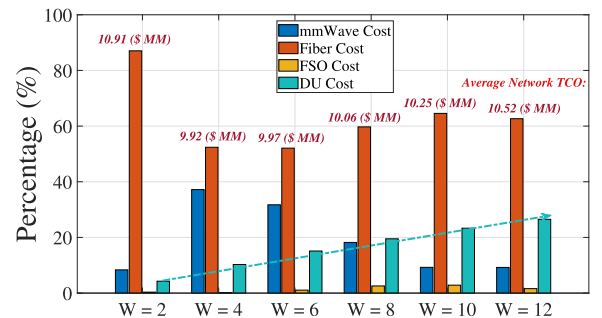


Fig. 8. Surplus capacity vs. number of groups (G) across different levels of decentralized processing and benchmarks.

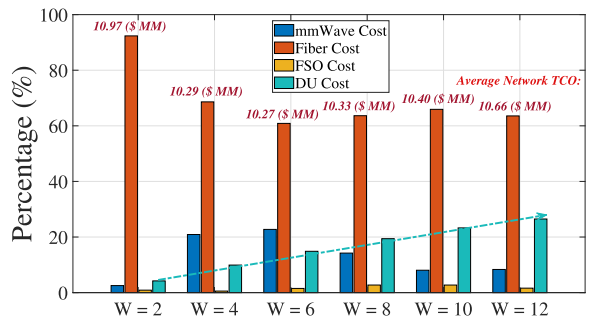
the current optimization does not enforce any surplus constraint, network operators may extend the model by imposing minimum surplus thresholds to ensure future-proofing in static deployments.

These results emphasize that low-cost strategies must not compromise QoS requirements to ensure reliable network performance. Moreover, the findings presented in Figure 8 consistently demonstrate the superiority of FS7.2x, offering greater surplus capacity compared to FS8 in both schemes. When combined with the cost-effectiveness insights from Figure 7, this reinforces FS7.2x as the preferred O-RAN functional split for future mobile networks [7], [17]. Furthermore, FS7.2x aligns well with UDNs requirements, where the O-RAN architecture plays a role in complementing the vision of future UDNs deployment. Across all presented results, the heuristic method consistently yields lower surplus capacity and higher TCO compared to the optimized network. Therefore, we conclude that the best strategy in UDNs involves a well-planned and diversified mix of fronthaul technologies, as exemplified by the superior performance of our optimized network, in both FS options. The results highlights that cost considerations should not overshadow QoS requirements to reliably meet performance targets for SPs.

Remark 3: Surplus fronthaul capacity is not used as a measure of algorithm performance but rather as an indicator of network resilience and headroom. It quantifies how much additional traffic the fronthaul infrastructure could support beyond the minimum required by the selected functional split. Most importantly, the surplus plots can be used to account for the additional control plane traffic as discussed in (4) and (5). This metric helps assess the network's future readiness



(a) FS7.2x average network TCO and Tier 2 cost breakdown.



(b) FS8 average network TCO and Tier 2 cost breakdown.

Fig. 9. Average network TCO and technology cost distribution for different DU counts in RS-based CF-mMIMO.

under rising throughput demands, sudden traffic surges, and evolving service requirements. It is further important to note that any observed surplus fronthaul capacity is not the result of

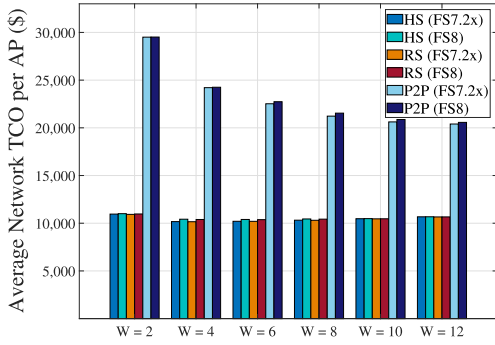


Fig. 10. Comparison of average TCO per AP among small cells (P2P fronthaul), RS, and HS topologies under varying FS options and different levels of decentralized processing.

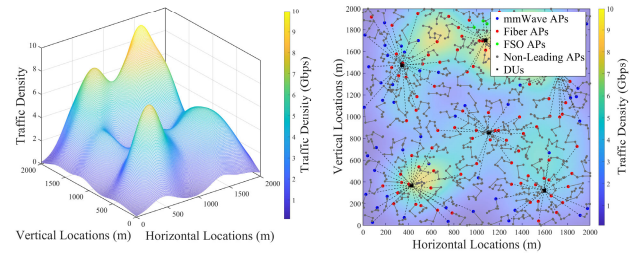
intentional over-dimensioning, but rather an artefact of discrete fronthaul technology selection under strict cost minimization and feasibility constraints, and is therefore interpreted solely as a resilience and headroom indicator.

D. Radio-Stripes, Hierarchical and Small Cells TCO

To underscore the cost-effectiveness of the optimized fronthaul connection schemes in supporting ultra-dense CF-mMIMO deployments, Figure 10 compares the average TCO per AP in RS, HS, and the traditional P2P fronthaul connection scheme for small cells networks [14]. To ensure comparison fairness, each of the aforementioned systems deploys an identical configuration of L APs within the same coverage area and satisfies the same capacity requirements (ψ^{FSX}) in constraint (29), for both FS7.2x and FS8. Moreover, both RS and HS setups are configured with a fixed number of groups ($G = 200$). Figure 10 reveals that both RS and HS setups significantly reduce the average TCO per AP by sharing fronthaul resources among AP within the same group, particularly pronounced at lower levels of decentralized processing (e.g., $W < 4$). These schemes provide a practical and efficient alternative to conventional P2P connections for UDNs. When combined with the proposed optimization framework, RS and HS topologies further enhance the economic viability and scalability of UDN architectures for next-generation mobile networks.

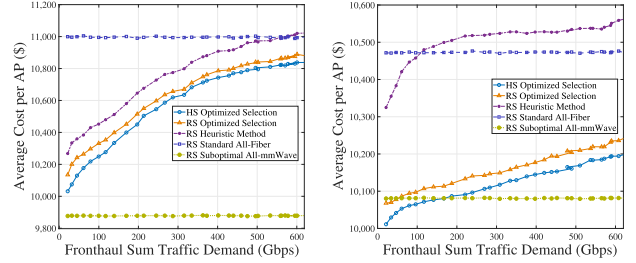
E. Non-Homogeneous Traffic Analysis

To further assess the robustness and scalability of the proposed framework beyond standardized functional split operating conditions, we conduct a systematic stress test by increasing the minimum fronthaul bandwidth requirement using a traffic-aware methodology, similar to what is described in Section II-B3. Importantly, this analysis does not involve redesigning fronthaul technologies or modifying their underlying models. Motivated by our previous work [14], we replace the fixed functional split requirement ψ^{FSX} in (29) with a spatially varying threshold ψ_ℓ^{het} assigned to each leading AP $\ell \in \mathcal{M}_w$. This modification preserves the original optimization structure and cost formulation, replacing only the static traffic requirement with a traffic-aware threshold that captures operator-style spatial demand field. Specifically, the minimum



(a) Traffic distribution mesh plot. (b) Optimized fronthaul selection.

Fig. 11. Sample realization of combined RS-enabled CF-mMIMO network and traffic distribution with optimized fronthaul technology selection for $G = 150$ and $W = 6$.



(a) TCO per AP for $W = 2$. (b) TCO per AP for $W = 6$.

Fig. 12. Average TCO per AP vs. fronthaul sum traffic thresholds for different DU counts and $G = 150$.

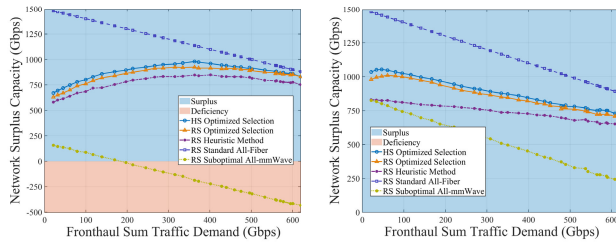
fronthaul requirement for each leading AP is defined as a function of its spatial coordinates:

$$\psi_\ell^{\text{het}} = f_{\text{traffic}}(x_\ell, y_\ell), \quad \forall \ell \in \mathcal{M}_w. \quad (34)$$

The spatial traffic field is generated using a hotspot-based Gaussian mixture model that emulates operator-style demand distributions [14]. A total of N_s traffic hotspots are randomly placed over the coverage area, as shown in Figure 11. Moreover, the peak demand is limited by \mathbf{X}^{max} , which has to be less than the highest physical capacity of the used technologies (Fiber, 10 Gbps in this study). Accordingly, the fronthaul capacity constraint in (29) becomes as follows:

$$x_{w\ell} R_{w\ell}^{\text{Fiber}} + z_{w\ell} R_{w\ell}^{\text{mmW}} + u_{w\ell} R_{w\ell}^{\text{FSO}} \geq \psi_\ell^{\text{het}}, \quad \forall \ell \in \mathcal{M}_w, \forall w \in \mathcal{W}. \quad (35)$$

Finally, each leading AP $\ell \in \mathcal{M}_w$ is assigned a minimum capacity threshold according to its location through (34), which directly governs the feasibility of (35) and, consequently, the fronthaul technology selection. As ψ_ℓ^{het} increases, the optimization progressively prioritizes higher-capacity and more reliable fronthaul technologies for these APs. This trend clearly leads to an increase in the optimized TCO and network capacity, along with a gradual convergence toward more fiber-dominated deployments, especially under highly centralized deployments (i.e., $W = 2$), as backed by previous results in [14], and demonstrated in Figures 12 and 13. This behavior reflects that by increasing fronthaul traffic demand, under which low- and medium-capacity links become infeasible, high-capacity fronthaul solutions dominate the feasible design space. The resulting technology selection patterns shown in Figure 11 is consistent with the core objective of the framework. It trades cost for capacity and reliability only when needed, favoring fiber in high-traffic regions while preserving

(a) Surplus capacity for $W = 2$. (b) Surplus capacity for $W = 6$.Fig. 13. Network surplus capacity vs. fronthaul sum traffic thresholds for different DU counts and $G = 150$.

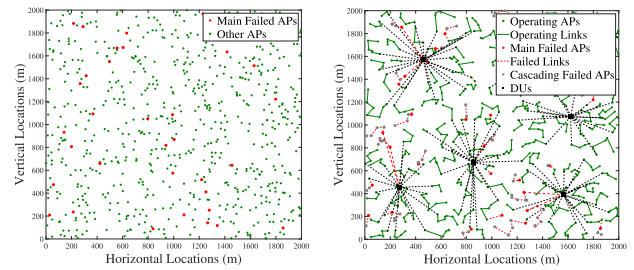
feasibility within the physical capabilities of the available fronthaul technologies.

Remark 4: Infeasible links under extreme demand is a direct consequence of physical fronthaul capacity limits rather than a limitation of the proposed optimization framework. If ψ_ℓ^{het} at a given leading AP exceeds the maximum achievable rate supported by all available technologies (R^{Fiber} , R^{mmW} , or R^{FSO}), no feasible fronthaul assignment exists by construction. In such cases, the framework correctly identifies these locations instead of forcing invalid assignments. From a planning perspective, this outcome signals that the deployment has reached a technology-imposed boundary and motivates actionable remedies such as network densification, topology reconfiguration, or upgrading the fronthaul technology set. Therefore, this analysis serves as a diagnostic scalability tool that clearly distinguishes algorithmic behavior from fundamental physical capacity ceilings.

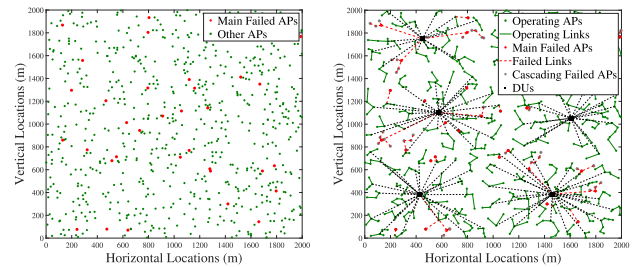
F. Network Resilience Analysis

Despite the economic advantages of RS, the serial architecture makes its resilience against link failures questionable. Specifically, the sequential connectivity in RS implies that a single fronthaul link failure could trigger cascading failures affecting the rest of subsequent APs. Conversely, HS uses a branched topology that reduces interdependence among APs within the same group and mitigates the cascading failure effects. Let p represent the fraction of APs that experience an outage due to a failure in the fronthaul link of any deployed AP $\ell \in \mathcal{L}$, forming the main failed APs set $\mathcal{K}_m \subseteq \mathcal{L}$. This failure applies equally to both RS and HS-enabled CF-mMIMO systems. In RS, a single AP failure cascades to subsequent APs in the stripe, owing to the serial fronthaul connection. On the other hand, HS reduces this cascading effect through its branched connections. In both schemes, we denote APs failing due to cascading effects as the set $\mathcal{K}_c \subseteq \mathcal{L}$. The worst-case scenario occurs when the fronthaul link of the leading AP (Tier 2) fails, resulting in the complete loss of all dependent APs within the group, as depicted in Figures 14 and 15. Hence, the set of all failed APs, \mathcal{K} , is the total sum of main and cascaded failures, i.e., $\mathcal{K} = \mathcal{K}_m \cup \mathcal{K}_c \subseteq \mathcal{L}$.

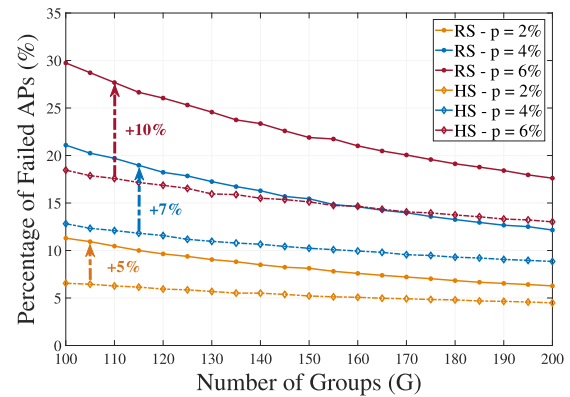
The severity of cascading failures depends on the fraction of main AP failures p , the number of AP groups G , the group size L_{G_i} , and the specific locations of main failing APs. To provide a broader perspective, Figure 16 shows the average percentage of total AP failures for various G and p values. We observe that the RS scheme exhibit high vulnerability to



(a) Sample of main failures in RS. (b) Sample of cascaded failures in RS.

Fig. 14. RS-enabled CF-mMIMO network resilience sample realization for $G = 150$ and $p = 5\%$.

(a) Sample of main failures in HS. (b) Sample of cascaded failures in HS.

Fig. 15. HS-enabled CF-mMIMO network resilience sample realization for $G = 150$ and $p = 0.05$.Fig. 16. Comparison of average total AP failures for varying numbers of groups G and main failure fractions p .

cascading failures, with $p = 6\%$ leading to about 30% total AP outages for $G = 100$. The HS topology, however, shows improved resilience, reducing total outages to around 19% under similar conditions. Given its comparable cost-efficiency and superior resilience, HS emerges as a compelling fronthaul connectivity solution for future UDNs and CF-mMIMO deployment. Additionally, Figure 16 reveals that the severity of cascading failures inversely correlates with the number of deployed groups G . Increasing the number of groups G , while keeping L fixed, naturally decreases the total number of APs per group L_{G_i} , thereby lowering interdependence and limiting the impact of cascading failures. This trend approaches the behavior of P2P small cells, which isolate failures through dedicated links and offer higher overall robustness.

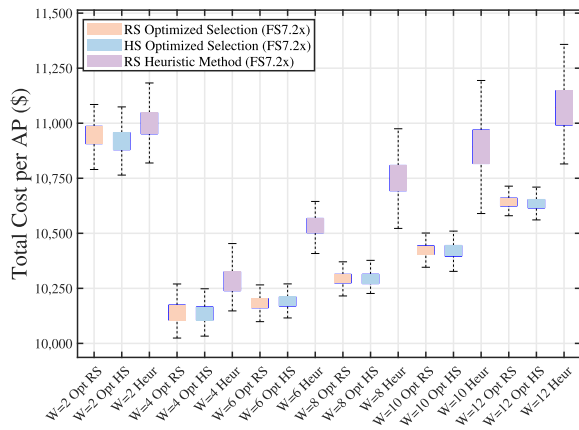


Fig. 17. Distribution of TCO per AP over 500 network realizations for different DU counts and $G = 150$, comparing optimized and heuristic solutions.

G. Variance Analysis Across Randomized Deployments

To complement the average-case performance results presented earlier in this section, we now investigate the per-realization variability of the proposed optimization framework and compare it against the heuristic method. This analysis is motivated by the need to understand how consistently the solution performs across randomized network topologies. Figure 17 presents a boxplot showing the distribution of total cost per AP across 500 randomized realizations for both the optimized and heuristic approaches under FS 7.2x, different DU pooling configurations ($W = 2, 4, \dots, 12$), and $G = 150$. Each W value includes three configurations: the optimized RS method, the optimized HS method, and the heuristic method and shows the spread, median, and interquartile range of deployment cost.

As observed, the optimized method consistently yields lower median cost and narrower variance across all W values. In contrast, the heuristic method exhibits greater variability and a wider spread, especially at higher W . This trend highly supports our results in Figure 9, highlighting that as the number of DU pools increases, the complexity of the planning problem grows, and the benefits of joint optimization over simple heuristics become more pronounced. The variance gap between these methods increases with W , reinforcing the value of global cost-aware planning in more decentralized deployment scenarios. This confirms that the optimized strategy not only lowers average cost but also reduces the risk of extreme deployment scenarios, providing stronger guarantees for real-world planning under uncertainty. The figure also shows that RS and HS architectures have comparable costs across different network realizations; however, the HS advantage lies in the greater resilience to AP failures, as demonstrated in Figure 16.

VI. CONCLUSION

This paper presented a two-tiered optimization framework for hybrid fronthaul network planning that minimizes TCO while ensuring scalability and high performance in HS- and RS-enabled CF-mMIMO, as well as broader UDN deployments within O-RAN. The framework jointly optimizes fiber, mmWave, and FSO fronthaul options under capacity and

reliability constraints, achieving efficient infrastructure utilization and robust performance. Results demonstrate that hybrid deployments outperform single-technology (all-fiber or all-mmWave) and heuristic schemes by balancing cost, capacity, and resilience. HS-enabled CF-mMIMO further enhances reliability compared to RS-based systems, offering a cost-efficient yet failure-resilient architecture. An important extension of this framework is the explicit incorporation of future traffic growth or reserve capacity requirements into the optimization itself, for example through location-dependent demand scaling or minimum headroom constraints. Such extensions would enable a direct trade-off between cost efficiency and future-proofing and are left for future investigation. These findings underscore the need to balance economic and performance objectives when designing future fronthaul networks, providing actionable insights for service providers in deploying scalable and cost-effective UDNs that meet evolving 6G and O-RAN requirements.

REFERENCES

- [1] S. Parkvall et al., "5G NR release 16: Start of the 5G evolution," *IEEE Commun. Standards Mag.*, vol. 4, no. 4, pp. 56–63, Dec. 2020.
- [2] H. Q. Ngo, G. Interdonato, E. G. Larsson, G. Caire, and J. G. Andrews, "Ultradense cell-free massive MIMO for 6G: Technical overview and open questions," *Proc. IEEE*, vol. 112, no. 7, pp. 805–831, Jul. 2024.
- [3] C.-X. Wang et al., "On the road to 6G: Visions, requirements, key technologies, and testbeds," *IEEE Commun. Surv. Tut.*, vol. 25, no. 2, pp. 905–974, 2nd Quart., 2023.
- [4] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.
- [5] J. G. Andrews, T. E. Humphreys, and T. Ji, "6G takes shape," *IEEE BITS Inf. Theory Mag.*, vol. 4, no. 1, pp. 2–24, Mar. 2024.
- [6] M. A. Habibi, M. Nasimi, B. Han, and H. D. Schotten, "A comprehensive survey of RAN architectures toward 5G mobile communication system," *IEEE Access*, vol. 7, pp. 70371–70421, 2019.
- [7] M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "Understanding O-RAN: Architecture, interfaces, algorithms, security, and research challenges," *IEEE Commun. Surv. Tut.*, vol. 25, no. 2, pp. 1376–1411, 2nd Quart., 2023.
- [8] F. Conceição, M. Gomes, V. Silva, R. Dinis, and C. H. Antunes, "Joint spectral and power efficiency optimization in uplink radio stripes," *IEEE Trans. Commun.*, vol. 72, no. 8, pp. 5209–5225, Aug. 2024.
- [9] P. Franger, J. Hederen, M. Hessler, and G. Interdonato, "Improved antenna arrangement for distributed massive MIMO," WO Patent 2 018 103 897 A1, World Intellectual Property Organization, 2018. [Online]. Available: <https://patents.google.com/patent/WO2018103897A1/en>
- [10] U. Demirhan and A. Alkhateeb, "Enabling cell-free massive MIMO systems with wireless millimeter wave fronthaul," *IEEE Trans. Wireless Commun.*, vol. 21, no. 11, pp. 9482–9496, Nov. 2022.
- [11] P. Agheli, M. J. Emadi, and H. Beyranvand, "Designing cost- and energy-efficient cell-free massive MIMO network with fiber and FSO fronthaul links," *AUT J. Electr. Eng.*, vol. 53, no. 2, pp. 159–170, 2021.
- [12] I. Chiotis and A. L. Moustakas, "Uplink performance optimization of limited-capacity radio stripes," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 12382–12395, Sep. 2024.
- [13] A. Azarbahram, O. L. A. López, P. Popovski, and M. Latva-Aho, "On the radio stripe deployment for indoor RF wireless power transfer," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2024, pp. 1–6.
- [14] A. S. Mohammed, H. A. Ammar, K. S. Tharakan, H. ElSawy, and H. S. Hassanein, "Traffic-aware cost-optimized fronthaul planning for ultra-dense networks," *IEEE Wireless Commun. Lett.*, vol. 14, no. 2, pp. 529–533, Feb. 2025.
- [15] Z. H. Shaik, E. Björnson, and E. G. Larsson, "MMSE-optimal sequential processing for cell-free massive MIMO with radio stripes," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7775–7789, Nov. 2021.
- [16] A. Fayad, T. Cinkler, and J. Rak, "Toward 6G optical fronthaul: A survey on enabling technologies and research perspectives," *IEEE Commun. Surv. Tut.*, vol. 27, no. 1, pp. 629–666, Feb. 2025.

- [17] Ö. T. Demir, M. Masoudi, E. Björnson, and C. Cavdar, "Cell-free massive MIMO in O-RAN: Energy-aware joint orchestration of cloud, fronthaul, and radio resources," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 2, pp. 356–372, Feb. 2024.
- [18] *NR; Physical Channels and Modulation (Release 18)*, document TS 38.211, 3GPP, Jul. 2025.
- [19] O-RAN Alliance. *O-RAN Working Group 4 (Open Fronthaul Interfaces WG) Control, User and Synchronization Plane Specification*. Accessed: Oct. 2025. [Online]. Available: <https://specifications.o-ran.org/specifications>
- [20] *Study on Channel Model for Frequencies From 0.5 to 100 GHz (Release 18)*, document TR 38.901, 3GPP, 2024.
- [21] M. Alzenad, M. Z. Shakir, H. Yanikomeroglu, and M.-S. Alouini, "FSO-based vertical backhaul/fronthaul framework for 5G+ wireless networks," *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 218–224, Jan. 2018.
- [22] X. Zhu and J. M. Kahn, "Free-space optical communication through atmospheric turbulence channels," *IEEE Trans. Commun.*, vol. 50, no. 8, pp. 1293–1300, Aug. 2002.
- [23] D. Schulz et al., "Robust optical wireless link for the backhaul and fronthaul of small radio cells," *J. Lightw. Technol.*, vol. 34, no. 6, pp. 1523–1532, Mar. 15, 2016, doi: [10.1109/JLT.2016.2523801](https://doi.org/10.1109/JLT.2016.2523801).
- [24] J. Tang et al., "Cooperative ISAC-empowered low-altitude economy," *IEEE Trans. Wireless Commun.*, vol. 24, no. 5, pp. 3837–3853, May 2025.
- [25] *Final Catalog of Eligible Expenses and Estimated Costs*, FCC, Sep. 2021. [Online]. Available: <https://docs.fcc.gov/public/attachments/DA-21-947A4.pdf>
- [26] *Gurobi Optimizer Reference Manual*, Gurobi Optimization, LLC, Mar. 2026. [Online]. Available: https://docs.gurobi.com/_downloads/optimizer/en/current/pdf/



Anas S. Mohammed (Student Member, IEEE) received the B.Sc. degree (Hons.) in electrical engineering, with a concentration in energy efficiency from the King Fahd University of Petroleum and Minerals (KFUPM), Saudi Arabia, in 2023, and the M.A.Sc. degree in electrical and computer engineering from Queen's University, Canada, in 2025. During his graduate studies, he was a Graduate Research Fellow with the Telecommunications Research Laboratory (TRL) and the Smart Wireless Massive Systems Laboratory (SWIMS), Queen's

University. His background spans both advanced research and industry practice. His research interests lie in the broad areas of electrical engineering, including wireless communication systems, mobile networks, electrical and electronic systems design, power systems, smart grids, networks strategic planning, and AI-driven optimization. Over the years, he conducted several hands-on research projects in the broad fields of electrical engineering, resulting in high-quality publications, and participation in global conferences, events, and demonstration sessions.



Krishnendu S. Tharakan (Member, IEEE) received the B.Tech. degree in electronics and communications engineering from the National Institute of Technology Calicut, India, in 2016, and the Ph.D. degree from the Electrical Engineering Department, IIT Indore, India, in 2023. From 2016 to 2017, she was an Engineer with Tata Elxsi, Trivandrum, India. She was a recipient of Visvesvaraya Ph.D. Fellowship from MeitY, Government of India. She is currently a Post-Doctoral Fellow with the School of Electrical Engineering and Computer Science,

KTH Royal Institute of Technology, Stockholm, Sweden. Prior to that, she was a Post-Doctoral Fellow with the School of Computing, Queen's University, Kingston, Canada. Her current research interests include wireless communications, distributed optimization, federated learning, and statistical learning theory. She was recognized as an Excellent Reviewer of IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING in 2024.



Hussein A. Ammar (Member, IEEE) received the M.A.Sc. degree in electrical and computer engineering from the American University of Beirut (AUB) in 2017 and the Ph.D. degree in electrical and computer engineering from the University of Toronto in 2023. In 2023, he completed a four-month internship with Ericsson on developing artificial intelligence (AI) solutions for wireless communications. In 2018, he worked as a Research Assistant with the Mobile and Distributed Computing Laboratory, AUB, and as a Research and Development Engineer in the information and communications technology industry. Since 2024, he has been an Assistant Professor with the Department of Electrical and Computer Engineering, Royal Military College of Canada. His research interests include wireless communications, AI for wireless networks, scalable and resilient deep reinforcement learning, coordinated distributed MIMO, user-centric cell-free massive MIMO, statistical signal processing, and mathematical optimization. He received the University of Toronto Fellowship and the Edward S. Rogers Sr. Graduate Scholarship. He was recognized as an Exemplary Reviewer of IEEE COMMUNICATIONS LETTERS in 2023.



Hesham ElSawy (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Manitoba, Canada, in 2014. He is currently an Associate Professor with the School of Computing, Queen's University, Kingston, ON, Canada. Prior to that, he was an Assistant Professor with the King Fahd University of Petroleum and Minerals (KFUPM), Saudi Arabia, a Post-Doctoral Fellow with the King Abdullah University of Science and Technology (KAUST), Saudi Arabia, and a Research Assistant with TRTech, Winnipeg, MB,

Canada. He conducts research in the broad area of wireless communications and networking, with a special focus on 5G/6G networks, joint communications and sensing, the Internet of Things, edge computing, non-terrestrial networks, and wireless security. He was a recipient of the IEEE ComSoc Outstanding Young Researcher Award for Europe, Middle East, and Africa Region in 2018. He also received several best paper awards, including the IEEE COMSoc Best Tutorial Paper Award in 2020 and IEEE COMSoc Best Survey Paper Award 2017. He is an Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, and the IEEE COMMUNICATIONS LETTERS.



Hossam S. Hassanein (Fellow, IEEE) is currently a leading Researcher in the areas of broadband, wireless and mobile networks architecture, protocols, control, and performance evaluation. His record spans more than 700 publications in journals, conferences, and book chapters; in addition to numerous keynotes and plenary talks in flagship venues. He has received several recognition and best paper awards at top international conferences. He is also the Founder and the Director of the Telecommunications Research Laboratory (TRL), School of Computing,

Queen's University, with extensive international academic and industrial collaborations. He was a recipient of the 2016 IEEE Communications Society Communications Software Technical Achievement Award for outstanding contributions to routing and deployment planning algorithms in wireless sensor networks and the 2020 IEEE IoT, Ad Hoc and Sensor Networks Technical Achievement and Recognition Award for significant contributions to technological advancement of the Internet of Things, ad hoc networks, and sensing systems. He is the Former Chair of the IEEE Communication Society Technical Committee on Ad Hoc and Sensor Networks (TC AHSN). He is an IEEE Communications Society Distinguished Speaker [a Distinguished Lecturer (2008–2010)].