Fronthaul Compression Control for Shared Fronthaul Access Networks

Sandra Lagén, Xavier Gelabert, Andreas Hansson, Manuel Requena, and Lorenza Giupponi

ABSTRACT

There is a widely held belief that future radio access network architectures will be characterized by increased levels of virtualization, whereby base station functionalities, traditionally residing at a single location, will be scattered across different logical entities while being interfaced via high-speed fronthaul (FH) links. For the deployment of such FH links, operators are faced with the challenge of maintaining acceptable radio access performance while at the same time keeping deployment costs low. A common practice is to exploit statistical multiplexing by allowing several cells to utilize the same FH link. As a result, in order to cope with the resulting aggregated traffic, different techniques can be used to reduce the required FH data rates. Herein, we focus on FH compression control strategies for multiple-cell/multiple-user scenarios sharing a common FH link. We propose various methods for sounding reference signal (SRS) handling, and analyze different FH-aware modulation data compression and scheduling strategies. Considering a full system setup, including the radio and FH access networks, numerical evaluation is conducted using a 5G NR system-level simulator implemented in rs-3. Simulation results show that under stringent FH capacity constraints, optimized modulation compression strategies provide significant user-perceived throughput gains over baseline strategies (between 5.2× and 6.9×). On top of them, SRS handling methods achieve additional 2 to 41 percent gains.

INTRODUCTION

The design of future radio access networks (RANs) often needs to fulfill the demanding requirements across different and competing axes. While increasing the spectral efficiency has a major impact on the perceived quality of service for the end user, implementing a scalable and low-power solution is often regarded as important for network operators. Furthermore, efficient use of computing resources via pooling and the ability to provide cross-layer solutions are also desired features. With this in mind, the centralized RAN (C-RAN) architecture paradigm [1] has emerged and is being considered by the 3rd Generation Partnership Project (3GPP) and Open-RAN (O-RAN) as a key design alternative in next-generation RANs. Among its features, C-RAN advocates for the disaggregation of baseband processing functions between different physical entities that can be either distributed and residing close to the antenna or centralized in a given location. Specifically, a base station, also known as a gNB in 3GPP 5G New Radio (NR), will break up into a centralized unit (CU) communicating with at least a distributed unit (DU) via the so-called midhaul interface [2]. In turn, several radio units (RUs) will interface toward a DU via the fronthaul (FH) [3]. When designing and deploying C-RAN, it is important to consider both capacity constraints and latency requirements of the FH [4]; more so considering the increased bandwidths in 5G NR, in addition to antenna densification, increased modulation orders, and enhanced carrier aggregation features [5]. All of these contribute to the increase in the required FH capacity [6].

In general, the dimensioning of the FH may respond to the peak rate requirements given during the planning phase of a specific network technology deployment (e.g., 4G LTE). Nonetheless, under normal network operation, several reasons cause the FH to undergo resource underprovisioning at specific moments in time [7]. One example is the ever growing adoption of new features, new functional splits, or new algorithms requiring increased information exchange between RUs and DUs/CUs. Other examples are to allow a seamless rollout of new radio access technologies (e.g., 5G NR) or to facilitate a continuous layout of low-power small cells in specific traffic-demanding areas while at the same time allowing this new data to be handled through the available and pre-dimensioned FH network. From the equipment vendor’s viewpoint, some interest may be rooted in offering a set of new features requiring minimal disturbance on the existing FH network, where only software updates would be necessary to upgrade firmware, algorithms, and so on. On the contrary, a dimensioning change in the FH network may be seen as a costly measure, that is, by exchanging optical interface adapters, network switches, and, at worst, optical fiber layouts themselves (e.g., switching from single-mode to multi-mode). All the above provides a good motivation to consider the case where the FH can run into a capacity underprovisioning problem. To lessen the demand in FH capacity and address the above-mentioned FH underprovisioning problem, FH compression schemes become essential [1]. Briefly, FH compression involves the partial reduction, or total removal, of information sent over the FH. FH compression methods...
have received wide attention from both the information theory and signal processing communities, in particular after the emergence of C-RAN architectures in 4G LTE. Recently, more practical schemes have been defined in 3GPP 5G NR and O-RAN [3]. Among the envisioned techniques, modulation compression is highlighted given its ability to reduce the required FH capacity due to constraining the modulation order. This can be effectively achieved with no degradation of the transmitted samples sent over the FH network and with reduced algorithmic complexity [6].

Besides FH compression, aimed at jointly reducing the deployment and operational costs along with fostering deployment scalability, the ability to multiplex and aggregate data from multiple cells over a single shared FH interface is of great interest to mobile network operators [8]. In this case, the same fixed FH link is shared by multiple cells (i.e., data from/to several RUs are multiplexed by using a layer-2 switch over a single FH link toward/from their respective DUs), as shown in Fig. 1. These scenarios present multiple technical challenges arising from the shared FH use by data originating from/to different cells. Essentially, effective FH compression control schemes have to especially consider and exploit the fact that many cells share a fixed capacity full duplex FH link, and should enhance the multiplexing gain by using clever combinations of different compression techniques, providing efficient methods to control the use of FH resources while minimizing air interface performance degradation. Consequently, an evaluation approach relying on end-to-end system-level simulations in a multi-cell environment is carried out in this work. In [9], baseline along with improved FH-aware packet scheduling methods of dynamic modulation compression were derived. Therein, both the scheduling and modulation compression decisions were dynamically adjusted according to the monitored FH capacity.

When considering the full system, even if we focus on downlink data transmission, the main part of the data bulk sent through the FH interface comes from the physical downlink shared channel (PDSCH) (used to send data) and the uplink sounding reference signals (SRSs) (used to estimate the channel). Modulation compression-based methods reduce the PDSCH bulk part in the FH downlink link. However, when using SRS-based channel estimation for beamforming/precoding design, as standardized for time-division duplex (TDD) 5G systems, the FH uplink utilization can be reduced by selectively compressing/removing unnecessary SRSs, which may further impact the beamforming/precoding design in the downlink. Here, SRS handling methods can alleviate the FH uplink load by properly handling the allocation of SRS signals through the full duplex FH interface. In this article, different from [6] where disaggregated architectures and FH compression methods were reviewed, and from [9] where dynamic modulation compression methods were proposed to reduce only the PDSCH bulk, we provide a summary and a new vision on the integral solutions for dynamic FH compression control in shared FH architectures. In particular, we review methods for FH-aware downlink data scheduling, and we provide a new study for SRS handling methods in the uplink. Finally, the numerical evaluation of the aforementioned schemes is carried out using a seasoned dynamic system-level simulator developed in ns-3 [10].

The rest of the article is structured as detailed hereafter. We discuss the system model and overview FH compression methods, with special attention to SRS handling methods for uplink SRSs. We describe the simulation scenario and assess the end-to-end performance. Finally, we highlight future research lines and conclude the article.

**Fronthaul Compression Control**

This section introduces the system model and discusses solutions for shared FH architectures integrating FH-aware scheduling methods and SRS handling methods, to reduce the FH downlink and FH uplink loads, respectively.

**System Model**

We consider herein a multiple-cell/multiple-user-equipment (UE) cellular deployment following a C-RAN architecture and consisting of co-located CUs/DUs and geographically distributed RUs. We consider the PDCP-RLC split (a.k.a. Option 2) for the CU-DU [2] and the intra-PHY split (Option 7.2x) for the DU-RU [3], as per 3GPP and O-RAN specifications, respectively. Regarding the CU/DU/RU deployment, we follow the so-called Scenario B, as highlighted by O-RAN [11]. In this case, CUs and DUs (for all cells) are located together in a centralized location (edge or regional cloud), whereas RUs are scattered at operator-owned cell sites. Consequently, the centralized entity implements the high-PHY, medium access control (MAC), and above processing for all cells, while the RU of each site executes the low-PHY and RF processing of every cell [12]. DUs are interconnected with the RUs through a low-layer full-duplex FH interface of limited capacity in downlink and uplink directions, according to a star FH topology [3]. This way, RUs share the same full duplex FH link. The deployment scenario is illustrated in Fig. 1.

We focus on FH compression control for downlink data transmission in multi-cell TDD systems with a shared full duplex FH interface. In particular, we:

1. Analyze FH-aware scheduling methods
FIGURE 2. SRS transmission through the full-duplex FH interface, when using a TDD pattern of (D D D F U) in the air interface (i.e, three downlink slots, followed by a flexible slot and one uplink slot) and SRSs being sent in the F slots: a) symbol-by-symbol transmission; b) prioritized FH transmission; c) prioritized FH transmission with partial SRS transmissions.

Based on modulation compression to meet the available FH downlink capacity
2. Propose SRS handling methods to meet the available FH uplink capacity

Downlink FH-Aware Scheduling Methods for PDSCH
In typical cellular systems, the downlink modulation and coding scheme (MCS) for each UE is determined based on the reported channel quality indicator (CQI). Given the MCSs and the amount of data in the RLC buffers, the MAC scheduler decides the number of resource blocks (RBs) allocated to each UE by following specific scheduling rules (e.g., proportional fair). To meet the shared FH downlink capacity constraint, several FH-aware scheduling methods have been proposed in the literature. Two baseline options are:
1. Dropping packets at the PHY layer
2. Postponing scheduling decisions at the MAC layer [9]

Another option, is to limit the MCS per cell [6] using modulation compression. To enhance these solutions, centralized optimized methods were proposed in [9], in which the resource allocation (i.e., number of RBs) and modulation compression (i.e., MCS) of each UE are dynamically set. These methods are reviewed in what follows.

Drop Packets at High-PHY: Assuming typical MCS selection and RB assignment at the MAC scheduler, packet dropping can be implemented at the high-PHY layer in the DUs. In this case, a centralized logic decides to drop those MAC packet data units (PDUs) (including new data and/or hybrid automatic repeat request, HARQ, retransmissions and their associated control across all cells) that cannot fit in the available shared FH downlink capacity [9].

Postpone Scheduling Decisions at MAC: Assuming typical MCS selection and RB assignment at the MAC scheduler, discarding/dropping MAC scheduling decisions can be executed at the MAC layer in the DUs. This way, data is not dropped, but its transmission (or retransmission) is postponed. In this case, a centralized logic decides to drop/discard those scheduling decisions (related to new data and/or HARQ retransmissions across all cells) that cannot fit in the available shared FH downlink capacity, for which its associated transmission is postponed [9].

MCS Limits at MAC: By exploiting semi-static modulation compression [6], the system can limit the maximum MCS (and thus the maximum modulation order) that is allowed per cell according to the available shared FH downlink capacity. Note that per-cell MCS limits can be combined with dynamic methods, like drop and postpone strategies. In particular, their joint operation may result in fewer packet drops and scheduling decision postponements because of the inherent FH load reduction when using lower MCSs.

Resource Allocation and MCS Optimization at MAC: By using dynamic modulation compression, a centralized control entity can manage the MAC schedulers of all the cells (placed in the collocated DUs in Fig. 1) and determine the most appropriate MAC scheduling decisions (including scheduling of users, MCS assignment, and resource allocation) across all cells dynamically so that the shared FH downlink capacity is properly exploited and certain QoS per user is satisfied. In particular, two solutions were derived in [9] to optimize the RB allocation and the modulation compression applied to each UE of each cell for every time instant.

Uplink FH-Aware Handling Methods for SRSs
In TDD systems, SRSs can be used for beam management in downlink and uplink, due to beam and channel reciprocity. In particular, SRS receptions at the base station are typically used to estimate the channel and then determine the beamforming/precoding. To get an accurate acquisition of the SRS signal, the bulk needed to send SRSs through the FH uplink interface (RU-to-DU) can be very large, since its size depends on the number of antennas used for channel estimation and the number of resource elements carrying SRS samples.

A key observation is that SRS signals, different from downlink/uplink data in PDSCH/PUSCH, are non-delay-sensitive and do not need to be transmitted through the FH as soon as they arrive at the RU. Basically, we can exploit the fact that the FH interface is full duplex, while the air interface is half duplex; thus, by leveraging on the TDD pattern, SRSs can be sent through the FH uplink interface when there is a downlink slot.

Figure 2 illustrates examples of the impact on the peak FH bandwidth requirements of SRS transmissions, assuming that for a given SRS signal, a certain amount of bits needs to be processed at the DU before some target deadline. There are two main examples of how to meet the latency constraint:
• Symbol-by-symbol FH transmission (Fig. 2a): After each symbol is received at the RU, it is immediately transmitted over the FH. The required FH capacity is then given by the maximum between the SRS and PUSCH bulk requirements. This option causes high load peaks on the FH, and makes it so that the dimensioning of the required FH capacity is determined by SRS bulks, which require more samples than PUSCH, especially for multi-antenna and wide-bandwidth systems.
• Prioritized FH transmission (delayed transmission) (Figs. 2b and 2c): A certain delay is allowed when conveying SRSs over the FH. For example, in Fig. 2b, PUSCH transmission is prioritized, and SRS samples are buffered...
Prioritized FH transmissions reduce the FH uplink capacity requirement at the cost of a different delay on the reception of the full SRS signal at the DU. This consequently implies a delay on the beamforming/precoding update (illustrated by the light blue downlink slot in Fig. 2), which may affect the downlink end-to-end performance and should be properly evaluated through system-level simulations. In the case of multiple RUs sharing the same FH interface, and potentially overlapping SRS transmissions (e.g., because different cells/RUs use the same TDD pattern), SRS handling methods need to be designed. Two options appear: time multiplexing or frequency multiplexing. In the frequency multiplexing option, all the SRS bulks experience the same FH delay, as shown in Figs. 3a and 3b. We can use a worst case partition of the available FH bandwidth among the multiple RUs that share the FH (as shown in Fig. 3a, for the case of 3 RUs sharing the FH interface, where only 2 RUs send SRSs at the same time). In this case, the FH capacity may not be fully exploited in the uplink direction. Otherwise, we can adopt a dynamic FH bandwidth allocation to the RUs that have data to be sent on a particular slot (as shown in Fig. 3b), where the total delay can be reduced compared to the hard bandwidth distribution option. On the other hand, through the time multiplexing option, dynamicity is naturally achieved. In this case, SRS bulks are sent sequentially (Fig. 3c) so that some of them are received more quickly at the DU/CU for processing. This option allows the beam update to be done sooner compared to the frequency multiplexing options. Here, SRS priority handling methods are needed to decide the order/priority to send the UEs’ SRS bulks.

End-to-End Simulation

For the evaluation, the ns-3 5G-LENA system-level simulator is used [10]. We extended the 5G-LENA simulator with a new centralized intelligence that controls the MAC/PHY operations of all the DUs and implements the proposed FH-aware scheduling procedures and SRS handling methods.

Scenario

We consider a hexagonal site deployment with three sites, according to an Urban Micro scenario. Each site is composed of 3 cells and 3 uniform planar antenna arrays, covering 120° in azimuth each. Frequency reuse 1 is assumed. The rest of the deployment and network parameters are:

- Number of cells (RUs): 9
- Number of UEs per cell: 10
- Inter-site distance: 200 m
- RU antenna height: 10 m
- RU transmit power: 30 dBm
- RU antenna: 5 × 2 directional elements
- UE antenna height: 1.5 m
- UE antenna: 1 isotropic element
- Carrier frequency: 2 GHz

Bandwidth: 100 MHz

Numerology: 1 (30 kHz subcarrier spacing)

RB overhead: 0.04

Duplexing mode: TDD, with pattern [D D D F U].

SRS: in F slot, spanning over 1 orthogonal frequency-division multiplex (OFDM) symbol. Two SRS periodicities:

- 50 ms (SRS config1)
- 25 ms (SRS config2)

MAC scheduler: Round-Robin

MCS Table: 2 (up to 256-QAM: quadrature amplitude modulation)

Channel update period: 40 ms

HARQ: incremental redundancy, 20 HARQ processes

RLC: unacknowledged mode

Transport protocol: UDP

Traffic: File Transport Protocol (FTP) Model 1

- File generation rate: 50 files/s
- FH: start topology, shared full-duplex FH link of 0.5 Gb/s capacity in each direction (downlink/uplink)
- Simulation duration: 10 s

As the key performance indicator we consider the user-perceived throughput (UPT), measured at the IP layer. The UPT corresponds to the fraction between the received bytes per file and the time period needed to complete the file transfer.

Results

In the end-to-end evaluation, we assess the impact of using different FH compression control methods. For downlink data, we consider the following FH-aware scheduling methods (described earlier):

- Drop: High-PHY drop of MAC PDUs
- Postpone: Discard MAC scheduling decisions
- RB: RB assignment optimization per active UE
- MCS: MCS optimization per active UE

We evaluate each strategy in combination with the following SRS handling methods (detailed earlier):

- fixedDelay: The FH uplink bandwidth is equally distributed among all the cells that share the FH.
- dyndelay freqMux: The FH uplink bandwidth is equally distributed among active cells (delay is time-dependent).
- dyndelay timeMux: The FH uplink bandwidth is fully used by each SRS bulk (delay is time-dependent).

Figure 3 shows the cumulative density function (CDF) of the UPT (in megabits per second) when using SRS config, for different scheduling state-
The benefits are significant in all percentiles. The range of UPT is from 5.2× to 6.9× in the mean UPT over the Postpone strategy. Indeed, the MCS strategy outperforms the RB strategy in the 5th percentile UPT because, by reducing the modulation order, higher robustness is achieved against propagation/interference variations. Conversely, RB strategy outperforms MCS in the 95th percentile UPT, because it provides a more efficient RB distribution due to a fine-grained control mechanism, thus enabling a larger amount of data to be served.

Regarding the SRS handling methods, we observe that the impact of using beams which are not well adjusted to the channel is appreciable in the RB optimization strategy, while for other strategies the impact is reduced (Figs. 4 and 5). This is because Drop/Postpone strategies are dropping/postponing many packets, so the performance is dominated by the losses; while the MCS strategy is more robust to signal-to-interference-plus-noise ratio (SINR) degradation.

The impact of the different SRS handling methods on the UPT depends on the SRS configuration. Specifically, in SRS config1 (Fig. 4), both the frequency and time multiplexing with dynamic adaptation improve the UPT performance, as do the worst case frequency multiplexing, and both provide similar performance. This is because with the considered frame pattern and SRS periodicity (50 ms), there are 20 available slots for SRSs within the SRS periodicity, and the deployment considers 10 UEs per cell. Therefore, not all the cells are active at each SRS opportunity. In this way, the dynamic freqMux also allows reducing beam update delays compared to the worst case freqMux option, and gets similar UPT performance as the timeMux option, which has lower delays, because SRS bulks are sent one after the other. Under SRS config2, we have considered the same scenario but with an SRS periodicity of 25 ms, which results in 10 opportunities for SRSs. All SRS opportunities are then used by one of the UEs to send SRSs. In this case, as expected and shown in Fig. 5, the two frequency multiplexing options (with fixed or time-dependent delays) achieve the same end-to-end performance, because the two options are equivalent. Interestingly, the time multiplexing option outperforms both frequency multiplexing options, since a major part of the SRS bulks experience lower delay updates. In summary, dynamic SRS handling methods achieve mean UPT gains ranging from 2 to 41 percent.

Finally, the presented results have allowed interesting observations regarding the behaviors of FH-scheduling and SRS-handling methods as a function of the type of served UEs (i.e., cell edge, cell middle, or cell center). A summary of the main conclusions is shown in Table 1. For cell-edge UEs, the best option is to use MCS strategy, and the SRS handling method does not impact the performance. For cell middle and cell center UEs, the best option is to use RB strategy in downlink. In uplink, the cell center UEs are not affected by the SRS handling method, while for cell middle UEs, the recommended option is the time multiplexing.

**Future Research Directions**

Based on the presented study, analysis, and...
obtained simulation results, we envision the following research lines:

- **SRS priority handling methods**: The results from earlier show that users in different conditions (cell edge/middle/center) are affected differently by the delay updates of the beams. Accordingly, if a time-multiplexing option is adopted, clever SRS priority handling methods for the SRS bulks have to be defined. Under shared FH capacity, SRS bulks associated with specific UEs could be prioritized to reduce their delays. For example, when using an RB strategy, cell edge UEs could be prioritized, because they are more affected by a delay increase in the beam update. Instead, when using a combined MCS/RB strategy, cell middle UEs could be prioritized. A control entity at the FH interface could implement the SRS priority handling method by knowing the SINR associated with each SRS bulk to distinguish among cell center/middle/edge UEs.

- **SRS control methods**: The control entity could also keep track of the actual FH delay for each SRS bulk, leading to SRS control methods. For example, if the buffering delay surpasses the channel coherence time, such an SRS bulk could be dropped from the buffer of packets to be sent through the FH, because when the DU receives it, the measurement will already be outdated. This would leave FH capacity available for other transmissions.

- **SRS priority handling methods with partial transmissions and partial beam updates**: Partial transmissions of the SRS blocks, as shown in Fig. 2c, constitute an improvement for shared FH interfaces. In this case, part of the SRS bulks of different UEs can be sent earlier so that the DU/CU can do a first channel estimation and beamforming update with partial information (e.g., half or part of the SRS samples). Later, once the full SRS information is sent through the FH, a second beamforming update can be implemented based on complete information.

- **Uplink data FH compression**: In the present study, we have focused on compression of FH information related to downlink data transmissions, for which the downlink data (in downlink) and SRS (in uplink) constitute the bulk FH part. Future studies could include compression of uplink data in PUSCH.

- **Joint flexible splits and FH compression**: There has been wide interest in flexible functional split selection recently. However, the interaction between the split selection and the scheduling/resource allocation strategies has been less studied. An interesting area for further research is to analyze joint strategies that optimize the functional split and the FH compression control for shared FH multi-cell scenarios.

**Conclusions**

In this article, we have presented an integral design and a thorough end-to-end evaluation of shared FH scenarios where multiple FH compression control techniques are proposed. In particular:

1. We have analyzed FH-aware scheduling methods to compress user data that goes through the downlink FH.

2. We have proposed SRS bulk handling methods to handle uplink SRS bulks that go through the uplink FH.

Then end-to-end simulations over a 5G-aligned scenario have been presented. In multi-cell scenarios with shared FH link, we have evaluated the impact of four main FH-aware scheduling methods for downlink data compression — Drop, Postpone, RB, and MCS — combined with three methods for SRS handling: fixed frequency multiplexing, dynamic frequency multiplexing, and dynamic time multiplexing. Results have shown that when there is tight FH capacity, centralized and optimized scheduling strategies (MCS and RB methods) are essential to maintain an acceptable end-to-end user experience. SRS handling methods are shown to affect the RB optimization strategy, for which our results have exhibited that the time-multiplexing option always provides the best performance and improves all the other SRS handling methods for configurations in which all the SRS opportunities are used to send SRBs. However, when not all the SRS opportunities are used to send SRBs, dynamic frequency multiplexing can also achieve similar performance. Interestingly, the degradation in the end-to-end performance depends on the quality/condition of the target UE, and it is more pronounced in the cell edge/middle users, which get a lower SINR in the downlink as a result of the path loss degradation and larger errors in the SRS-based channel estimation, for which future research lines have been highlighted.

Based on our findings above, we advocate for the following recommendations. First, operators and vendors should seriously consider the FH underprovisioning problem, whereby a properly dimensioned FH at the planning stage may become underprovisioned over time. Second, considering shared FH link segments is key to exploit multiplexing gains arising from traffic inhomogeneities. Third, when capacity-limited FH problems arise, leveraging FH compression strategies helps alleviating the problem. We conclude that dynamic compression of data is essential to maintain acceptable user experience, while at the same time noting that compression of reference signals (especially SRBs) plays a relevant role.

**Acknowledgments**

This work has been partially funded by Huawei Technologies and Spanish MINECO grant TSI-063000-2021-56/57 (6G-BLUR).

**References**

[1] M. Peng et al., “Recent Advances in Cloud Radio Access Networks: System Architectures, Key Technologies, and Open Issues,” *IEEE Commun. Surveys & Tutorials*, vol. 18, no. 3, 2016, pp. 2282–2308.

[2] 3GPP TS 38.401, TSG RAN, “NG-RAN; Architecture Description,” Release 16, v. 16.1.0, Mar. 2020.

[3] O-RAN Fronthaul Working Group, “Technical Specification; Control, User and Synchronization Plane Specification,” v. 03.00, Apr. 2020.
[4] L. M. P. Larsen, A. Checko, and H. L. Christiansen, “A Survey of the Functional Splits Proposed for 5G Mobile Crosshaul Networks,” IEEE Commun. Surveys & Tutorials, vol. 21, no. 1, 2019, pp. 146–72.

[5] S. Parkvall et al., “NR: The New 5G Radio Access Technology,” IEEE Commun. Standards Mag., vol. 1, Dec. 2017, pp. 24–30.

[6] S. Lagen et al., “Modulation Compression in Next Generation RAN: Air Interface and Fronthaul Trade-offs,” IEEE Commun. Mag., vol. 59, no. 1, Jan. 2021.

[7] M. Peng et al., “Fronthaul-Constrained Cloud Radio Access Networks: Insights and Challenges,” IEEE Wireless Commun., vol. 22, no. 2, Apr. 2015, pp. 152–60.

[8] A. Limesh et al., “Overview of O-RAN Fronthaul Specifications,” Special Articles on Standardization Trends towards Open and Intelligent Radio Access Networks, NTT Docomo Tech. J., vol. 21, July 2019.

[9] S. Lagen et al., “Fronthaul-Aware Scheduling Strategies for Dynamic Modulation Compression in Next Generation RANs,” IEEE Trans. Mobile Computing, Nov. 2021.

[10] 5G-LENA Project; https://5glena.ccttc.es/.

[11] O-RAN Alliance White Paper, “O-RAN Use Cases and Deployment Scenarios,” Feb. 2020.

[12] 3GPP TR 38.801, TSG RAN, “Study on New Radio Access Technology: Radio Access Architecture and Interfaces,” Release 14, v. 14.0.0, Mar. 2017.

[13] 3GPP TR 36.814, “Evolved Universal Terrestrial Radio Access (EUTRA); Further Advancements for EUTRA Physical Layer Aspects,” Release 9, v. 9.2.0, Feb. 2017.

[14] Samsung, “Application of Stochastic Geometry in Modeling Future LTE-A and 5G Wireless Networks”; https://web.ma.utexas.edu/conferences/simons2015/Modeling3GPPSamsungFinal.pdf.

**Biographies**

**Sandra Lagen** (sandra.lagen@ccttc.es) holds a Ph.D. from UPC (2016). She is a senior researcher and head of the Open Simulations research unit at CTTC.

**Xavier Gelabert** (xavier.gelabert@huawei.com) is a senior research engineer at Huawei Technologies Sweden AB. He has 15+ years of experience working across RAN L1, L2, and L3, as well as 3GPP standardization.

**Andreas Hansson** (andreas.hansson@huawei.com) is a principal baseband software engineer within Huawei Technologies Sweden AB, working with software architecture, modeling, and systemization, with 20 years of experience.

**Manuel Requena** (manuel.requena@ccttc.es) is a senior researcher at CTTC and responsible for the EXTREME Testbed of the Services as Networks research unit.

**Lorenza Giupponi** (lorenza.giupponi@ccttc.es) holds a Ph.D. from UPC (2007). She is a senior researcher at CTTC and a member of the CTTC Executive Committee.