

Enabling Large-Scale Networked Media Events through Cell-Free Massive MIMO

Alex Piccioni*, Claudia Rinaldi[†], Carlo Centofanti*, Roberto Valentini*, and Andrea Marotta*

**Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila, L'Aquila, ITALY*
{alex.piccioni, carlo.centofanti1, roberto.valentini, andrea.marotta}@univaq.it

[†]*National Inter-University Consortium for Telecommunications (CNIT), L'Aquila, ITALY* {claudia.rinaldi@cnit.it}

Abstract—Large live events such as concerts, arena shows, sport competitions, and large-scale public performances are becoming dense Internet of Sounds environments, where massive spectator traffic coexists with wireless sound and media production devices. This coexistence poses a severe challenge to radio access: spectators generate downlink-dominant best-effort demand, whereas production devices such as wireless microphones, immersive-audio capture nodes, cameras, and on-site media equipment require stable uplink contribution links to feed the live production chain. For networked-audio and immersive-media applications, satisfying a guaranteed bit rate serves both as a throughput objective and as a prerequisite for supporting continuous packetized media transport, buffer stability, and higher-layer timing requirements. In this paper, Cell-Free massive Multiple-Input Multiple-Output is studied as a radio access solution for jointly serving these two traffic classes in crowded live-event venues. We develop a system-level simulator that compares a conventional cellular deployment and a Cell-Free architecture under the same total number of antennas. The evaluation considers both spectator and production performance, while accounting for the impact of crowd-induced blockage. Results show that both architectures sustain the guaranteed-rate uplink, but Cell-Free massive MIMO maintains a substantially more uniform uplink service across the venue while preserving spectator connectivity, making it a promising 6G radio-access architecture for Internet of Sounds live production scenarios.

Index Terms—Cell-Free, Massive MIMO, Internet of Sounds (IoS), Networked Audio Production, Immersive Audio Capture.

I. INTRODUCTION

High-density live events, such as concerts, arena shows, sport competitions, and large-scale public performances, routinely gather tens of thousands of people within a few thousand square meters. Beyond the local reproduction of sound, these venues are increasingly evolving into Internet of Multimedia Things (IoMT) environments, where Internet of Sounds (IoS) applications coexist with video, sensing, and telemetry services over a common networked media infrastructure. Audio-visual content is captured, processed, transported, and redistributed through a media chain involving wireless microphones, immersive-audio capture nodes, cameras, telemetry units, and production equipment. This evolution turns the live venue into a highly heterogeneous wireless ecosystem, where massive consumer traffic coexists with professional

multimedia devices generating guaranteed-rate contribution streams that are essential to the continuity and quality of the production workflow.

Providing radio access in such scenarios is particularly challenging. Spectators are densely concentrated over a limited area and generate strong downlink (DL) best-effort demand for content access, social sharing, and interactive services. At the same time, the crowd itself becomes part of the propagation environment: human bodies introduce additional blockage and shadowing, especially for low-height links and dense in-venue deployments. Conventional macro-cellular solutions are therefore affected by both capacity pressure and geometry-dependent coverage limitations, which become critical when public connectivity and professional media contribution must be supported over the same physical venue [1].

Two fundamentally different classes of users share the venue. The first class is the large population of *spectators*, i.e., ordinary attendees whose traffic is DL-dominant, best-effort, and without stringent priority. The second class is a small set of *production devices* that generate the live audio-visual feed, including wireless microphones, networked and immersive-audio capture nodes, broadcast cameras, and on-site mixing and telemetry equipment. These devices form the *contribution links* of the media production chain [2]; their traffic is uplink (UL)-dominant and, unlike spectator traffic, requires a stable connection at a *guaranteed, constant* bit rate, since dropped or throttled streams directly corrupt the broadcast and the networked-media experience. The growing relevance of wirelessly captured and transported sound at live events, from networked music performances to immersive-audio capture, further raises the reliability requirements of these ULs within the emerging IoS [3]. Although packet-level jitter, clock distribution, and synchronization protocols are outside the scope of this work, a stable guaranteed-rate UL is the necessary access-layer condition on which networked-audio and immersive-media applications rely.

Cell-Free (CF) massive Multiple-Input Multiple-Output [4], [5] is a beyond-5G/6G radio-access paradigm in which many geographically distributed Access Points (APs), coordinated by one or more Central Processing Units (CPUs), jointly serve the UEs through coherent processing without conventional cell boundaries. Each User Equipment (UE) is served by a *user-centric* subset of nearby APs, providing macro-diversity and

This work has been supported by the European Union - NextGenerationEU under the Italian Ministry of University and Research (MUR) National Innovation Ecosystem grant ECS00000041 - VITALITY - CUP E13C22001060006.

mitigating the cell-edge effect typical of single-Base Station (BS) association. The comparison in [4] showed that, for an equal total number of antennas, CF massive MIMO can outperform both small cells and cellular massive MIMO, highlighting that the spatial distribution of radio resources is itself an architectural advantage, making CF a natural candidate for dense venues.

This architectural property is especially relevant for IoS live venues because CF massive MIMO naturally matches the spatial organization of distributed media production. Its distributed radio resources shorten the average distance between production UEs and serving APs, increase robustness against crowd-induced blockage, and better support IoS sensing and media acquisition. Most CF studies target spectral efficiency, scalability, and power control under nominal conditions [5], [6], typically assuming a homogeneous user population. The coexistence of a massive best-effort DL crowd and a limited number of guaranteed-rate UL contribution links has received little attention. Moreover, the conventional 3GPP Urban Macro channel adopted in much of the CF literature neglects the dominant propagation impairment of dense live events, i.e., crowd-induced body blockage. Consequently, the suitability of CF massive MIMO for IoS live-production scenarios remains largely unexplored.

To address these gaps, this paper formulates the uplink media-contribution problem in high-density live events, where a guaranteed-bit-rate UL slice must support professional production devices while the surrounding crowd simultaneously generates traffic and acts as a dense field of radio blockers. A channel model is developed by augmenting the 3GPP Urban Micro Street-Canyon propagation model with a geometric crowd-blockage component, in which human bodies are represented as a Poisson field of blockers whose impact depends on the relative elevation of the APs and the UEs. Finally, a vectorized system-level Monte Carlo framework is employed to compare CF and cellular massive MIMO under identical antenna and transmit-power budgets, quantifying the benefits of spatially distributing the radio resources across the venue and evaluating how these benefits evolve as the production fleet scales.

II. RELATED WORK

The IoS frames connected sound-related devices as an ecosystem for sensing, processing, exchanging, and actuating sound-related information [3]. In the musical domain, this vision is rooted in the Internet of Musical Things, where smart instruments, wearables, and connected musical devices support new forms of interaction among performers, audiences, and production infrastructures [7]. Networked music performance studies have further shown that real-time musical interaction imposes strict requirements on delay, audio quality, and synchronization, since temporal instability can directly affect musical coordination and perceived quality [8]. In professional environments, audio-over-IP standards such as AES67 specify interoperable mechanisms for high-performance packetized

audio transport, including synchronization, media clock identification, network transport, encoding, streaming, and session management [9].

In parallel, broadcast and media-production communities have identified 5G and beyond-5G systems as key enablers for professional contribution workflows, where wireless cameras, microphones, and audio devices generate upstream media feeds for live production [2], [10]. These works motivate the need for stable wireless contribution links, but they do not quantify how a dense radio-access infrastructure should jointly serve professional UL media flows and massive co-located spectator traffic. Beyond the media-production workflow itself, the wireless infrastructure supporting live events must operate under extremely challenging propagation and traffic conditions. Large venues such as stadiums, arenas, and concert halls are among the most demanding wireless deployment scenarios, due to extreme UE density, strong traffic heterogeneity, and severe spatial concentration of demand.

Existing dense-network studies and surveys generally focus on capacity provisioning, user association, interference, and spectral efficiency in cellular or small-cell systems [1]. However, in a packed live venue, the crowd is not only a traffic source but also part of the propagation environment. Human bodies introduce additional attenuation, shadowing, and geometry-dependent blockage, especially for low-height radio links and dense in-venue deployments. Human-body blockage has been modeled in the wireless literature as a function of blocker density, body dimensions, link distance, and transmitter/receiver height [11]; nevertheless, such models are rarely embedded into system-level evaluations of large live events with both public connectivity and professional media-production traffic. This gap is particularly relevant when comparing high concentrated cellular masts with low distributed APs, since the two architectures interact differently with crowd-induced blockage. Addressing these challenges requires radio-access architectures capable of providing both high capacity and uniform service quality across large and densely populated venues. In this context, CF massive MIMO has emerged as a promising distributed radio-access paradigm.

The CF paradigm originates from the network-MIMO and distributed-antenna literature and was formalized in [4], which showed that a large number of distributed single-antenna APs serving all UEs coherently outperforms small cells when the total antenna count is fixed. Scalability is achieved through user-centric clustering, where each UE is served only by a subset of APs [5], [12], and through practical combining/precoding such as Linear Minimum Mean Square Error (L-MMSE) and conjugate beamforming [5]. Surveys [1], [6] review spectral-efficiency, power-control, and fronthaul aspects. These works establish CF massive MIMO as a promising beyond-5G/6G radio-access architecture, but they generally focus on homogeneous traffic and nominal propagation conditions. As a result, the IoS live-event scenario, characterized by the coexistence of a massive best-effort DL crowd and guaranteed-rate UL contribution links under crowd-induced

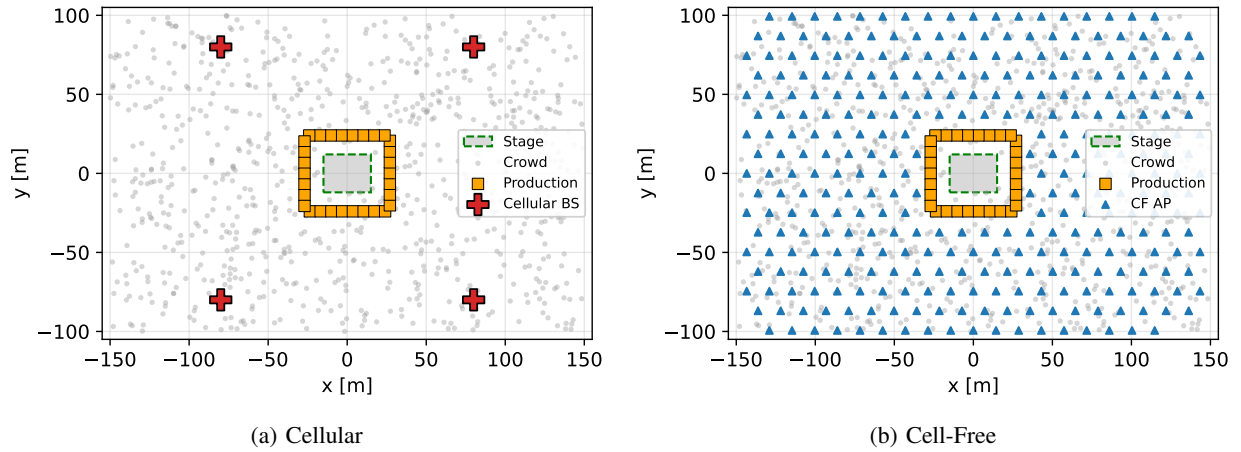


Fig. 1: Reference scenario for cellular and CF architectures.

blockage, remains largely unexplored.

Overall, existing works address IoS applications, professional media contribution, dense-venue connectivity, human-body blockage, and CF massive MIMO as largely separate problems. What remains insufficiently quantified is whether a distributed radio-access architecture can jointly sustain massive spectator traffic and guaranteed-rate wireless sound/media contribution links in a crowded live-event venue. This paper targets this intersection by comparing CF and cellular massive MIMO under equal radio-resource budgets and by explicitly accounting for the venue geometry and crowd-induced blockage.

III. SYSTEM MODEL

Let us consider a concert venue modeled as a stadium of size (l, w) with a central stage of size (l_s, w_s) surrounded by a crowd area (Fig. 1). The two architectures operate over the same area and serve the same set of devices under an equivalent radio-resource budget (respectively, Fig. 1a for traditional cellular architecture, and Fig. 1b for CF). In the CF case, N_{cf} APs, each with M_{cf} antennas, are distributed over the venue at height h_{ap} and connected to a CPU. In the cellular case, the same antenna and transmit-power budget is concentrated into N_{cell} BSs equipped with M_{cell} antennas of height h_{bs} placed around the venue.

The set of UEs $\mathcal{U}$ is partitioned into two classes. The first class, denoted by $\mathcal{C}$, represents the spectators attending the event. It consists of K_{sp} single-antenna users uniformly distributed over the venue, whose traffic is predominantly downlink-oriented and characterized by best-effort services, with only occasional UL transmissions. The second class, denoted by $\mathcal{P}$, represents the media-production devices responsible for generating the live audio-visual content. This set comprises $K_{pr} \ll K_{sp}$ production nodes, including cameras, microphones, immersive-audio capture systems, and mixing or telemetry equipment, deployed around the stage area. These two user classes generate fundamentally different traffic requirements, with the crowd primarily demanding downlink

connectivity and the production devices requiring reliable UL contribution links to support real-time media delivery.

The generic production device $p_k \in \mathcal{P}$, with $k = 1, \dots, K_{pr}$, requires a guaranteed UL bitrate $R_{req,k}$, whose value depends on the specific media application. Unlike spectators, whose traffic is predominantly best-effort, production devices must continuously deliver audio-visual contribution streams to the media-production infrastructure. Consequently, their communication requirements must be protected from fluctuations in public-network demand in order to guarantee stable contribution bitrates. Both the cellular and CF architectures employ an Orthogonal Frequency-Division Multiple Access (OFDMA)-based radio interface, where the available bandwidth B is divided into Physical Resource Blocks (PRBs). Resource isolation between the two user classes is achieved through network slicing. In particular, a fraction ϕ of the UL resources is reserved for media-production devices, while the remaining $(1 - \phi)$ portion is allocated to spectator traffic. This configuration guarantees service continuity for contribution links and reflects the operational practices commonly adopted in professional live-event production environments.

Both architectures operate in Time-Division Duplex (TDD) mode. Since spectator traffic is predominantly downlink-oriented whereas production traffic mainly consists of UL contribution streams, different TDD resource allocations can be assigned to the two slices. Although both user classes are explicitly modeled, the focus of this work is on the UL traffic generated by the production devices. In fact, the primary objective is to assess whether the considered radio-access architectures can reliably support the bitrate requirements of professional audio-visual contribution links while simultaneously serving a dense population of spectators.

Since the following propagation model applies identically to both spectators and production devices, a generic UE $k \in \mathcal{U}$ is considered. The wireless channel is characterized by the large-scale fading coefficient between radio node m and UE k , defined as

$$\beta_{mk}^{dB} = -(\text{PL}_{mk} + \xi_{mk} + L_{mk}^{\text{crowd}}), \quad (1)$$

where PL_{mk} denotes the path loss modeled as 3GPP Urban Micro Street-Canyon (UMi-SC) [13], $\xi_{mk} \sim \mathcal{N}(0, \sigma_{sf}^2)$ models the log-normal shadow fading component, and L_{mk}^{crowd} accounts for the additional attenuation introduced by the audience. The UMi-SC model is particularly suitable for dense in-venue deployments, characterized by relatively short communication distances and low-height access points. However, unlike conventional urban scenarios, live-event venues exhibit an additional propagation impairment caused by the presence of a dense crowd, whose bodies act as physical blockers.

To capture this effect, human bodies are modeled as a two-dimensional Poisson field with density ρ (expressed in persons/m²), determined by the crowd size and venue area ($\rho = K_{sp}/(l w)$). Each blocker is characterized by a representative height h_b and width w_b . For a link of length d_{mk} , the line-of-sight ray partially remains below h_b , thus being susceptible to blockage. Specifically, the ray remains blockable only over the near-UE segment of length

$$x_{mk}^{\text{blk}} = d_{mk} \frac{\max(h_b - h_{ue}, 0)}{h_m - h_{ue}} \quad (2)$$

with $0 \leq x_{mk}^{\text{blk}} \leq d_{mk}$, where $h_m \in \{h_{ap}, h_{bs}\}$ is the height of the serving radio node (an AP in the CF case, a BS in the cellular one), and $h_{ue} = 1.5$ m is the UE height.

The number of blockers intersecting this segment follows a Poisson distribution, $g_{mk} \sim \text{Poisson}(\rho w_b x_{mk}^{\text{blk}})$. Assuming that each blocker introduces an attenuation A_b , the overall crowd-induced loss is expressed as

$$L_{mk}^{\text{crowd}} = g_{mk} A_b. \quad (3)$$

This formulation naturally captures the dependence of crowd blockage through the ratio $d_{mk}/(h_m - h_{ue})$, with the blockable segment increasing with the link length and decreasing with the node elevation.

A. Cell-Free SINR

In the CF architecture all N_{cf} APs are connected to one or more CPUs and cooperate to serve the UEs coherently, without cell boundaries. Distributing M_{cf} -antenna nodes throughout the venue keeps the radio resources spatially close to both the crowd and the stage-side production devices, providing the macro-diversity that counteracts the crowd-induced blockage of Eq. (3). In fact, a device that is shadowed towards one AP is typically in line-of-sight towards another. Serving every UE from all N_{cf} APs is, however, neither scalable nor necessary, since only the nearby APs contribute an appreciable channel gain. We therefore adopt a user-centric formulation in which the CPU performs coordinated processing over the serving set $\mathcal{S}_k$ of each UE, defined as the $N_s \leq N_{cf}$ APs with the strongest large-scale fading,

$$\mathcal{S}_k = \left\{ \bar{m} : \beta_{\bar{m}k} \text{ among the } N_s \text{ largest } \beta_{mk} \right\}, \quad (4)$$

for $m = 1, \dots, N_{cf}$, and $|\mathcal{S}_k| = N_s$. Channel estimation relies on τ_p orthogonal UL pilots, which are reused across UEs. Considering the pilot power p_p and the set $\mathcal{U}_k$ of UEs sharing

the pilot of UE k , the MMSE estimation-quality coefficient is [5]

$$\gamma_{mk} = \frac{\tau_p p_p \beta_{mk}^2}{\tau_p p_p \sum_{j \in \mathcal{U}_k} \beta_{mj} + \sigma^2}, \quad (5)$$

where σ^2 is the noise power.

Pilots are assigned according to a large-scale fading-aware strategy that minimizes co-pilot overlap, thereby reducing pilot contamination within each slice. The spectator UE class employs scalable Maximum-Ratio (MR) combining over its serving set $\mathcal{S}_k$. The corresponding achievable Signal-to-Interference-plus-Noise Ratio (SINR) of UE k is given by

$$\text{SINR}_k^{\text{ul,cf}} = \frac{M_{cf} P_{ul} \left(\sum_{m \in \mathcal{S}_k} \gamma_{mk} \right)^2}{\sigma^2 \sum_{m \in \mathcal{S}_k} \gamma_{mk} + \sum_{j \in \mathcal{K}_k^{\text{sp}}, j \neq k} P_{ul} \sum_{m \in \mathcal{S}_k} \gamma_{mk} \beta_{mj}}, \quad (6)$$

where $\mathcal{K}_k^{\text{sp}}$ denotes the set of simultaneously active spectator UEs, and P_{ul} is the UL transmitted power. It is worth noticing that pilot contamination is captured through the estimation-quality coefficient γ_{mk} of Eq. (5), while the coherent interference term due to co-pilot users is neglected.

Conversely, the production UE class employs centralized Linear Minimum Mean Square Error (L-MMSE) combining at the CPU to suppress intra-slice interference. Denoting by $\hat{\mathbf{h}}_k$ the stacked channel estimates of UE k over its serving antennas, the corresponding achievable SINR is

$$\text{SINR}_k^{\text{ul,mmse}} = \frac{P_{ul} z_k}{1 - P_{ul} z_k}, \quad (7)$$

where $z_k = \hat{\mathbf{h}}_k^H \mathbf{A}^{-1} \hat{\mathbf{h}}_k$, and $\mathbf{A} = \sum_{j \in \mathcal{K}_k^{\text{pr}}} P_{ul} \hat{\mathbf{h}}_j \hat{\mathbf{h}}_j^H + \mathbf{C}$, with $\mathcal{K}_k^{\text{pr}}$ denoting the set of simultaneously active production UEs including UE k , and $\mathbf{C}$ accounting for the channel-estimation error and receiver noise. Given the limited size of the production fleet, these UEs are served by the full distributed AP set rather than by a restricted user-centric cluster. The clustering mechanism is therefore applied only to the massive spectator population, where it is required to preserve scalability. Since spectator and production traffic are mapped onto separate slices, the interference term for each UE includes only the simultaneously active users of the same class, so that the two slices are evaluated as logically isolated CF subsystems sharing the same physical infrastructure.

B. Cellular SINR

In the cellular baseline, each UE associates with the BS m providing the strongest large-scale fading coefficient β_{mk} . The same receiver-processing schemes adopted in the CF realization are employed, namely MR combining for spectator UEs and centralized L-MMSE combining for the production slice. For the spectator class, the SINR experienced by the generic UE k associated with BS m is

$$\text{SINR}_k^{\text{ul,cell}} = \frac{M_{\text{cell}} P_{ul} \gamma_{mk}}{\sigma^2 + \sum_{j \in \mathcal{K}_k^{\text{sp}}, j \neq k} P_{ul} \beta_{mj}}, \quad (8)$$

where, with a single serving node, the estimation quality affects only the useful signal term, while the interference depends on the large-scale fading coefficients. The production slice is instead processed according to Eq. (7), with the only difference that $\hat{\mathbf{h}}_k$ now collects the M_{cell} co-located antennas of the serving BS rather than those of the distributed array, and $\mathbf{A}$ is accordingly formed at BS level. Hence, the two architectures differ only in the spatial distribution of the antennas while employing the same receiver-processing strategy.

C. Link adaptation and throughput

In practical wireless systems, transceivers are affected by hardware impairments such as phase noise, power-amplifier non-linearities, and finite-resolution quantization. These effects introduce distortion components whose power scales with the received signal power and can be characterized through the transmitter/receiver Error Vector Magnitude (EVM). By modeling the aggregate hardware-induced distortion as an additional noise term with variance proportional to EVM^2 , we can express $\text{SINR}_{\text{max}} = \frac{1}{\text{EVM}^2}$. Thus, the effective SINR available at the demodulator can be expressed as

$$\text{SINR}_{\text{eff}} = \frac{\text{SINR}}{1 + \text{EVM}^2 \text{SINR}} = \frac{\text{SINR}}{1 + \text{SINR}/\text{SINR}_{\text{max}}}. \quad (9)$$

Unlike thermal noise, the distortion term increases with the received signal power, causing SINR_{eff} to saturate at the hardware-limited ceiling SINR_{max} . This bounds the achievable spectral efficiency and prevents physically unrealistic performance gains under highly favorable propagation conditions.

The effective SINR is mapped to the highest 5G NR Modulation and Coding Scheme (MCS) satisfying the target Block Error Rate (BLER) requirement [14]. We consider the 256-QAM table for the production slice, whereas the spectator slice employs the 64-QAM table. Each MCS is associated with a spectral efficiency $\varepsilon_k^{\text{mcs}}$, while the residual packet errors are captured through the corresponding BLER value, BLER_k . The resulting goodput spectral efficiency is therefore expressed as $\varepsilon_k = \varepsilon_k^{\text{mcs}} (1 - \text{BLER}_k)$, which represents the average amount of successfully delivered information per unit bandwidth. Since the UL resources are partitioned through slice isolation, production devices and spectators are served over distinct bandwidth portions, denoted by $B_{s,\text{pr}} = \phi B$, and $B_{s,\text{sp}} = (1 - \phi)B$. In high-density live-event scenarios, the achievable capacity is primarily determined by the spatial multiplexing capability of the radio-access architecture, whose hard ceiling is set by the spatial degrees of freedom available at the total deployed antenna.

Accordingly, the K^{act} active users of each slice are partitioned into co-scheduled groups of at most K_g simultaneous streams, which share the slice bandwidth in time. The SINR of Eq. (6)–(8) is evaluated within each group, while the resulting time-sharing factor is $\mu = 1/\lceil K^{\text{act}}/K_g \rceil$. The group size K_g depends on the adopted receiver-processing scheme. While interference-suppressing receivers can efficiently support a larger number of simultaneous streams, with MR combining

the achievable SINR rapidly degrades as the multiplexing order increases. Therefore, K_g is selected independently for the two receiver architectures. The throughput achieved by each UE is then

$$R_k = \varepsilon_k B_s f_{\text{dir}} \mu \left(1 - \frac{\tau_p}{\tau_c} \right), \quad (10)$$

where f_{dir} denotes the TDD resource fraction allocated to the considered transmission direction and $B_s \in \{B_{s,\text{pr}}, B_{s,\text{sp}}\}$, while the last factor accounts for the pilot overhead within the coherence block τ_c , explicitly considering the resources reserved for the pilots of the production slice. A production device is considered successfully served whenever its achieved throughput satisfies the application requirement, i.e., $R_k \geq R_k^{\text{req}}$.

IV. SIMULATION SETUP

To compare the performance of the CF and cellular architectures under equivalent radio-resource budgets, we develop an ad hoc system-level simulator that reproduces the live-event scenario described in Section III. The simulator generates the venue geometry, places spectators and production devices according to the considered deployment assumptions, computes the large-scale fading coefficients according to Eq. (1), evaluates the UL SINR experienced by both user classes, and applies the 5G NR link-adaptation and throughput models described in Section III. In computing the large-scale fading, the per-link crowd excess loss of Eq. (3) is saturated at a maximum value $L_{\text{max}} = 20$ dB, i.e. $L_{mk}^{\text{crowd}} = \min(g_{mk} A_b, L_{\text{max}})$, reflecting the finite absorption of a dense crowd, beyond which the link is sustained by diffraction and scattering around the bodies. Pilot allocation satisfies the coherence-block constraint $\tau_p^{\text{pr}} + \tau_p^{\text{sp}} \leq \tau_c$, with $\tau_p^{\text{pr}} = \min(K_{\text{pr}}, \tau_p^{\text{max}})$. For $K_{\text{pr}} > \tau_p^{\text{max}}$, the production slice also reuses pilots, and the resulting pilot contamination is accounted for in (5) through $\mathcal{U}_k$. For a fair comparison, the CF and cellular deployments are configured with the same total antenna budget ($N_{\text{cf}} M_{\text{cf}} = N_{\text{cell}} M_{\text{cell}}$).

The venue hosts a realistic attendance of $K_{\text{sp}} = 120000$ spectators, which determines the crowd-blockage density ρ of Eq. (3). Since this density is a statistical (Poisson) quantity, it does not require placing every attendee individually. Thus, we drive the blockage with the full attendance while explicitly simulating a representative sample of 7500 spectators as the simultaneously-active users, which also keeps the computational cost of the Monte Carlo campaign tractable. As the production contribution is carried on an isolated slice, its results are independent of the simulated crowd size. Table I summarizes the simulation parameters. The considered production devices and their corresponding guaranteed UL bitrates are selected to reflect representative wireless media-contribution applications, including compressed UHD/8K video streams, multichannel and immersive-audio capture systems, wireless microphones, and telemetry devices. Finally, the effective-SINR ceiling is set to $\text{SINR}_{\text{max}} = 30$ dB, corresponding to an EVM of approximately 3.2%, which is representative of a high-quality RF front-end supporting modulation orders up to 256-QAM.

V. SIMULATION RESULTS

The performance comparison between the CF and cellular architectures is first evaluated in terms of UL signal quality. Fig. 2 shows the SINR distributions experienced by both production devices and spectators for $K_{\text{pr}} = 30$ production UEs. For the production class, both architectures operate close to the maximum usable SINR. The CF deployment reaches an average UL SINR of approximately 29.9 dB, with the distribution tightly concentrated around the hardware limit imposed by Eq. (9). The cellular baseline achieves a comparable average of about 27.6 dB, but exhibits a substantially wider distribution. These results indicate that both architectures provide sufficient signal quality to support the contribution traffic, whereas the main advantage of the distributed deployment lies in the spatial uniformity of the received SINR. By exploiting multiple geographically distributed APs, CF effectively mitigates location-dependent performance variations, while the cellular architecture remains more sensitive to the position of the serving BS.

The difference between the two architectures becomes considerably more pronounced for the spectator class, where thousands of simultaneously active users compete for the same radio resources. In this case, the CF architecture maintains the UL SINR distribution around an average of 3.1 dB, whereas the cellular baseline decreases to -8.0 dB and exhibits a substantially heavier lower tail, reaching up to -44 dB at the 5th percentile compared with -1.4 dB for CF. These results indicate that the distributed deployment provides a substantially more uniform service across the spectator popula-

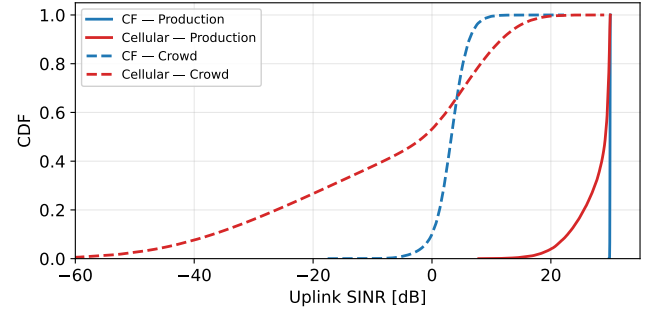


Fig. 2: UL SINR CDF comparison between CF and cellular for both production and spectator UE classes.

tion, effectively mitigating the severe degradation experienced by cell-edge users in the cellular architecture. It is worth noting that this comparison is performed under identical pilot resources and aggregate UL transmit power, ensuring that the observed differences are exclusively determined by the spatial distribution of the radio infrastructure rather than by more favorable channel estimation or power budgets.

Consequently, the SINR improvement directly translates into the UL throughput CDFs reported in Fig. 3 for both CF and cellular architectures with $K_{\text{pr}} = 30$. The CF architecture provides approximately 124 Mbps to all production devices, fully satisfying the bitrate requirements of the considered applications. The cellular deployment achieves a comparable average throughput of about 120 Mbps and the same median value, but exhibits a much broader throughput distribution, with values starting from 37.3 Mbps with an average around 119.8 Mbps. Although both architectures provide similar average throughput, the CF deployment offers a significantly more uniform service across the production fleet, ensuring the throughput reliability required by guaranteed-rate multimedia contribution services.

Fig. 4 compares the average achieved throughput against the target bitrate required by each application considering $K_{\text{pr}} = 30$ production UEs. With the production slice allocated 80% of the available UL resources and processed through interference-suppressing L-MMSE combining, both architectures satisfy the requirements of all application classes, including the 8K cameras requiring 80 Mbps. The CF deployment achieves

TABLE I: Simulation parameters

Parameter	Cell-Free	Cellular
Nodes	$N_{\text{cf}} = 320$ APs	$N_{\text{cell}} = 4$ BSs
Antennas per node	$M_{\text{cf}} = 2$	$M_{\text{cell}} = 160$
Node height	$h_{\text{ap}} = 10$ m	$h_{\text{bs}} = 25$ m
Combining	MR / L-MMSE	MR / L-MMSE
Serving set, N_s	32 APs (sp.), 320 (pr.)	Strongest β
Co-sched. streams, K_g (sp. / pr.)	256 / 640	128 / 640
Shared parameters		
Total DL power	320 W	
Venue, (l, w)	(300, 200) m	
Stage, (l_s, w_s)	(30, 24) m	
Spectators, K_{sp}	120000 ($\approx 2/\text{m}^2$)	
Production, K_{pr}	[30, 720]	
Required data rate, $R_{\text{req},k}$	80 Mbps (8K Cam.), 25 Mbps (UHD Cam.), 18 Mbps (Imm. Audio), 8 Mbps (Multich. Audio), 2 Mbps (Wireless Mic.), 1 Mbps (Telemetry)	
Carrier fr. / Bw., f_c / B	3.5 GHz / 100 MHz (273 RB, 30 kHz SCS)	
TDD frame, f_{dir}	20 slots (14 DL / 5 UL / 1 gd.): UL 0.25, DL 0.70	
UL transmit power, P_{ul}	23 dBm	
Pilots, τ_p / τ_c	Orthogonal (prod.), Reused ≤ 20 (crowd) / 200	
Noise Figure / SINR Cap	7 dB / 30 dB (EVM)	
Production slice, ϕ	80% of UL band	
Path loss	3GPP UMi-Street Canyon [13]	
Crowd specifications	$h_b = 1.7$ m, $w_b = 0.5$ m	
Crowd blockage	$\rho \approx 2/\text{m}^2$, $A_b = 3$ dB, $L_{\text{max}} = 20$ dB	
Multiplexing / Target BLER	Spatial multiplexing / 10^{-1}	
Realizations	1000	

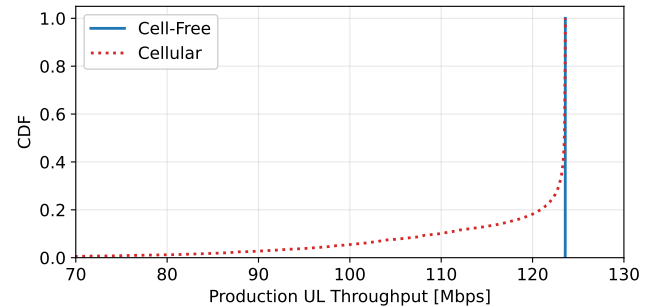


Fig. 3: Production UL throughput CDF.

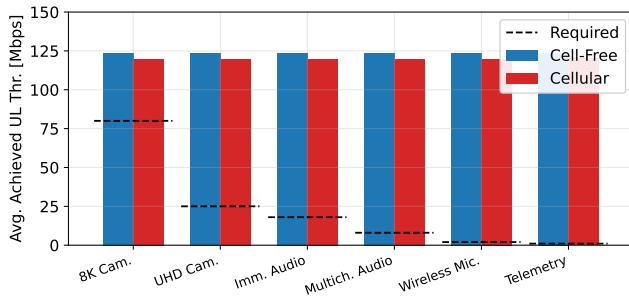


Fig. 4: Production UL average achieved rate vs required rate per device type.

a 100% satisfaction rate across the entire production fleet, whereas the cellular baseline reaches an overall satisfaction of 99.9%. The only residual degradation is observed for the 8K cameras, whose bitrate requirement is sufficiently close to the achievable throughput for the lower tail of the throughput distribution to produce a limited number of unsatisfied devices.

Finally, Fig. 5 investigates the scalability of the two architectures as the number of production devices increases while maintaining the same total number of antennas and transmit-power budget. Both architectures achieve full satisfaction for the reference production fleet and provide comparable performance up to approximately 60 devices. Beyond this operating point, however, the performance progressively diverges. At 240 production devices, the CF architecture still satisfies 93% of the links, compared with 72% for the cellular baseline, while at 720 devices the satisfaction rate decreases to 52% and 39%, respectively. The average UL throughput follows a similar trend, decreasing from approximately 124 to 10 Mbps for CF and from 120 to 10 Mbps for the cellular deployment.

VI. CONCLUSION

This work investigated the adoption of CF massive MIMO as a radio-access architecture for large-scale IoS and live media-production events, where professional audio-visual production devices coexist with a dense population of spectators. To this end, a system-level framework jointly modeling crowd-induced blockage, network slicing, spatial multiplexing, and heterogeneous traffic requirements was developed. The results

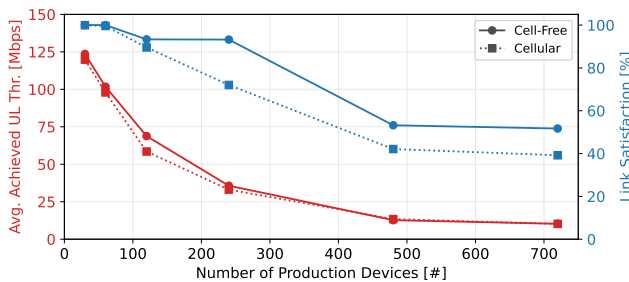


Fig. 5: Average achieved UL throughput vs the number of production devices and link satisfaction.

showed that, although both CF and conventional cellular deployments sustain the guaranteed-rate uplink under the reference operating conditions, the distributed architecture provides substantially more uniform service across both production and spectator devices, while preserving a higher satisfaction rate as the production fleet scales, making CF a promising radio-access solution for future multimedia environments.

Several research directions remain open. From a validation perspective, the proposed system-level study will be complemented by measurement campaigns and experimental testbeds in real live-event venues to calibrate the crowd-blockage model and assess the impact of practical implementation aspects, including hardware impairments and synchronization inaccuracies. From a modeling perspective, future work will investigate indoor arenas, user mobility and motions, dynamic crowd evolution. Additional research efforts will address pilot-assignment strategies specifically tailored to media-contribution traffic, as well as the joint optimization of cloud-edge processing and radio-resource management.

REFERENCES

- [1] S. Elhoushy, M. Ibrahim, and W. Hamouda, "Cell-Free Massive MIMO: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 492–523, 2022.
- [2] EBU, "5G for the Distribution of Audiovisual Media Content and Services," European Broadcasting Union, Technical Report (TR) 054, 5 2020.
- [3] L. Turchet, M. Lagrange, C. Rottondi, G. Fazekas, N. Peters, J. Østergaard, F. Font, T. Bäckström, and C. Fischione, "The Internet of Sounds: Convergent Trends, Insights, and Future Directions," *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11 264–11 292, 2023.
- [4] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-Free Massive MIMO Versus Small Cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.
- [5] E. Björnson and L. Sanguinetti, "Scalable Cell-Free Massive MIMO Systems," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4247–4261, 2020.
- [6] G. Interdonato, E. Björnson, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Ubiquitous Cell-Free Massive MIMO Communications," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 197, 2019.
- [7] L. Turchet, C. Fischione, G. Essl, D. Keller, and M. Barthet, "Internet of Musical Things: Vision and Challenges," *IEEE Access*, vol. 6, pp. 61 994–62 017, 2018.
- [8] C. Rottondi, C. Chafe, C. Allocchio, and A. Sarti, "An Overview on Networked Music Performance Technologies," *IEEE Access*, vol. 4, pp. 8823–8843, 2016.
- [9] Audio Engineering Society, "AES67-2023: AES Standard for Audio Applications of Networks – High-Performance Streaming Audio-Over-IP Interoperability," AES Standard, 12 2023.
- [10] 3GPP, "Service requirements for Video, Imaging and Audio for Professional Applications (VIAPA) (Release 19)," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 22.263, 10 2025, Version 19.0.0.
- [11] M. Gapeyenko, A. Samuylov, M. Gerasimenko, D. Moltchanov, S. Singh, E. Aryafar, S.-p. Yeh, N. Himayat, S. Andreev, and Y. Koucheryavy, "Analysis of Human-Body Blockage in Urban Millimeter-Wave Cellular Communications," in *2016 IEEE International Conference on Communications (ICC)*, 2016, pp. 1–7.
- [12] S. Buzzi and C. D'Andrea, "Cell-Free Massive MIMO: User-Centric Approach," *IEEE Wireless Communications Letters*, vol. 6, no. 6, pp. 706–709, 2017.
- [13] 3GPP, "Study on Channel Model for Frequencies from 0.5 to 100 GHz (Release 19)," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 38.901, 06 2026, Version 19.4.0.
- [14] —, "NR; Physical Layer Procedures for Data (Release 19)," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.214, 06 2026, Version 19.4.0.