Ubiquitous Cell-Free Massive MIMO Communications

Giovanni Interdonato, Emil Björnson, Hien Quoc Ngo, Pål Frenger and Erik G. Larsson

Abstract

Since the first cellular networks were trialled in the 1970s, we have witnessed an incredible wireless revolution. From 1G to 4G, the massive traffic growth has been managed by a combination of wider bandwidths, refined radio interfaces, and network densification, namely increasing the number of antennas per site \[1\]. Due its cost-efficiency, the latter has contributed the most. Massive MIMO (multiple-input multiple-output) is a key 5G technology that uses massive antenna arrays to provide a very high beamforming gain and spatially multiplexing of users, and hence, increases the spectral and energy efficiency (see \[2\] and references herein). It constitutes a centralized solution to densify a network, and its performance is limited by the inter-cell interference inherent in its cell-centric design. Conversely, ubiquitous cell-free Massive MIMO \[3\] refers to a distributed Massive MIMO system implementing coherent user-centric transmission to overcome the inter-cell interference limitation in cellular networks and provide additional macro-diversity. These features, combined with the system scalability inherent in the Massive MIMO design, distinguishes ubiquitous cell-free Massive MIMO from prior coordinated distributed wireless systems. In this article, we investigate the enormous potential of this promising technology while addressing practical deployment issues to deal with the increased back/front-hauling overhead deriving from the signal co-processing.
I. INTRODUCTION

One of the primary ways to provide high per-user data rates--requirement for the creation of a 5G network--is through network densification, namely increasing the number of antennas per site and deploying smaller and smaller cells [1]. A communication technology that involves base stations (BSs) with very large number of transmitting/receiving antennas is Massive MIMO [2], where MIMO stands for multiple-input multiple-output. This key 5G technology leverages aggressive spatial multiplexing. In the uplink, all the users transmit data to the BS in the same time-frequency resources. The BS exploits the massive number of channel observations to apply linear receive combining, which discriminates the desired signal from the interfering signals. In the downlink, the users are coherently served by all the antennas, in the same time-frequency resources but separated in the spatial domain by receiving very directive signals. By supporting such a highly spatially-focused transmission (precoding), massive MIMO provides higher spectral efficiency, and reduces the inter-cell interference compared to existing mobile systems.

The inter-cell interference is however becoming the major bottleneck as we densify the networks. It cannot be removed as long as we rely on a network-centric (cell-centric) implementation, since the inter-cell interference is inherent to the cellular paradigm. In a conventional cellular network, each user equipment (UE) is connected to the access point (AP) in only one of the many cells (except during handover). At a given time instance, the APs have different numbers of active UEs, causing inter-cell interference (Fig. 1a).

Cellular networks are suboptimal from a channel capacity viewpoint because higher spectral efficiency (SE) (bit/s/Hz/user) can be achieved by co-processing each signal at multiple APs [4]. The signal co-processing concept is present in [5], network MIMO [6], coordinated multipoint with joint transmission (CoMP-JT) [7], and multi-cell MIMO cooperative networks [8]. It is conventionally implemented in a network-centric fashion, by dividing the APs into disjoint clusters as in Fig. 1b. The APs in a cluster transmit jointly to the UEs residing in their joint coverage area, thus it is equivalent to deploying a conventional cellular network with distributed antennas in each cell. Despite the great theoretical gains, the 3GPP LTE standardization of CoMP-JT has not achieved much practical gains [7].

Conversely, when the co-processing is implemented in a user-centric fashion, each user is served by its closest APs (user-specific cluster), while all the APs that affect the user take its interference into consideration, as illustrated in Fig. 1c. This approach eliminates the cell
(a) A conventional cellular network where each UE is connected to only one AP.

(b) A conventional network-centric implementation of CoMP-JT, where the APs in a cluster cooperate to serve the UEs residing in their joint coverage area.

(c) A user-centric implementation of CoMP-JT, where each UE communicates with its closest APs.

(d) A “cell-free” Massive MIMO network is a way to implement a user-centric network.

Fig. 1: With the same collection of APs, we can create many different types of communication networks.
boundaries resulting in no inter-cell interference, hence the terminology “cell-free”.

Ubiquitous cell-free Massive MIMO enhances the conventional CoMP-JT by implementing the user-centric approach and leveraging the benefits of using Massive MIMO, i.e., high spectral efficiency, system scalability, and close-to-optimal linear processing. In such a system, the APs are connected via front-haul connections to central processing units (CPUs), which are responsible for the coordination. The CPUs are interconnected by back-haul (Fig. 1d).

Note that, it is convenient to have AP-specific synchronization and reference signals, which can be viewed as cell identifiers, also in cell-free networks. Hence, “cell-free” refers to the lack of cell boundaries in the data transmission.

II. SYSTEM OPERATION AND RESOURCE ALLOCATION

To give a first sense of the paradigm shift that cell-free Massive MIMO constitutes, Fig. 2 shows the user performance at different locations in an area with nine APs: Fig. 2a shows that the SEs in a cellular network is poor at the cell edges due to strong inter-cell interference, while Fig. 2b shows that a cell-free network can avoid interference by co-processing over the APs and provide more uniform performance among the users. The SE is only limited by signal propagation losses.

A. Ubiquitous Cell-Free Massive MIMO: The Scalable Way to Implement CoMP-JT

The first challenge in implementing a cell-free Massive MIMO network is to obtain sufficiently accurate channel state information (CSI) so that the APs can simultaneously transmit (receive) signals to (from) all UEs and cancel interference in the spatial domain. The conventional approach of sending DL pilots and letting the UEs feed back channel estimates is unscalable since the feedback load is proportional to the number of APs. To circumvent this issue, we note that each AP only requires local CSI to perform its tasks [9], which refers to the channel between the AP and to each of the UEs. Local CSI is conveniently acquired in time-division duplex (TDD) operation since, when a UE sends a pilot, each AP can simultaneously estimate its channel to the UE. Hence, the overhead is independent of the number of APs. By exploiting channel reciprocity, the UL channel estimates can be also utilized as DL channel estimates, as in cellular Massive MIMO [2]. Just like Massive MIMO is the scalable way to implement multi-user MIMO [2], ubiquitous cell-free Massive MIMO is the scalable way to implement CoMP-JT [3], [10], [11]. In cell-free networks there are L of geographically distributed APs that jointly serve a
Fig. 2: SE achieved by UEs at different locations in an area covered by nine APs that are deployed on a regular grid. Note that 8 bit/s/Hz was selected as the maximal SE, which corresponds to uncoded 256-QAM.

relatively smaller number $K$ of users: $L \gg K$. Cell-free Massive MIMO can provide ten-fold improvements in 95%-likely SE for the UEs over a corresponding cellular network with small cells [3], [10]. There are two key properties that explains this result.

The first property is the increased macro-diversity. Fig. 3a illustrates this with single-antenna APs deployed on a square-grid with varying inter-site distance (ISD): 5, and 100 m. The figure shows the cumulative distribution function (CDF) of the channel gain for a UE at a random position with channel vector $\mathbf{h} = [h_1 \ldots h_L]^T \in \mathbb{C}^L$, where $h_l$ is the channel from the $l$th AP. The channel gain is $\|\mathbf{h}\|^2$ in cell-free Massive MIMO and $\max_l |h_l|^2$ in a cellular network. With
Fig. 3: Evaluation of (a) macro-diversity, and (b) favorable propagation in cell-free Massive MIMO. The simulation setup considers 2500 single-antenna APs deployed on a square-grid with wrap-around and varying ISD. We consider independent Rayleigh small-scale fading and three-slope path-loss model from [3].
a large ISD, the UEs with the best channel conditions have almost identical channel gains in both cases, but the most unfortunate UEs gains 5 dB from cell-free processing. With a small ISD of 5 m, which is reasonable for connected factory applications, all UEs obtain 5-20 dB higher channel gain by the cell-free network.

The second property is favorable propagation, which means that the channel vectors $h_1, h_2$ of any pair of UEs are nearly orthogonal, leading to little inter-user interference. The level of orthogonality can be measured by the squared inner product $|h_1^H h_2|^2 / (\|h_1\|^2 \|h_2\|^2)$. A smaller value represents greater orthogonality. In a cellular network with single-antenna APs, $h_1$ and $h_2$ are scalars and thus the measure is one. Favorable propagation will, however, appear in cell-free Massive MIMO where $h_1, h_2 \in \mathbb{C}^L$, since the combination of small-scale and large-scale fading makes the large-dimensional channel vectors pairwise nearly orthogonal [11]. This is illustrated in Fig. 3b which shows the CDF of the orthogonality measure for two randomly located UEs. The inner product is very small for all the considered ISDs.

B. TDD Protocol

The TDD protocol recommended for cell-free Massive MIMO is illustrated in Fig. 4. Each AP estimates the uplink (UL) channel from each UE by measurements on UL pilots. By virtue of reciprocity, these estimates are also valid for the DL channels. Hence, the pilot resource requirement is independent of the number of AP antennas and no UL feedback is needed.

After applying precoding, each UE sees an effective scalar channel. The UE needs to estimate the gain of this channel to decode its data. Note that in cellular Massive MIMO, owing to channel hardening, the UE may rely on knowledge of the average channel gain for decoding [2]. In cell-free Massive MIMO, in contrast, there is less hardening and DL effective gain estimation is desirable at the terminal [11], [12]. This estimate can either be obtained from DL pilots sent by the AP during a DL training phase [12] (Fig. 4a) or, potentially, blindly from the DL data transmission if there are no DL pilots (Fig. 4b).

Fig. 4 shows two possible TDD frame configurations, with and without DL pilot transmission. The configuration including the pilot-based DL training, depicted in Fig. 4a, consists of four phases: (i) UL training; (ii) UL data transmission; (iii) Pilot-based DL training; and (iv) DL data transmission. Fig. 4b illustrates the TDD frame without DL pilot transmission. This implies that, for data decoding, the UEs either rely on channel hardening or blindly estimate the DL channel from the data.
The channel coherence interval is defined as the time-frequency interval during which the channel can be approximately considered as static. It is determined by the propagation environment, UE mobility, and carrier frequency [2]. The TDD frame should be equal or shorter than the smallest coherence time among the active UEs. For simplicity, we herein assume it is equal.

Let \( \tau = T_c B_c \) the length of TDD frame in symbols, where \( B_c \) is the coherence bandwidth and \( T_c \) indicates the coherence time. It is partitioned as \( \tau = \tau_{u,p} + \tau_{u,d} + \tau_{d,p} + \tau_{d,d} \), where \( \tau_{u,p}, \tau_{u,d}, \tau_{d,p} \) and \( \tau_{d,d} \) denote the total number of symbols per frame spent on transmission of UL pilots, UL data, DL pilots and DL data, respectively. Importantly, \( \tau \) as well as \( \tau_{p}, \tau_{d,d}, \tau_{u,d} \) can be adjusted over time to accommodate the coherence interval variation and the traffic load change. However, such frame reconfiguration should occur slowly to limit the amount of control signaling required by the resource re-allocation.

The maximum number of mutually orthogonal pilots is upper-bounded by \( \tau \). Hence, allocating orthogonal pilots is physically impossible in networks with \( K \geq \tau \), and non-orthogonal pilots are necessary. UEs that send non-orthogonal pilots cause mutual interference that make the respective channel estimates correlated, a phenomenon known as pilot contamination. This results in a degradation of the channel estimation quality and in coherent interference.

C. Power Control

Power control is important to handle the near-far effect, and protect UEs from strong interference. The power control can be governed by the CPU, which tells the APs and UEs which power-control coefficients to use. By using closed-form capacity bounds that only depend on
the large-scale fading, the power control can be well optimized and infrequently updated, e.g.,
a few times per second.

When maximum-ratio (MR) precoding is used, at AP $l$, the symbol intended for UE $k$, $q_k$, is
first weighted by $\hat{g}_{lk}^*$ and $\sqrt{\rho_{lk}}$, where $\hat{g}_{lk}$ is the estimate of the channel from AP $l$ to UE $k$ and
$\rho_{lk}$ is the power-control coefficient. The weighted symbols of all $K$ UEs will be then combined
and transmitted to the UEs. In the UL, at UE $k$, the corresponding symbol $q_k$ is weighted by a
power-control coefficient $\sqrt{\rho_k}$ before transmission to the APs.

In general, the power-control coefficients should be selected to maximize a given performance
objective. This objective may, for example, be the max-min rate or sum rate:

- Max-Min Fairness Power Control: The goal of this power-control policy is to deliver the
same rate to all UEs and maximizing that rate. In a large network, some UEs may have very
bad channels to all APs, thus it is necessary to drop them from service before applying this
policy, otherwise the service will be bad for everyone. As in cellular Massive MIMO, the
max-min fairness power-control coefficients can be obtained efficiently by means of linear
and second-order cone optimization [3, Section IV-B].
- Power Control with User Prioritization: The rate requirements are typically different among
the UEs, which can be taken into account in the power-control policy. For instance, UEs
that use real-time services or have more expensive subscriptions have higher priority. The
max-min fairness power control can be extended to consider weighted rates, where the
individual weights represent the priorities. Minimum rate constraints can be also included.
- Power Control with AP Selection: Due to the path-loss, APs far away from a given UE will
modestly contribute to its performance. AP selection is implemented by setting non-zero
power-control coefficients to the APs designed to serve that UE.

D. Pilot Assignment

To limit pilot contamination, efficient pilot assignment is important. Pilot assignment is deter-
mined at the CPU, and communicated to all the APs which forward it to the UEs. This message
would map the UE identifier to the pilot index. This UE-to-pilot mapping can be transmitted
either in the broadcast control channel within the system information acquisition process or in
the random access channel during the random access procedure. Pilot assignment can be done
in several ways:
• Random pilot assignment: Each UE is randomly assigned one of the $\tau_p$ mutually orthogonal pilots. This method requires no coordination, but there is a substantial probability that closely located UEs use the same pilot, leading to bad performance.

• Brute-force optimal assignment: A search over all possible pilot sequences can be performed to maximize a utility of choice, such as the max-min rate or sum rate. This method is optimal but its complexity grows exponentially with $K$.

• Greedy pilot assignment: The $K$ UEs are first assigned pilot sequences at random. Then this assignment is iteratively improved by performing small changes that increase the utility. To achieve good network performance, pilot assignment and power control can be performed jointly.

III. PRACTICAL DEPLOYMENT ISSUES

The cost and complexity of deployment, limited capacity of back/front-haul connections, and network synchronization are three major issues that need be solved in a practical deployment.

A. Radio Stripes System

The cabling and internal communication between APs is challenging in practical cell-free massive MIMO deployments. An appropriate, cost-efficient architecture is the radio stripe system (patent WO/2018/103897), presented next.

In a radio stripe system, the antennas and the associated antenna processing units (APUs) are serially located inside the same cable, which also provides synchronization, data transfer, and power supply via a shared bus; see Fig. 5a. More specifically, the actual APs consist of antenna elements and circuit-mounted chips (including power amplifiers, phase shifters, filters, modulators, A/D and D/A converters) inside the protective casing of a cable or a stripe. Each radio stripe is then connected to one or multiple CPUs. Since the total number of distributed antennas is assumed to be large, the transmit power of each antenna can be very low, resulting in low heat-dissipation, small volume and weight, and low cost. Small low-gain antennas are used. For example, if the carrier frequency is 5.2 GHz then the antenna size is 2.8 cm, thus, the antennas and processing hardware can be easily fitted in a cable/stripe.

The receive/transmit processing of an antenna is performed right next to itself. On the transmitter side, each APU receives up to $K$ streams of input data (e.g., one stream per UE, one UE with $K$ streams, or some other UE-stream allocation) from the previous APU via the
(a) Example of a radio stripe system.

(b) Radio stripes, here illustrated in white, enable invisible installation in existing construction elements.

**Fig. 5:** The radio stripe system design is shown in (a), while different deployment concepts are exemplified in (b). Each radio stripe sends/receives data to/from one or multiple CPUs through a shared bus (or internal connector), which also provides synchronization and power supply to each APU.
shared bus. In each antenna, the input data streams are scaled with the pre-calculated precoding vector and the sum-signal is transmitted over the radio channel to the receiver(s). By exploiting channel reciprocity, the precoding vector may be a function of the estimated uplink channels. For example, if the conjugate of the estimated uplink channel is used, MR precoding is obtained. This precoding requires no CSI sharing between the antennas.

On the receiver side, the received radio signal is multiplied with the combining vector previously calculated in the uplink pilot phase. The output gives $K$ data streams. The processed streams are then combined with the data streams received from the shared bus and sent again on the shared bus to the next APU. More specifically, the $m$th APU sums its received data streams to the input streams from APU $m-1$ consisting of combined signals from APUs $1, \ldots, m-1$, for one or more UEs. This cumulative signal is then outputted to APU $m+1$. The combination of signals is a simple per-stream addition operation.

The radio stripe system facilitates a flexible and cheap cell-free Massive MIMO deployment. Cheapness comes from many aspects: $(i)$ deployment does not require highly qualified personnel. Theoretically, a radio stripe needs only one (plug and play) connection either to the front-haul network or directly to the CPU; $(ii)$ a conventional distributed massive MIMO deployment requires a star topology, i.e., a separate cable between each APs and a CPU, which may be economically infeasible. Conversely, radio stripe installation complexity is unaffected by the number of antenna elements, thanks to its compute-and-forward architecture. Hence, cabling becomes much cheaper; $(iii)$ maintenance costs are cut down as a radio stripe system offers increased robustness and resilience: highly distributed functionality offer limited overall impact on the network when few stripes being defected; $(iv)$ low heat-dissipation makes cooling systems simpler and cheaper. While cellular APs are bulky, radio stripes enable invisible installation in existing construction elements as exemplified in Fig. 5b. Moreover, a radio stripe deployment may integrate for example temperature sensors, microphones/speakers, or vibration sensors, and provide additional features such as fire alarms, burglar alarms, earthquake warning, indoor positioning, and climate monitoring and control.

B. Front-haul and Back-haul Capacity

While there is no need to share CSI between antennas, the CPUs must provide each APU with the data streams. The data is delivered from the core network via the back-haul and then forwarded to the APU over the front-haul; see Fig. 5a. Similarly, the CPU receives the cumulative
signals from its the radio stripes over the front-haul and decodes them. The data will then be delivered to the core network over the back-haul.

The required front-haul capacity of a radio stripe is proportional to the number of simultaneous data streams that it supports at maximum network load. The required back-haul capacity of a CPU corresponds to the sum rate of the data streams that its radio stripes will transmit/receive at maximum network load. The way to limit these capacity requirements is to constrain the number of UEs that can be served per AP (e.g., radio stripe) and CPU. To avoid creating cell boundaries, a user-centric perspective must be used when selecting which subset of APs that serve a particular UE [3], [13], as illustrated in Fig. 1c.

Suppose a UE is alone in the network and all APs transmit to it with full power. Since the path-loss decays rapidly with the propagation distance, 95% of the received power will originate from a subset of the APs, called the 95%-subset. When the ISD is large, as in a conventional cellular network, the 95%-subset might only contain a handful of APs. As the ISD reduces (i.e., the number of APs per km² grows), the 95%-subset is larger. This property can be used to limit the back-haul signaling. For example, it is shown in [3] that only 10-20% of the APs in the 1 km² area surrounding a UE belongs to the 95%-subset.

C. Synchronization

To serve a UE by coherent joint transmission from multiple APs, the network infrastructure needs to be synchronized. The network might have an absolute time (phase) reference, but the APs are unsynchronized. This means that, effectively, the transmitter and receiver circuits of each AP have their own time references. The difference in time reference between the transmitter and receiver in a given AP represents the reciprocity calibration error. The difference in, say, transmitter time reference, between any pair of APs represents the synchronization error between these two APs. To limit the reciprocity and synchronization errors, a synchronization process need to be applied at regular intervals.

Suppose the transmitter of AP_1 has a clock bias of \( t_i \) (i.e., its local time reference clock shows zero at absolute time \( t_i \)) and the receiver has a clock bias of \( r_i \) (i.e., its clock shows zero at absolute time \( r_i \)). A simple synchronization protocol may then work as follows:

1) At local time zero (absolute time \( t_1 \)), AP_1 transmits a known pulse. AP_2 receives this pulse at time \( t_1 - r_2 \), according to its clock, and timestamps it with \( \delta_{12} = t_1 - r_2 \). Similarly, AP_3 timestamps the pulse with \( \delta_{13} = t_1 - r_3 \).
2) At its local time zero, AP$_2$ transmits a known pulse. AP$_1$ timestamps the received pulse with its local reception time $\delta_{21} = t_2 - r_1$. AP$_3$ timestamps it with $\delta_{23} = t_2 - r_3$.

3) Finally, at its local time zero, AP$_3$ transmits a known pulse. AP$_1$ timestamps this received pulse with $\delta_{31} = t_3 - r_1$. AP$_2$ timestamps it with $\delta_{32} = t_3 - r_2$.

The quantities $\delta_{ij}$ are known from the measurements, but $t_1, r_1, t_2, r_2, t_3, r_3$ cannot be obtained from $\delta_{ij}$ since the corresponding linear equation system is singular. However, the reciprocity and synchronization errors are easily recovered:

$$t_1 - r_1 = \delta_{12} + \delta_{31} - \delta_{32}, t_2 - r_2 = \delta_{21} + \delta_{32} - \delta_{31}, t_3 - r_3 = \delta_{31} + \delta_{23} - \delta_{21}, t_1 - t_2 = \delta_{13} - \delta_{23}, t_1 - t_3 = \delta_{12} - \delta_{32}, t_2 - t_3 = \delta_{21} - \delta_{31}.$$ 

This enables synchronization between the three APs.

This synchronization method can be applied in a differential manner. Consider measurements $\delta_{ij}$ taken at a first point in time at which the biases are $t_1, r_1, t_2, r_2, t_3, r_3$, and then measurements $\delta'_{ij}$ taken at a second point in time at which the biases are $t'_1, r'_1, t'_2, r'_2, t'_3, r'_3$. The application of the above method to $\delta'_{ij} - \delta_{ij}$ yields the evolution of clock biases, up to a drift that is common to the whole group.

Extension to synchronization between two groups is straightforward. Consider two groups A and B, each group comprising three APs. The reciprocity and synchronization errors within each group may be calibrated through the above-described procedure. Each group will, however, have an unknown remaining clock bias. Let $\delta_{ij}^{A,B} \triangleq t_i^A - r_j^B$ the time discrepancy measured at AP$_j$ in group B, following the known pulse transmission by AP$_i$ in group A. The inter-group synchronization error can be easily obtained by $t_i^A - t_j^B = \delta_{ii}^{A,B} + \delta_{ii}^{B,A} - \delta_{ii}^{B,B}$. Extensions to synchronization between more than two groups follows the same methodology as above. Note that, in a radio stripe system, groups of APs are sequential. Hence, synchronization is only required between a group and its neighbor.

IV. PERFORMANCE OF UBIQUITOUS CELL-FREE MASSIVE MIMO

We will analyze the anticipated performance in two case studies of practical interest: (i) an industrial indoor scenario, and (ii) an outdoor piazza scenario.

A. Industrial Indoor Scenario

Ubiquitous coverage, low latency, ultra-reliable communication, and resilience are key for wireless communications in a factory environment. The flexible distributed cell-free architecture,
with its macro-diversity gain and inherent ability to suppress interference, is suitable to cope with the requirements of this scenario.

We consider the industrial environment described in [14]: a 7-8 m high building with metal ceiling and concrete floors and walls. The industrial inventory mainly consists of metal machinery. The radio stripes are deployed in an area of $100 \times 100$ meters in such a way that 400 APs shape a $20 \times 20$ regular grid, as shown in Fig. 6a (left). The end-most antennas are 5 m apart. They are placed at 6 m above ground level, while the UE antenna height is 2 m.

The performance, in terms of per-user SE and the impact of power control is shown in Fig. 6a (right). We consider $K = 20$ uniformly distributed UEs, mutually orthogonal pilots, and four different DL power control settings, assuming a maximum per-AP radiated power of 200 mW:

1) CD-FPT: Channel-dependent full power transmission. All APs transmit with full power and the power-control coefficients for a given AP $l$ are the same for all $k = 1, \ldots, K$. The power-control coefficient between AP $l$ and UE $k$ is $\rho_{lk} = \left( \sum_{k' = 1}^{K} \gamma_{lk'} \right)^{-1}$, where $\gamma_{lk'}$ is the variance of the corresponding channel estimate $\hat{g}_{lk'}$;

2) MMF: Max-min fairness power control, as described earlier, with all APs serving all UEs.

3) MMF-CQB AP selection: Max-min fairness power control with channel-quality-based AP selection. Each UE is served by a UE-specific subset of APs. This subset consists of the APs that contribute to at least 95% of the channel gain to a given UE [3].

4) MMF-RPB AP selection: Max-min fairness power control with received-power-based AP selection. The subset consists of the APs that contribute to at least 95% of the power assigned to a given UE [3].

Max-min fairness power control doubles the 95%-likely SE compared to the baseline CD-FPT case. Thanks to optimal power control, the radio-stripe system can guarantee each UE almost 4.5 bit/s/Hz. The performance with AP selection is also exploited (dashed and dashed dotted lines). We can see that the SE reduction is minor if the RPB AP selection strategy is used, while the CQB criterion leads to a 20% reduction. The performance gap is attributable to the cardinality of the corresponding AP subsets; on average, CQB uses 17% of the APs and RPB uses 42% of the APs.

B. Piazza Scenario

Installations causing a big visual impact on the environment can be prohibited in areas like piazzas and historic places. In such a scenario, a radio-stripe system can provide all the
(a) Industrial indoor scenario. The carrier frequency is 5200 MHz. We use the one-slope path-loss model defined in [14], with reference distance $d_0 = 15$ m, path-loss at reference distance $PL(d_0) = 70.28$, path-loss exponent $n = 2.59$, and log-normal shadowing standard deviation $\sigma = 6.09$.

(b) Piazza scenario. The carrier frequency is 2 GHz. The large-scale fading is modeled as in [3], assuming uncorrelated shadow fading.

Fig. 6: Left figures illustrate the APs deployments. On the right, the CDF for the per-user SE, as defined in [3 Section III-B]. We choose $L = 400$, $K = 20$ and bandwidth $B = 20$ MHz. The small-scale fading follows i.i.d. Rayleigh distribution. We implement a wrap-around technique to simulate no cell boundaries.
advantages previously described with an unobtrusive deployment. We consider a radio-stripe system that covers a $300 \times 300 \text{ m}^2$ square. The radio stripes are placed along the perimeter of the square at 9 m height, for example, on building facades. There are 400 APs in total, as shown in Fig. 6b (left). We consider $K = 20$ uniformly distributed UEs, mutually orthogonal pilots, and the same power-control policies as before. The maximum per-APs radiated power is set to 400 mW to deal with the large coverage area.

The numerical results are shown in Fig. 6b (right). With max-min power control, we can provide a SE around 4.5 bit/s/Hz/user, doubling the 95%-likely SE compared to the baseline CD-FPT. Due to the AP deployment symmetry, the AP selection strategies perform almost equally well; CQB and RPB select around $1/3$ of the APs on average. The performance gap with respect the case with no AP selection is negligible, thus $2/3$ of the APs can be left out in the transmission towards a given UE.

V. OPEN ISSUES: WHERE THERE’S A WILL, THERE’S A WAY

Cell-free Massive MIMO brings the best of two worlds: the macro-diversity from distributing many APs and the interference cancellation from cellular Massive MIMO. The TDD operation ensures system scalability and distributed processing as the channel estimation and precoding occur at each AP, thus no instantaneous CSI is exchanged over the front-haul. The user-centric data transmission suppresses the inter-cell interference and also contributes to reduce the front-haul overhead. Thanks to all these features, cell-free Massive MIMO succeeds where all the prior coordinated distributed wireless systems failed.

While this article has outlined the basic processing and implementation concepts, many open issues remain, ranging from communication theory to measurements and engineering efforts:

- **Power control**: While (weighted) max-min power-control is computationally tractable and provides uniform quality of service, it does not take actual traffic patterns into account. New power control algorithms are needed to balance fairness, latency, and network throughput, while permitting a distributed implementation.

- **Distributed signal processing**: MR precoding/detection and synchronization can be distributed, as described earlier, but the data encoding/decoding must be carried out at one or multiple CPUs. The distribution of such signal processing tasks over the network is non-trivial, when looking for a good tradeoff between high rates and limited back-haul signaling.
• **Resource allocation and broadcasting:** Scheduling, paging, pilot allocation, system information broadcast, and random access are basic functionalities that traditionally rely on a cellular architecture. New algorithms and protocols are needed for these tasks in cell-free networks.

• **Channel modeling:** The performance analysis of cell-free networks have primarily considered Rayleigh fading channels. Practical channels are likely to contain a mix of line-of-sight and non-line-of-sight paths, and will likely differ substantially depending on the carrier frequency. Dedicated channel measurements followed by refined channel modeling are necessary to better understand the channel characteristics and fine-tune resource allocation algorithms.

• **DL channel estimation:** Recent works [11], [12] show that cell-free networks provide a low degree of channel hardening. DL channel estimates, needed for data decoding, can either be obtained from DL pilots, which increases the pilot overhead, or by blind estimation techniques that uses the DL data. Dedicated algorithms for this estimation are needed.

• **Compliance with existing standards:** The 5G standard is intended to be forward-compatible and only relies on cell-identities for the basic functionalities. It is likely that cell-free data transmission can be implemented in 5G, but further work in standardization and conceptual development is needed.

• **Prototype development:** The step from a promising communication concept to a practical network requires substantial prototyping. The first working cell-free prototype may be pCell, where [15] describes a setup with 32 APs serving 16 UEs. Since every AP in a cell-free network has low cost and footprint, prototyping can be carried out using rather simple components. One can begin by demonstrating the synchronization and joint processing capabilities with a small number of APs in a limited area, and then continue with more APs and larger coverage area.

**References**

[1] J. G. Andrews, X. Zhang, G. D. Durgin, and A. K. Gupta, “Are we approaching the fundamental limits of wireless network densification?” *IEEE Commun. Mag.*, vol. 54, no. 10, pp. 184–190, Oct. 2016.

[2] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.

[3] 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.
[4] S. Shamai and B. M. Zaidel, “Enhancing the cellular downlink capacity via co-processing at the transmitting end,” in Proc. IEEE VTC-Spring, vol. 3, May 2001, pp. 1745–1749.

[5] S. Zhou, M. Zhao, X. Xu, J. Wang, and Y. Yao, “Distributed wireless communication system: A new architecture for future public wireless access,” IEEE Commun. Mag., vol. 41, no. 3, pp. 108–113, Mar. 2003.

[6] G. Caire, S. Ramprashad, and H. Papadopoulos, “Rethinking network MIMO: Cost of CSIT, performance analysis, and architecture comparisons,” in Proc. Information Theory and Applications Workshop (ITA), Jan. 2010, pp. 1–10.

[7] A. Osseiran, J. F. Monserrat, and P. Marsch, 5G Mobile and Wireless Communications Technology. Cambridge University Press, 2016, ch. 9, Coordinated Multi-Point Transmission in 5G.

[8] D. Gesbert, S. Hanly, H. Huang, S. Shamai, O. Simeone, and W. Yu, “Multi-cell MIMO cooperative networks: A new look at interference,” IEEE J. Sel. Areas Commun., vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[9] E. Björnson, R. Zakhour, D. Gesbert, and B. Ottersten, “Cooperative multicell precoding: Rate region characterization and distributed strategies with instantaneous and statistical CSI,” IEEE Trans. Signal Process., vol. 58, no. 8, pp. 4298–4310, Aug. 2010.

[10] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, “Precoding and power optimization in cell-free Massive MIMO systems,” IEEE Trans. Wireless Commun., vol. 16, no. 7, pp. 4445–4459, Jul. 2017.

[11] Z. Chen and E. Björnson, “Channel hardening and favorable propagation in cell-free Massive MIMO with stochastic geometry,” IEEE Trans. Commun., pp. 1–1, 2018.

[12] G. Interdonato, H. Q. Ngo, E. G. Larsson, and P. Freenger, “How much do downlink pilots improve cell-free Massive MIMO?” in Proc. IEEE Global Commun. Conf. (GLOBECOM), Dec. 2016, pp. 1–7.

[13] S. Buzzi and C. D’Andrea, “Cell-free massive MIMO: User-centric approach,” IEEE Wireless Commun. Lett., vol. 6, no. 6, pp. 706–709, Dec. 2017.

[14] E. Tanghe, W. Joseph, L. Verloock, L. Martens, H. Capoen, K. V. Herwegen, and W. Vantomme, “The industrial indoor channel: Large-scale and temporal fading at 900, 2400, and 5200 MHz,” IEEE Trans. Wireless Commun., vol. 7, no. 7, pp. 2740–2751, Jul. 2008.

[15] A. Forenza, S. Perlman, F. Saibi, M. D. Dio, R. van der Laan, and G. Caire, “Achieving large multiplexing gain in distributed antenna systems via cooperation with pCell technology,” in Proc. Asilomar Conf. Signals, Syst., Comput., Nov. 2015, pp. 286–293.