Cell-Free Massive MIMO: A New Next-Generation Paradigm

Jiayi Zhang, Member, IEEE, Shuaifei Chen, Yan Lin, Jiakang Zheng, Bo Ai, Senior Member, IEEE, and Lajos Hanzo, Fellow, IEEE

Abstract—Cell-free (CF) massive multiple-input multiple-output (MIMO) systems have a large number of individually controllable antennas distributed over a wide area for simultaneously serving a small number of user equipment (UE). This solution has been considered as a promising next-generation technology due to its ability to offer a similar quality of service to all UEs, despite its low-complexity signal processing. In this article, we provide a comprehensive survey of CF massive MIMO systems. To be more specific, the benefit of so-called channel hardening and favorable propagation conditions are exploited. Furthermore, we quantify the advantages of CF massive MIMO systems in terms of their energy- and cost-efficiency. Additionally, the signal processing techniques invoked for reducing the fronthaul burden for joint channel estimation and for transmit precoding are analyzed. Finally, the open research challenges in both its deployment and network management are highlighted.

Index Terms—Cell-free massive MIMO, 5G, signal processing, performance analysis.

I. INTRODUCTION

The multiple-input multiple-output (MIMO) concept proposed in their respective seminal papers by Paulraj [1], Foschini [2] and Telatar [3] has been recognized as one of the most disruptive technologies of the recent decades. By employing multiple antennas at the transmitter (TX) and/or the receiver (RX), MIMO systems provide spatial multiplexing and/or diversity gains. Hence they are capable of dramatically increasing the spectral efficiency (SE) without extra frequency, power, or time resources. From the perspective of the number of user equipment (UE) supported on the same time-frequency resources, MIMO systems can be generally divided into two broad categories [4]:

- Single-user MIMO (SU-MIMO) systems enable the base station (BS) to serve a single UE with the aid of multiple antennas.
- Multi-user MIMO (MU-MIMO) systems allow the BS to simultaneously serve multiple UEs, each assigned either one or several antennas.

It is widely recognized that the throughput of SU-MIMO systems is roughly proportional to the minimum of $[N_{TX}, N_{RX}]$, where $N_{TX}$ and $N_{RX}$ are the number of antennas at the TX and RX in rich scattering scenarios. MU-MIMO systems offer advantages over SU-MIMO arrangements, since they significantly reduce the complexity of UEs and improve the sum-rate of downlink broadcast channels. MU-MIMO systems rely on the UE’s separation in the time-domain, frequency-domain, and spatial-domain, where the latter was titled as Spatial Division Multiple Access (SDMA) [5]. To elaborate a little further, SDMA distinguishes the users with the aid of their unique, user-specific CIRs, which requires their accurate estimation.

In order to improve the SE further, the radical new concept of “massive MIMO” has been proposed by Marzetta [6]. As a paradigm shift from MU-MIMO, massive MIMO uses a large number of antennas for simultaneously serving only a few UEs on each time-frequency resource slot by exploiting SDMA principle [7]–[9]. By exploiting the so-called channel hardening\(^1\) phenomena and the resultant favorable propagation conditions, massive MIMO systems substantially reduce both the transmission energy required and the inter-cell interference imposed, despite using low-complexity signal processing in either a centralized or in a distributed manner.

- Centralized massive MIMO systems [6] employ a compact large-scale antenna array at the BS, and they have the advantage of low data-sharing overhead and fronthaul requirements.
- Distributed massive MIMO systems [11] rely on a large number of antennas geographically spread out over a cell, hence they achieve a high diversity gain against shadow fading by exploiting the independent fading of their signals to provide uniformly good service for all UEs.

Distributed massive MIMO systems have been investigated in single-cell and multi-cell scenarios [12], [13]. In principle, the multi-cell distributed massive MIMO concept is a general

\(^1\)The concept of channel hardening claims that the fading channel between the multiple antennas and UEs becomes a near-deterministic scalar channel [10].

This work was supported in part by the National Natural Science Foundation of China (Grant No. 61601020), the Beijing Natural Science Foundation (Grant Nos. 4182049 and L171005), and the open research fund of State Key Laboratory of Integrated Services Networks (Xidian University) (Grant No. JSN20-04). L. Hanzo would like to acknowledge the financial support of the European Research Council’s Advanced Fellow grant QuantCom and by the EPSRC projects EP/Noo4558/1, EP/PO34284/1.

J. Zhang, S. Chen and J. Zheng are with the School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, P. R. China. They are also with State Key Laboratory of Integrated Services Networks (Xidian University) (e-mail: jiayizhang@bjtu.edu.cn).

Y. Lin is with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: yalin@njust.edu.cn).

B. Ai is with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China (e-mail: boai@bjtu.edu.cn).

L. Hanzo is with the School of ECS, University of Southampton, UK (e-mail: lh@ecs.soton.ac.uk).
Figure 1. Comparison of CF massive MIMO and network MIMO systems. In CF massive MIMO, a large number of geographically distributed antennas jointly serve a small number of UEs via a fronthaul network and a CPU in the same time-frequency resource. In network MIMO, a small number of antennas geographically spreads out over a cell and jointly serve every UE in the cell with global CSI. Each cell has a CPU to exchange the local CSI to others.

Recent, the so-called cell-free (CF) massive MIMO network infrastructure has been proposed in [17] as a beneficial incarnation of the general distributed massive MIMO concept. By relying on time-division duplex (TDD) operation, a large number of geographically distributed antennas jointly serve a lower number of UEs with the aid of a fronthaul network and a central processing unit (CPU) operating in the same time-frequency resource. In contrast to the network MIMO concept of Fig. 1, the cellular or cell boundary concepts disappear in CF massive MIMO and the UEs are served simultaneously by all antennas, hence the name. More explicitly, a large number of distributed access points (APs) employing single or multiple antennas simultaneously serve all UEs by exploiting local CSI and performing joint transmission. The CPU sends the downlink (DL) data and power control coefficients to the APs, while the APs feed back the data received from the UEs in the uplink (UL) to the CPU via fronthaul link. In CF mMIMO, only channel statistics are utilized at the CPU to apply joint detection with good performance. The CF massive MIMO solution is eminently suitable for next-generation indoor and hot-spot coverage scenarios, such as a smart factory, train stations, small villages, shopping malls, stadiums, subways, hospitals, community centers and/or a college campus.

Again, the basic premise behind CF massive MIMO systems is to reap all the benefits of network MIMO solutions by providing many more antennas than the number of users which allows us to invoke transmit preprocessing (TPC) for eliminating the interference at the UEs. The key properties of massive MIMO can be beneficially exploited for supporting scalable implementations. Then, the noise, fading and inter-user interference can be averaged out by relying on the law of large numbers. As a benefit, CF massive MIMO systems are capable of providing better service than conventional uncoordinated small cells. Compared with CoMP, each AP of CF massive MIMO systems only requires local CSI to perform both TPC and signal detection [17]. Moreover, CF mMIMO can provide strong macro diversity to select the best out of many around APs [18]. Since the CF massive MIMO concept has attracted a lot of research interests [18]–[24], it is ripe for presenting a survey of the-state-of-the-art, and a taxonomy of its signal processing techniques.

Motivated by the above discussion, the contribution of this paper can be summarized as follows.

- We provide an overview of the CF massive MIMO concept relying on the contemporary research on this topic.
- The substantial benefits of CF massive MIMO systems are quantified and a range of challenging open issues is described.

### II. BENEFITS OF CF MASSIVE MIMO

Compared to traditional cellular massive MIMO networks, the outstanding features of CF massive MIMO schemes are their strong macro diversity and multi-user interference suppression capabilities. Since there is a high probability that every UE is surrounded by a large number of serving APs, all UEs may have good channel conditions. More specifically, the benefits of CF massive MIMO solutions are summarized in Table I and some of them are explicitly detailed in the following.

#### A. Large Energy Efficiency

Energy efficiency is a key performance metric that should be considered when designing future wireless communication systems. It is widely recognized that centralized cellular massive MIMO achieves a high energy efficiency (EE) in terms of Mb/s/Joule, since the radiated power of the individual antennas can be substantially reduced at a given total energy consumption and throughput. Explicitly, this is because a high array gain is provided by centralized massive MIMO solutions, which translates into an improved EE. As a further benefit, for network MIMO, the inter-cell interference can

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Centralized</th>
<th>Network</th>
<th>CF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Antennas</td>
<td>large</td>
<td>moderate</td>
<td>large</td>
</tr>
<tr>
<td>Deployment Cost</td>
<td>high</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>Macro Diversity</td>
<td>small</td>
<td>moderate</td>
<td>large</td>
</tr>
<tr>
<td>Channel Hardening</td>
<td>strong</td>
<td>weak</td>
<td>moderate</td>
</tr>
<tr>
<td>Favorable Propagation</td>
<td>strong</td>
<td>weak</td>
<td>moderate</td>
</tr>
<tr>
<td>Uniform Coverage</td>
<td>bad</td>
<td>moderate</td>
<td>good</td>
</tr>
<tr>
<td>Energy Efficiency</td>
<td>large</td>
<td>small</td>
<td>very large</td>
</tr>
<tr>
<td>Channel Estimation</td>
<td>global</td>
<td>global</td>
<td>local</td>
</tr>
<tr>
<td>Fronthaul Resource</td>
<td>small</td>
<td>large</td>
<td>moderate</td>
</tr>
</tbody>
</table>
be reduced by coherent cooperation between the distributed APs via the high-rate, low-delay fronthaul links. However, the power consumption of fronthaul links reduces the overall EE of network MIMO solutions.

Recently, the EE of CF massive MIMO has been characterized by Ngo et al. [23], who considered the practical effects of power allocation, channel estimation, the number of antennas, the power consumption of the hardware and of the entire fronthaul. The key insight is that the effect of the fronthaul’s power consumption on the EE is significant, especially for a large number of antennas, since the increase in the number of antennas leads to an increase in fronthaul power consumption. Moreover, the EE first increases and then decreases with the number of antennas at the AP [23], because leveraging an excessive number of AP antennas results in a potentially high distance between the distant AP and the UEs. However, we can opt for selecting only a few of the closest APs to serve each UE instead of all APs. On the other hand, optimal power control algorithms [22] can be utilized for further improving the EE. It is surprising that more than a tenfold EE improvement can be achieved by sophisticated AP selection schemes over a centralized massive MIMO scheme.

Moreover, it was found in [25] that compared to cellular massive MIMO, doubled EE can be attained by CF massive MIMO relying on max-min power control in urban, suburban and rural scenarios. Interestingly, the 95% likely per user throughput dramatically increases, which means that CF massive MIMO can enjoy high SE and EE at the same time. This surprising fact is based on its path loss advantage and power control. While previous papers deal with the case in which a traditional sub-6 GHz frequency is used, the authors of [19] proposed an optimal power allocation scheme for CF massive MIMO operating at mmWave frequencies to maximize the global DL EE.

Figure 2 depicts the EE of the CF massive MIMO DL as a function of the number of APs. Here we assume that all APs have the same power consumption of $P_m = 0.2$ W, while all fronthaul links have the power consumption of $P_{b,m} = 0.875$ W. It is clear that the EE is maximized for a specific number of APs $M$. For example, the optimal number of APs for 10 UEs is $M = 20$. Furthermore, it can be seen from the figure for all cases that the EE was noticeably reduced as $M$ increased. These results imply that we should carefully select the number of APs to support the practical implementation of CF massive MIMO.

### B. Flexible and Cost-Efficient Deployment

In the DL the massive MIMO concept supports payload transmission based on low-complexity coherent transmit pre-processing such as maximum-ratio TPC for achieving near-optimal performance. By contrast, in the UL the signal received from each UE can be decoded by invoking receiver combining. Explicitly, the desired signal can be focused on each UE by invoking beamforming-based DL transmit precoding. This is a benefit of spatial diversity, which can be achieved by positioning the antennas more than half a wavelength ($\lambda/2$) apart for ensuring that their signals fade independently.

However, the physical space available at both the BS and UE is always limited. For example, at a carrier frequency of 6 GHz, the wavelength is 5 cm, hence a $\lambda/2$-spaced 100-element uniform planar array (UPA) based massive MIMO scheme employed by a conventional BS has a dimension of 0.5×0.5 m, which cannot be employed in space-constrained environments. By contrast, the CF AP only has a small number of antennas, hence it can be accommodated in space-constrained environments, such as offices, or mounted on lampposts and traffic lights. Given its substantial EE improvement over centralized massive MIMO systems, it is plausible that the total deployment-and-operation cost of CF massive MIMO systems is low. Therefore, CF massive MIMO schemes a cost-efficient solution for next-generation networks.

In network MIMO, a large number of BSs are employed, which requires a large physical space to rent. Furthermore, the EE of network MIMO is much less than the one of CF mMIMO, which means that more power is needed to achieve the same SE for network MIMO. Network MIMO requires both pilot and data signals to be gathered at the CPU, while only channel statistics and data signals are transferred to the CPU in CF mMIMO with less fronthaul overhead. However, if the CF massive MIMO network is infinitely large, it becomes infeasible for all APs to be connected to a single CPU to jointly and simultaneously serve a single UE. In order to circumvent this problem, a distributed and scalable CF massive MIMO framework has been proposed in [26], in which multiple CPUs serving disjoint clusters of APs are considered. This framework brings about the two-fold benefits of having a reduced data payload in the fronthaul and an increased SE for all UEs.
C. The Channel Hardening and the Favorable Propagation Conditions

It is well known that massive MIMO is benefit from channel hardening and favorable propagation [27]. Let \( h_{jk} \) denote the propagation channel spanning from the \( j \)-th BS to the \( k \)-th UE. Channel hardening is defined as the condition of having \( \| h_{jk} \|^2 \rightarrow 1 \) when the number of antennas obeys \( M_j \rightarrow \infty \). The underlying benefit is that channel hardening results in a deterministic channel and hence potentially eliminates the need for combating small-scaling fading. Simple intuitive SE expressions can be obtained by exploiting this property. When \( M_j \rightarrow \infty \) and provided that
\[
E \left\{ \frac{1}{\| h_{jk} \|^2} \right\} \rightarrow 0,
\]
the channel directions of two UEs become asymptotically orthogonal. This property was termed as favorable propagation conditions, which makes it easy to mitigate the inter-user and intra-cell interference with the aid of low-complexity linear signal processing. Therefore, the performance of massive MIMO mainly depends on the large-scale fading, i.e. on the first- and second-order moments of the channels.

Similar to multi-cell massive MIMO, the key idea of CF massive MIMO is also to adopt hundreds or even thousands of antennas. By having many more antennas than users per area unit, the key properties of massive MIMO can be beneficially exploited for scalable implementations. However, the distributed antennas of CF massive MIMO may still impose some spatial channel correlation between the signals arriving from the transmit antennas of specific APs in each others vicinity at a particular receiver UE. Hence it remains unclear, whether the above-mentioned channel hardening and favorable propagation conditions are still valid for CF massive MIMO systems.

Recent results in [28] reveal that the conditions of channel hardening can only be observed for very low pathloss exponents (e.g., \( \alpha < 2 \)), since the channel fading becomes more deterministic in free-space propagation and indoor near-field propagation. Hence, in order to benefit from favorable propagations, we should increase the antenna density and therefore reduce the pathloss exponent with the aid of predominantly line-of-sight (LoS) propagation. In order to accurately capture the network changes, the multi-segment path loss model [29] may be more suitable for CF massive MIMO systems, where the path loss includes both the LoS and NLoS transmissions with a certain probability. This potentially leads to some channel hardening and favorable propagation in the CF massive MIMO-aided systems considered. A key implication is that the accurate performance analysis of CF massive MIMO systems cannot generally rely on the channel hardening and favorable propagation conditions, even though they have been routinely assumed in its centralized massive MIMO systems counterpart. For the uplink, a capacity lower bound was derived for spatially correlated single cell Massive MIMO systems in [30]. However, the exact downlink-rate expression is still unknown at the time of writing.

D. Appealingly Uniform Quality of Service

In CF massive MIMO, the average distance between the closest antennas and an arbitrary UE is substantially reduced by invoking a large number of distributed antennas. Compared to a small-cell system, where a user is served by the one and only closest antenna, an intuitive conclusion is that CF massive MIMO performs better.

To quantify the performance gain attained, the authors of [17] compared the net rate of the above-mentioned small-cell and of the CF massive MIMO systems by taking into account the realistic impairment of channel estimation errors, pilot contamination and of practical max-min power control. Note that the CPU performs power control purely based on the large-scale fading. For \( M = 100 \) and \( K = 40 \), the 95%-likely throughput for any UE of CF massive MIMO is nearly fivefold over the small-cell systems, when the shadow fading is uncorrelated. By contrast, under correlated shadow fading conditions a more than tenfold improvement can be achieved. Therefore, CF massive MIMO systems are capable of providing an improved service for all UEs across the entire coverage area.

As illustrated in Fig. 3, we generated 100 random realizations of the AP locations within a square area of size \( 1 \times 1 \) km\(^2\), and computed the SE per-UE for each realization. We used \( M = 100 \) and \( K = 10 \). The figure shows that the SE is much more concentrated around its median, of about 4.25 bits/s/Hz. More importantly, the SE recorded in all areas is higher than 3.5 bits/s/Hz.

III. SIGNAL PROCESSING IN CF MASSIVE MIMO SYSTEMS

The choice of the most appropriate signal processing techniques is of paramount importance in CF massive MIMO systems, which are different from those of centralized massive MIMO systems in terms of their distributed antenna configurations, different channel fading models and practical constraints imposed on the fronthaul network. These differences have far-reaching effects on the signal processing techniques of CF massive MIMO transceivers.

![Figure 3. SE per-UE of the CF massive MIMO downlink within a square area of 1 × 1 km\(^2\) (M = 100, K = 10).](image-url)
A. Channel Estimation

Accurate channel estimation is important for supporting efficient precoding and signal detection in order to reduce the inter-user interference. In the uplink, the UEs send pilots to all of the distributed antennas of the APs for estimating the channels based on for example the classic MMSE criterion in a decentralized fashion [31]. Pairwise-orthogonal pilot sequences can be used in the network to avoid the pilot contamination effect. Note that CF massive MIMO systems do not exchange the CSI among the APs in the uplink.

In channel hardening based propagation scenarios, which are encountered for example in the downlink of centralized massive MIMO systems experiencing Rayleigh fading, the instantaneous channel gain fluctuates only slightly around its mean value, i.e. we have $h_{mk} \approx \mathbb{E}\{h_{mk}\}$. Hence for example only the channel envelope average has to be known at the UE for reliably detecting the received signals without downlink pilots. By contrast, in CF massive MIMO systems, the channel does not always offer beneficial hardening [28]. Hence, the performance degrades upon simply using $\mathbb{E}\{h_{mk}\}$, since the instantaneous channel gain $h_{mk}$ fluctuates. In [32], a downlink training method was proposed for CF massive MIMO systems and the achievable throughput was considerably improved by the max-min fairness based power control advocated. Nevertheless, this study only considers the case of mutually orthogonal downlink pilots. In order to improve the number of UEs served, the study of non-orthogonal downlink pilots is a promising future subject in conjunction with efficient pilot assignment algorithms.

A preliminary study shows that the DL pilot based training is capable of believably improving the SE despite its low pilot overhead [33]. However, a blind algorithm dispensing with DL pilots and with the estimation of the downlink channel may perform better in CF massive MIMO, than in its cellular massive MIMO counterpart [34].

B. Uplink Signal Detection

During the UL signal detection phase, the UEs transmit their uplink data signals to the APs. The seminal papers [17], [35] advocated the use of matched filtering locally at each AP due to the fact that one can derive the closed-form SE expressions, taking into account the effects of both imperfect CSI, spatially correlated fading and/or hardware impairments. However, it has been shown in [21] that partially or fully centralized UL processing at the CPU can significantly increase the SE. In this context, four different levels of cooperation were introduced in [36] for UL signal detection. With channel statistics, the CPU is capable of maximizing the SE by optimizing the UL receiver weighting coefficient vector. This strategy was termed as large-scale fading decoding (LSFD) in cellular massive MIMO systems [37], and it was shown that the SE of LSFD receiver is much higher than that of the maximal-ratio (MR) receiver in CF massive MIMO networks. On the other hand, the optimization of the receiver filter coefficients at the CPU was considered in [38], which significantly improved the UL user rate. However, the problem is non-convex and has a high computation complexity, which is impractical in CF massive MIMO networks. In conclusion, investigating low-complexity, yet optimal distributed uplink signal detection schemes is still an open question at the time of counting.

IV. Transmission Scheme in CF Massive MIMO Systems

A. Transmit Precoding and Power Control

In CF massive MIMO systems, all APs are connected to a CPU via a fronthaul network for exchanging both their payload data and power control coefficients. Low-complexity maximum-ratio transmit pre-processing [17] or conjugate beamforming [39] is eminently suitable for CF massive MIMO systems, where the APs carry out transmit precoding locally in a distributed manner, rather than at the CPU [17]. Although the fronthaul rate requirement of TPC is low, it imposes a potentially high inter-user interference. More complex TPC techniques, such as zero-forcing [22] (ZF), can be invoked for reducing the inter-user interference and for significantly improving the system’s performance at the cost of an increased fronthaul traffic. It is of high importance for future studies to find the optimal tradeoff between the fronthaul rate requirement imposed and the system performance attained.

The CF massive MIMO systems relying on the ZF precoder are capable of attaining a tenfold improvement in terms of the 95%-likely per-UE rate over small-cell systems at an outage probability of 5% [21]. Both [17] and [21] quantify the benefits of TPC subject to long-term average power constraints formulated as $\sum_{k=1}^{K} \eta_{mk} \mathbb{E}\{|\hat{g}_{mk}|^2\} \leq 1$ for the $m$th AP, where $\eta_{mk}$, represents the power control coefficients and $\hat{g}_{mk}$ denotes the MMSE estimate of the channels gain. By contrast, a short-term average power constraint based TPC has been proposed in [40] obeying $\sum_{k=3}^{K} \eta_{mk} \leq 1$, where the average is only taken over the codewords and it is independent of the small-scale fading. Compared to the classic conjugate beamforming designed under a long-term average power constraint, the one under short-term average power constraint achieves a better coverage, provided that the number of APs is not excessive. Moreover, the authors of [39] proposed an imposed conjugate beamforming scheme, which completely eliminate the self-interference cell-free massive MIMO networks. The imposed conjugate beamforming scheme preserves the utter simplicity of the original conjugate beamforming. Further work may include the design of powerful linear and nonlinear precoders that take into account the distributed network structure and the frequency-selective fading channels effect, whilst operating at a moderate complexity.

The early CF massive MIMO papers [17], [24], [35] assumed that all pilot signals are transmitted at full power during the training phase. However, this may impose grave pilot contamination when a UE has poor channel response. To tackle this problem, one should allocate the most appropriate pilot power for all UEs during the training phase to improve the channel estimation quality and hence the total performance of CF massive MIMO [20].

In order to improve the service quality of all UEs, the classical max-min power control can be invoked by the CPU
The general problem of AP selection in CF massive MIMO systems can be classified as:

1) **Unconstrained**: The unconstrained AP selection problem of CF massive MIMO systems aims for maximizing the performance metric relying on a simple function of the AP selection index, such as coverage distance, received reference signal power, etc. The conceptually simplest method is that each UE may sort the APs and then select the top \( N \) APs according to the above performance metric. However, this method is impractical [42], since it has to know the value of \( N \). Additionally, this method is incapable of dealing with the AP selection problem subject to certain constraints.

2) **Constrained**: When taking into account practical constraints, the simple top-\( N \) AP based selection becomes unsuitable. For instance, the CF massive MIMO systems may suffer from the constraints imposed on the AP’s maximum traffic load, the AP’s maximum transmit power, the system’s maximum fronthaul capacity and so on. In order to solve this kind of problem efficiently, the solutions may rely on matching theory [43], greedy heuristic methods [44], bioinspired algorithms or reinforcement learning [45].

Additionally, a joint design of AP selection, power allocation and interference cancellation at UE was provided in [46] for mitigating the interference effects of CF heterogeneous networks treated as a weighted sum-rate maximization problem. As a further advance, a joint power control and load balancing problem is conceived in [47] for mitigating the inter-user interference of uplink CF massive MIMO systems. Both studies reveal that the proposed solution relying on jointly designing the AP selection and power allocation is beneficial for improving the performance of CF massive MIMO systems.

![Cumulative Distribution](image)

**Figure 4. Cumulative distribution of the per-UE downlink SE for CF massive MIMO systems operating both with and without max-min power control (\( M = 100, K = 10 \)).**

for downlink transmission in CF massive MIMO systems. More specifically, the optimization goal is to maximize the minimum SE of all UEs under the constraint of a per-AP transmit power, which can be determined, for example by the bisection algorithm. As shown in Fig. 4, the per-UE downlink SE achieved by accurate power control is much more concentrated around its median than that of the system operating without power control. All UEs have almost the same SE under accurate max-min power control. However, the bisection algorithm has quite a high complexity. Hence, near-optimal power control methods having significantly lower complexity are of great interest for future research.

**B. AP Selection**

The existing contributions typically assume that all APs jointly transmit their signals to all UEs in CF massive MIMO scenarios. This is however both power-inefficient and impractical, since only a fraction of the APs can beneficially transmit to a specific UE. A practical CF massive MIMO system dynamically allocates a group of APs that have the best channels to each UE. The selected-AP based approach outperforms the all-AP based approach in terms of the 95%-likely per-user rate even with the aid of simple channel estimation algorithms, as described in [18]. It remains however an open question as to whether an optimal power allocation scheme exists for improving the EE of CF massive MIMO systems.

Traditionally, AP selection is assumed to be based on a single connection for each UE in massive MIMO systems [41]. To be specific, each UE is only allowed to be associated with a single AP, albeit each AP can serve multiple UEs simultaneously. In order to exploit the benefits of AP cooperation, multiple AP-UE association has been developed for cell-edge UEs relying on joint transmission techniques.

**V. KEY CHALLENGES AND FUTURE DIRECTIONS**

The family of CF massive MIMO systems relies on a distributed way of deploying the antennas of the BS. The recent theoretical research has improved our conceptual understanding, but it also unveiled entirely new challenges that require further research.

**A. Millimeter Wave Communications**

The millimeter wave (mmWave) band spanning from about 10 GHz to 300 GHz constitutes an important domain for next-generation research [48]. The connection between CF massive MIMO and mmWave systems is straightforward, since both techniques are eminently suitable for indoor short-range communications. More specifically, in CF massive MIMO systems, the distance between the APs and UEs is significantly reduced for the sake of mitigating the pathloss, whilst simultaneously providing macro diversity gains for reducing the detrimental shadowing effects, which constitute the main impediment at mmWave frequencies. It is attractive to employ a large number of APs with small dimension antenna array to utilize mmWave with low hardware and algorithm complexity. Moreover, due to the distributed APs, the path diversity can be employed to circumvent the blockage effects of mmWave [19]. However, many open research problems persist in designing mmWave CF massive MIMO systems.
For example, the channel impulse response (CIR) of mmWave CF massive MIMO systems is typically sparse both in the time- and angular-domain [49]. Hence compressed sensing [50] may be readily invoked for exploiting this sparsity. The channel estimation accuracy may also be improved by the generalized approximate message passing technique of [51], provided that the distribution of the mmWave channel is known a-priori. The high bandwidth of mmWave CF massive MIMO systems is capable of supporting a high data-rate, but it also requires a high hackhaul rate. A particularly attractive solution is to use wireless mmWave fronthaul links instead of fiber-optic cables, which may reduce the operating costs.

B. Non-orthogonal Multiple Access

The concept of non-orthogonal multiple access (NOMA) has recently been proposed as a beneficial multiple access technique for increasing the SE. The so-called power-domain NOMA concept is capable of recognizing and exploiting the received power difference of two or more UEs for supporting them within the same time- and frequency-slot with the aid of successive interference cancellation at the UEs. The beneficial combination of massive MIMO and NOMA has attracted much interest, since both techniques are capable of further improving the SE. In CF massive MIMO systems the distances between the UEs and APs are more diverse than in centralized massive MIMO. Therefore, it is beneficial to combine the NOMA and CF massive MIMO concepts. The first attempt has been made in [52], where the performance of cell-free massive MIMO-NOMA systems was investigated. However, the uplink issues of cell-free massive MIMO-NOMA systems have not been addressed as yet. The performance analysis and the protocol stack design of such a system should be considered in future research.

C. High Mobility Scenarios

Most prior research has been conducted in the scenario of low-mobility UEs, moving at a velocity of less than 10 km/h. In the case of higher-velocity UEs communicating at a high rate and hence having short symbol-durations the multi-path dispersion may exceed the coherence interval. Moreover, when we have a limited number of orthogonal pilots for channel estimation, they cannot be reused in each others’ immediate vicinity, because the resultant pilot contamination may result in excessive interference, which corrupts the channel estimates. In addition, a significant phase-shift of multipath signals may occur in high-mobility scenarios. It is imperative to investigate the impact of phase shift on CF massive MIMO performance. Taking the phase shift into account, the performance loss is closely related to the pilot length [53].

Sophisticated transmit precoding and power allocation algorithms should be designed to cope with high vehicular speeds. A promising solution is to partition the UEs associated with different mobility into different groups. Then a few of the APs may be selected for supporting the UEs associated with high mobility, while the rest of them can be used for the remaining UEs. Alternatively, the UEs having a high mobility can be assigned orthogonal pilots, while the UEs with low mobility can reliably decode the data by using only long-term statistical CSI. Furthermore, noncoherent detection or even blind joint channel estimation dispensing with pilots may be used for mitigating the effects of pilot contamination.

D. Physical Layer Security

In next-generation wireless communications, physical layer security (PLS) is regarded as an attractive extension of the traditional cryptographic security approaches implemented at upper layers. In contrast to the traditional centralized massive MIMO networks, the density of APs and UEs in CF massive MIMO network is much higher, which significantly increases the probability of confidential information being exposed to a malicious eavesdropper. Although PLS in centralized massive MIMO networks has received much attention in the literature [54], the combination of PLS and CF massive MIMO is still limited in the existing literature.

Since the transmission protocol and pilot sequences are standardized and public, the pilot sequences for channel estimation could be exploited by a smart eavesdropper to actively attack the networks. Recent work reveals that an active eavesdropper pilot spoofing attack may result in a significant degradation of secrecy performance in CF massive MIMO [55]. In order to protect the networks from the active attacks, sophisticated attack detection and power control are needed. For example, a simple and efficient minimum description length scheme was proposed in [56] for detecting the pilot spoofing attacks by determining the subspace dimension of the received signal correlation matrix.

By solving power optimization problems, the authors of [24] provided a solution to deal with the case that a UE is suspected of being a eavesdropper in CF massive MIMO networks. The approach is based on a simple detection scheme, under the assumption that the pilot sequences assigned to each UE are mutually orthogonal. However, this assumption is not necessarily realistic in practical scenarios, hence pilot contamination becomes inevitable when the pilots are reused among the UEs. When an active eavesdropper pretends to be a legitimate user, a sophisticated and active eavesdropper detection is required in CF massive MIMO networks, when the pilot contamination effect is also taken account. Since there are numerous distributed APs in the networks which are connected via a fronthaul to a CPU and exchange information, it is feasible to rely on these APs for jointly detecting the presence of eavesdropper and hence for suppressing the attack. Another potential approach is that of exploiting the high spatial correlation of CF massive MIMO systems and then separating the networks into user-centric segments for eliminating the requirement of leaving orthogonal pilot resources. With the pilot contamination suppressed, active eavesdropper rejection becomes more realistic.

E. Hardware Impairments

The existing contributions [17] investigate the fundamental principles of CF massive MIMO relying on ideal hardware. However, both the power consumption and the hardware cost dramatically increase with the number of antennas, which is
considered as one of the bottlenecks in CF massive MIMO systems in practice. Furthermore, the limited fronthaul capability limits data traffic arriving from the APs. For all these reasons, CF massive MIMO systems should rely on low-cost low-power hardware components, especially for APs associated with multiple antennas. Evaluating the effects of realistic imperfect hardware components also requires further research. It is widely recognized that the aggregate impact of various hardware impairments on the centralized massive MIMO systems can be averaged out as a benefit of having unexploited excess degrees of freedom [57]. However, at the time of writing, the effect of hardware impairments on CF massive MIMO has not been widely characterized. Recent work takes a first look at this important topic to show that the impact of hardware impairments at the APs becomes less grave, as the number of APs grows [35].

Figure 5 shows the effect of transceiver hardware impairments on the downlink BE per-UE for CF massive MIMO systems, where we model the hardware impairments by additive Gaussian noise. The BE is reduced upon increasing the level of transceiver hardware impairments \( \kappa_t \) and \( \kappa_r \), which inevitably reduces the received signal power and imposes additional impairment. It is important to observe in Fig. 5 that the effect of hardware impairments at the AP is more pronounced than that at the UEs. Finally, a beneficial tradeoff has to be obtained between the hardware quality and the overall system performance attained.

**F. Prototypes**

While the concepts and benefits of CF massive MIMO have indeed been theoretically confirmed at the time of writings, constructing a comprehensive hardware prototype of such complex systems remains a challenge in practice. As a first step, it would be beneficial to conduct a link-level simulation of CF massive MIMO systems, to guide and validate the design and implementation. Since CF massive MIMO schemes rely on a distributed system philosophy, the precise time and frequency synchronization of the different APs pose a key challenge, which can be tackled by establishing a clock and trigger signal distribution network. To alleviate the processing burden imposed on the CPU, we can partition the whole CF massive MIMO prototype system into sub-systems, each connects the most appropriate number of APs.

**VI. Conclusions**

In this paper we have surveyed the state-of-the-art in CF massive MIMO systems and characterized their performance, algorithms as well as architectures. One of our important findings is that CF massive MIMO systems provide a better coverage than conventional collocated massive MIMO systems and uncoordinated small cells. However, further research is required for fully characterizing CF massive MIMO systems. For example, non-orthogonal downlink pilots should be considered for increasing the number of UEs served. And more sophisticated power control algorithms are desirable for enhancing the system’s performance. Finally, we highlighted a number of key challenges and future research directions for CF massive MIMO systems, which are expected to lead to powerful next-generation solutions.

**References**


Jiayi Zhang (S’08–M’14) received the B.Sc. and Ph.D. degree of Communication Engineering from Beijing Jiaotong University, China in 2007 and 2014, respectively. Since 2016, he has been a Professor with School of Electronic and Information Engineering, Beijing Jiaotong University, China. From 2014 to 2016, he was a Postdoctoral Research Associate with the Department of Electronic Engineering, Tsinghua University, China. From 2014 to 2015, he was also a Humboldt Research Fellow in Institute for Digital Communications, University of Erlangen-Nuernberg, Germany. From 2012 to 2013, he was a visiting Ph.D. student at the Wireless Group, University of Southampton, United Kingdom. His current research interests include Massive MIMO, unmanned aerial vehicle systems and performance analysis of generalized fading channels.

He was recognized as an exemplary reviewer of the IEEE COMMUNICATIONS LETTERS in 2015 and 2016. He was also recognized as an exemplary reviewer of the IEEE TRANSACTIONS ON COMMUNICATIONS in 2017 and 2018. He serves as an Associate Editor for IEEE COMMUNICATIONS LETTERS and IEEE ACCESS. He received the WCSP and IEEE APCC Best Paper Award in 2017.

Shuaifei Chen received his B.S. degree in communication engineering from Beijing Jiaotong University, China, in 2018. From 2018, he is a PhD student in Beijing Jiaotong University. His research interests include massive MIMO systems and machine learning for signal processing in communications.

Yan Lin received the M.S. and Ph.D. degree from Southeast University, China, in 2013 and 2018, respectively. She visited Southampton Wireless Group in University of Southampton, U.K. from Oct. 2016 to Oct. 2017. She is currently a lecturer in the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, China, from 2018. Her current research interests include user-centric networks, ultra-dense networks, vehicular networks, combinatorial optimization and reinforcement learning for wireless communication.

Jiakang Zheng received his B.S. degree in communication engineering from Beijing Jiaotong University, China, in 2019. From 2019, he is a PhD student in Beijing Jiaotong University. His research interests include massive MIMO systems and performance analysis of wireless systems.

Bo Ai (M’00–SM’10) received his Master degree and Ph. D. degree from Xidian University in China. He graduated from Tsinghua University with the honor of Excellent Postdoctoral Research Fellow at Tsinghua University in 2007. He was a visiting professor at EE Department, Stanford University in 2015. He is now working at Beijing Jiaotong University as a full professor and Ph. D. candidate advisor. He is the Deputy Director of State Key Lab of Rail Traffic Control and Safety, and the Deputy Director of International Joint Research Center. He is one of the main responsible people for Beijing “Urban rail operation control system” International Science and Technology Cooperation Base, and the backbone member of the Innovative Engineering Based jointly granted by Chinese Ministry of Education and the State Administration of Foreign Experts Affairs.

He has authored/co-authored 8 books and published over 300 academic research papers in his research area. He has hold 26 invention patents. He has been the research team leader for 26 national projects and has won some important scientific research prizes. Five papers have been the ESI highly-cited paper. He has been notified by Council of Canadian Academies (CCA) that, based on Scopus database, Prof. Bo Ai has been listed as one of the Top 1% authors in his field all over the world. Prof. Bo Ai has also been Feature Interviewed by IET Electronics Letters. His interests include the research and applications of channel measurement and channel modeling, dedicated mobile communications for rail traffic systems.

Prof. Bo Ai is a Fellow of the Institution of Engineering and Technology (IET Fellow), IEEE VTS Distinguished Lecturer. He is an IEEE VTS Beijing Chapter Vice Chair. IEEE BTS Xi’an Chapter Chair. He was as a Co-chair or a Session/Track Chair for many international conferences. He is an associate editor of IEEE Antennas and Wireless Propagation Letters, IEEE Transactions on Consumer Electronics and an Editorial Committee Member of the Wireless Personal Communications journal. He is the Lead Guest Editor for Special Issues on IEEE Transactions on Vehicular Technology, IEEE Antennas and Propagation Letters, International Journal on Antennas and Propagation. He has received many awards such as Distinguished Youth Foundation and Excellent Youth Foundation from National Natural Science Foundation of China, the Qiushi Outstanding Youth Award by Hong Kong Qiushi Foundation, the New Century Talents by the Chinese Ministry of Education, the Zhan Tianyou Railway Science and Technology Award by the Chinese Ministry of Railways, and the Science and Technology New Star by the Beijing Municipal Science and Technology Commission.

Lajos Hanzo (http://www-mobile.ecs.soton.ac.uk) FREng, FIEEE, FIET, Fellow of EURASIP, DSc received his degree in electronics in 1976 and his doctorate in 1983. In 2009 he was awarded an honorary doctorate by the Technical University of Budapest and in 2015 by the University of Edinburgh. In 2016 he was admitted to the Hungarian Academy of Science. During his 40- year career in telecommunications he has held various research and academic posts in Hungary, Germany and the UK. Since 1986 he has been with the School of Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He has successfully supervised 111 PhD students, co-authored 18 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, published 1700+ research contributions at IEEE Xplore, acted both as TPC and General Chair of IEEE conferences, presented keynote lectures and
has been awarded a number of distinctions. Currently he is directing a 60-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) UK, the European Research Councils Advanced Fellow Grant and the Royal Societys Wolfson Research Merit Award. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a Governor of the IEEE VTS. During 2008 - 2012 he was the Editor-in-Chief of the IEEE Press and a Chaired Professor also at Tsinghua University, Beijing. For further information on research in progress and associated publications please refer to http://wwwmobile.ecs.soton.ac.uk.