Foundations of User-Centric Cell-Free Massive MIMO

Other titles in Foundations and Trends® in Signal Processing

Foundations of User-Centric Cell-Free Massive MIMO

Özlem Tuğfe Demir
KTH Royal Institute of Technology
and Linköping University
ozlemtd@kth.se

Emil Björnson
KTH Royal Institute of Technology
and Linköping University
emilbjo@kth.se

Luca Sanguinetti
University of Pisa
luca.sanguinetti@unipi.it

Foundations and Trends® in Signal Processing

Foundations and Trends® in Signal Processing
Volume 14, Issue 3-4, 2021
Editorial Board

Editorial Scope

Topics

Foundations and Trends® in Signal Processing publishes survey and tutorial articles in the following topics:

- Adaptive signal processing
- Audio signal processing
- Biological and biomedical signal processing
- Complexity in signal processing
- Digital signal processing
- Distributed and network signal processing
- Image and video processing
- Linear and nonlinear filtering
- Multidimensional signal processing
- Multimodal signal processing
- Multirate signal processing
- Multiresolution signal processing
- Nonlinear signal processing
- Randomized algorithms in signal processing
- Sensor and multiple source signal processing, source separation
- Signal decompositions, subband and transform methods, sparse representations

- Signal processing for communications
- Signal processing for security and forensic analysis, biometric signal processing
- Signal quantization, sampling, analog-to-digital conversion, coding and compression
- Signal reconstruction, digital-to-analog conversion, enhancement, decoding and inverse problems
- Speech/audio/image/video compression
- Speech and spoken language processing
- Statistical/machine learning
- Statistical signal processing
    - Classification and detection
    - Estimation and regression
    - Tree-structured methods

Information for Librarians

Contents

Foundations of User-Centric Cell-Free Massive MIMO

Özlem Tuğfe Demir[1], Emil Björnson[2] and Luca Sanguinetti[3]

[1]*KTH Royal Institute of Technology and Linköping University; ozlemtd@kth.se*
[2]*KTH Royal Institute of Technology and Linköping University; emilbjo@kth.se*
[3]*University of Pisa; luca.sanguinetti@unipi.it*

ABSTRACT

Imagine a coverage area where each mobile device is com-
municating with a preferred set of wireless access points
(among many) that are selected based on its needs and co-
operate to jointly serve it, instead of creating autonomous
cells. This effectively leads to a user-centric post-cellular
network architecture, which can resolve many of the inter-
ference issues and service-quality variations that appear in
cellular networks. This concept is called User-centric Cell-
free Massive MIMO (multiple-input multiple-output) and
has its roots in the intersection between three technology
components: Massive MIMO, coordinated multipoint pro-
cessing, and ultra-dense networks. The main challenge is to
achieve the benefits of cell-free operation in a practically
feasible way, with computational complexity and fronthaul
requirements that are scalable to enable massively large
networks with many mobile devices. This monograph covers
the foundations of User-centric Cell-free Massive MIMO,
starting from the motivation and mathematical definition. It
continues by describing the state-of-the-art signal processing
algorithms for channel estimation, uplink data reception,

and downlink data transmission with either centralized or distributed implementation. The achievable spectral efficiency is mathematically derived and evaluated numerically using a running example that exposes the impact of various system parameters and algorithmic choices. The fundamental tradeoffs between communication performance, computational complexity, and fronthaul signaling requirements are thoroughly analyzed. Finally, the basic algorithms for pilot assignment, dynamic cooperation cluster formation, and power optimization are provided, while open problems related to these and other resource allocation problems are reviewed. All the numerical examples can be reproduced using the accompanying Matlab code.

1

Introduction and Motivation

The purpose of mobile networks is to provide devices with wireless access to a variety of data services anywhere in a wide geographical area. For many years, the main service of these networks was voice calls, but nowadays transmission of data packets is the dominant service [60]. Hence, the service quality of contemporary networks is mainly determined by the data rate (measured in bit per second) that can be delivered at different locations in the coverage area. The range of wireless transmission is determined by the propagation environment. Since the received signal power decays quadratically, or even faster, with the propagation distance, a traditional mobile network infrastructure consists of a set of geographically distributed transceivers that the connecting device can choose between. These are typically deployed at elevated locations (e.g., in masts and at rooftops) to provide unobstructed propagation to many places in the area. Each transceiver will be called an *access point (AP)* and each user device will be called a *user equipment (UE)* in this monograph.

Current mobile networks are built as *cellular networks*, which means that each UE connects to one AP, namely the one that provides the strongest signal. The UE locations for which a particular AP is selected
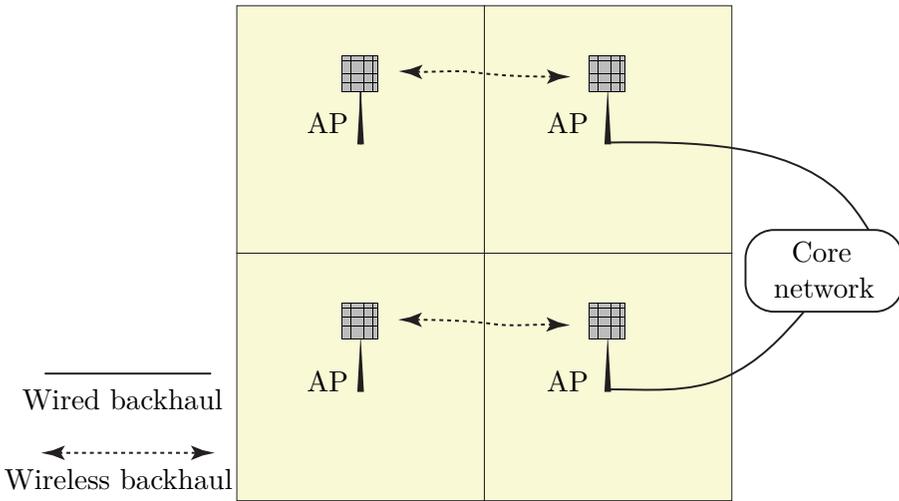
Figure 1.1: Cellular network with four APs, which are all connected to the core network via wired backhaul. Some APs are also interconnected by wireless backhaul.

is called a *cell*. Figure 1.1 shows the basic infrastructure of a cellular network with four APs, each equipped with a planar antenna array containing both the antenna elements and the associated radio units (also known as transceiver chains). The antenna elements emit and receive radio frequency (RF) waves, while the radios generate the analog RF signals to be emitted and process the received RF signals. The radios are connected to a baseband unit that processes the transmitted and received signals in the digital domain. This monograph is focused on the digital signal processing associated with the baseband, thus we will simply refer to each radio and its associated antenna element(s) as *an antenna*. The exact hardware implementation is thereby abstracted away. It is the number of such antennas that determines the dimensionality of the signals that will be generated and processed in the baseband.

The square area around each AP illustrates the cell that the AP provides service to. In reality, the cells will not have symmetric shapes (such as squares, triangles, or hexagons), but it is commonly illustrated like that when describing the fundamentals. The infrastructure of contemporary cellular networks can be divided into two parts: an edge and a core. The edge consists of APs and other hardware units that

are directly involved in the physical-layer communication with the UEs. The core network facilitates all the services requested by the UEs, including routing of data packages and connection to the Internet. The connections between the edge and core are called *backhaul* links and can either be fully wired (e.g., using fiber cables) or partially wireless (e.g., using fixed microwave links). Figure 1.1 shows an example where the APs to the right are connected via wired backhaul links to the core network. The APs to the left are connected wirelessly to the APs to the right, thus their backhaul traffic flows over both wireless and wired links.

An important consequence of the fact that the received signal power rapidly decays with the propagation distance is that the UEs that happen to be close to an AP (i.e., in the cell center) will experience a higher signal-to-noise ratio (SNR) than those that are close to the edge between two cells. A 10 000 times (40 dB) difference is common between the cell center and cell edge. Moreover, UEs at the cell edge are also affected by interference from neighboring APs, thus the signal-to-interference-plus-noise ratio (SINR) can be substantially lower than the SNR at these locations. The data rate is an increasing function of the SINR, thus there are large rate variations in each cell. Figure 1.2(a) exemplifies this behavior by showing the data rate achieved in the downlink by a UE at different locations, when each AP uses a traditional fixed-gain antenna and transmits with maximum power. When the UE is close to one of the APs, it achieves the maximum rate that is supported by the system, which is 80 Mbit/s in this example. In contrast, UEs at the cell edges achieve rates below 1 Mbit/s. This is insufficient for many data services but is nevertheless enough for making voice calls. Depending on the codec, a voice call requires as little as 10-100 kbit/s and this is supported everywhere in this example. Cellular networks were initially designed with this property in mind; we needed the SNR to be above a threshold everywhere in the coverage area to prevent dropped calls, but there was no benefit from being far above that threshold. This basic property has changed entirely when we started using cellular technology for data transmission. Since the UEs request the same data services everywhere in the coverage area, cell-center UEs only need to be connected part of the time, while the cell-edge UEs must be turned

(a) Each AP has a 9 dBi fixed-gain antenna.

(b) Each AP is equipped with 64 omni-directional antennas.
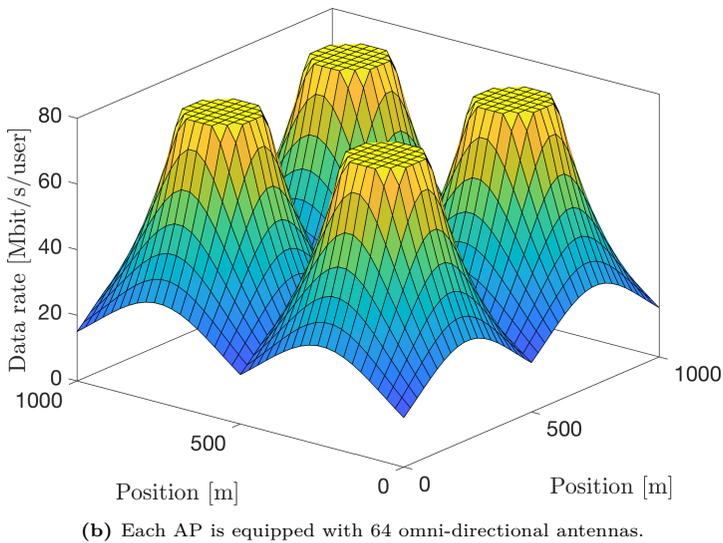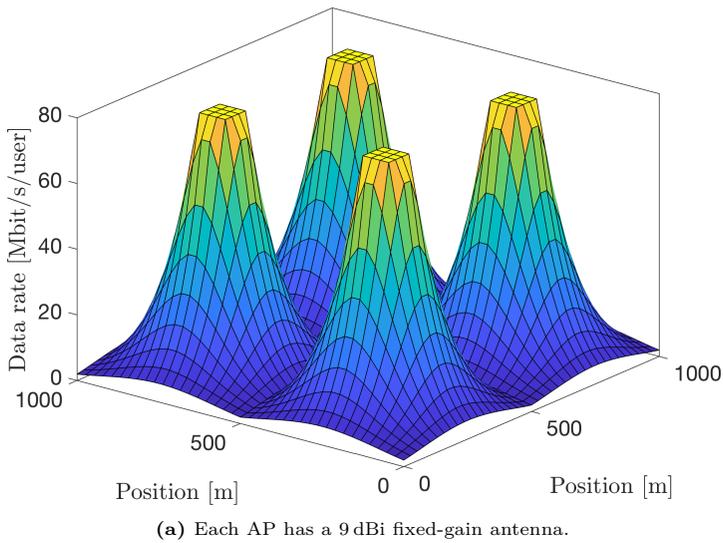
Figure 1.2: Example of the downlink data rate achieved by a UE at different locations in the cellular network in Figure 1.1, assuming each AP transmits with full power. The cell-edge SNR is 0 dB in (a) and the power is assumed to decay as the distance to the power of four. The bandwidth is 10 MHz, and the maximum spectral efficiency (SE) is 8 bit/s/Hz. The key observation is that the rates vary substantially in the network.

on for a much larger fraction of time (if the requested service can even be provisioned). Hence, at a given time instance, the majority of active UEs are at the cell edges and their performance will determine how the customers perceive the service quality of the network as a whole.

The large data rate variations are inherent to the cellular network architecture and remain even if the APs are equipped with advanced hardware, such as *Massive multiple-input multiple-output (MIMO)* [33], [113], [114]. The MIMO technology enables each AP to use an array of antennas (with integrated radios) to serve multiple UEs in its cell by directional transmission, which also increases the SNR and reduces inter-cell interference. More precisely, in the uplink, multiple UEs transmit data to the APs in the same time-frequency resource. The APs exploit the massive number of channel observations (made on the receive antennas) to apply linear receive combining, which discriminates the desired signal from the interfering signals using the spatial domain. In the downlink, the UEs are coherently served by all the antennas, in the same time-frequency resource, but separated in the spatial domain by receiving very directive signals. Figure 1.2(b) shows the downlink data rate achieved by a UE at different locations when each AP has an array of 64 antennas. The data rates are generally higher than in Figure 1.2(a). The cell-center area where the maximum data rate is delivered grows and large improvements are also seen at the cell-edge UEs, since beamforming from the antenna array at the AP can increase the SNR without increasing the inter-cell interference. Despite these gains, there are still substantial rate variations in each cell. Each AP could, in principle, optimize its transmit power to even out the differences (e.g., by reducing the power when serving UEs in the cell center) but this is undesirable since it results in serving all the UEs using the relatively low rates that can be delivered at the cell edge.

Current cellular networks can achieve high peak data rates in the cell centers, but the large variations within each cell make the service quality unreliable. Even if the rates are sufficiently high at, say, 80% of the locations in a cell, this is not sufficient when we are creating a society where wireless access is supposed to be ubiquitous. When payments, navigation, entertainment, and control of autonomous vehicles are all relying on wireless connectivity, we must raise the uniformity of the

data service quality. In summary, the primary goal for future mobile networks should not be to increase the peak rates, but the rates that can be guaranteed to the very vast majority of the locations in the geographical coverage area. The cellular network architecture was not designed for high-rate data services but for low-rate voice services, thus it is time to look beyond the cellular paradigm and make a clean-slate network design that can reach the performance requirements of the future. This monograph considers the cell-free network architecture that is designed to reach the aforementioned goal of uniformly high data rates everywhere.

The cell-free concept for wireless communication networks is defined in Section 1.1, which briefly describes how to operate such networks. Section 1.2 puts the new technology into a historical perspective. Section 1.3 describes three basic benefits that cell-free networks have compared to cellular networks. The key points are summarized in Section 1.4.

1.1   Cell-Free Networks

We will now describe the basic architecture and terminology of a cell-free network. The system and channel propagation models, including the mathematical notation, will be introduced in Section 2 on p. 41.

A cell-free network consists of L geographically distributed APs that are jointly serving the UEs that reside in the area. Each AP is connected via a *fronthaul* to a central processing unit (CPU), which is responsible for the AP cooperation. There can be multiple CPUs all connected via fronthaul links, which can be wired or wireless. An illustration of a cell-free network with single-antenna APs is provided in Figure 1.3. A cell-free network can be divided into an edge and a core, just as cellular networks. The APs and CPUs are at the edge and the connections between them are called fronthaul links, while the connections between the edge and core are still called backhaul links. Hence, the CPUs are connected to the core network via backhaul links, which are used to send/receive data from the Internet and other sources, to facilitate various data services. In contrast, the fronthaul links can be used for: 1) sharing physical-layer signals that will be transmitted in the downlink; 2) forwarding received uplink data signals that are yet to be
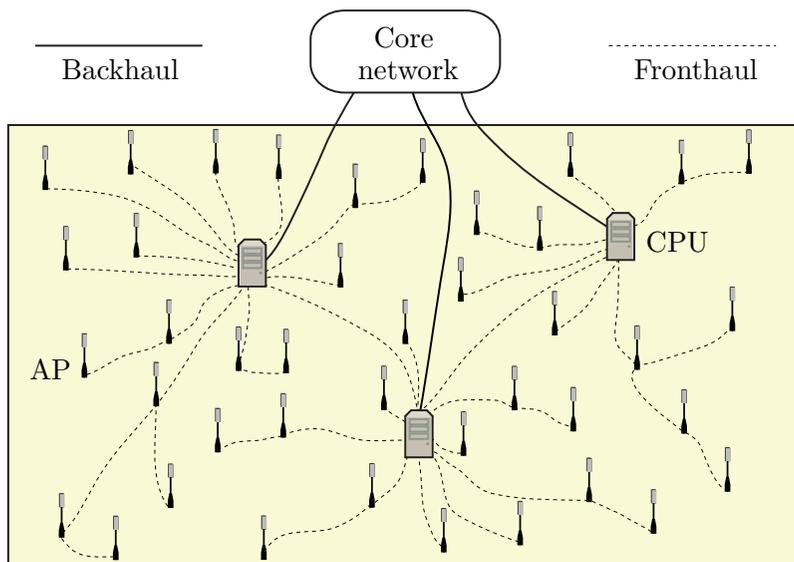
Figure 1.3: Illustration of a cell-free network with many geographically distributed APs connected to CPUs via fronthaul links. The CPUs are connected to the core network via backhaul links. The APs are jointly serving all the UEs in the coverage area.

decoded; and 3) sharing channel state information (CSI) related to the physical channels. The fronthaul also facilitates phase-synchronization between geographically distributed APs, for example, by providing a common phase reference.

A particular fronthaul topology is illustrated in Figure 1.3, where some APs are directly connected to a CPU while other APs are connected via a neighboring AP. We stress that this is only for illustration purposes. No specific assumption on the topology will be made in this monograph, except that the fronthaul links exist, have infinite capacity, negligible latency, and introduce no errors. This allows us to quantify the ultimate physical-layer performance of the cell-free network architecture. Practical constraints on the fronthaul infrastructure are briefly reviewed in Section 7.6 on p. 268. We also note that a CPU may not be a separate physical unit but may be viewed as a logical entity; for example, the CPUs may represent a set of local processors that can be either located
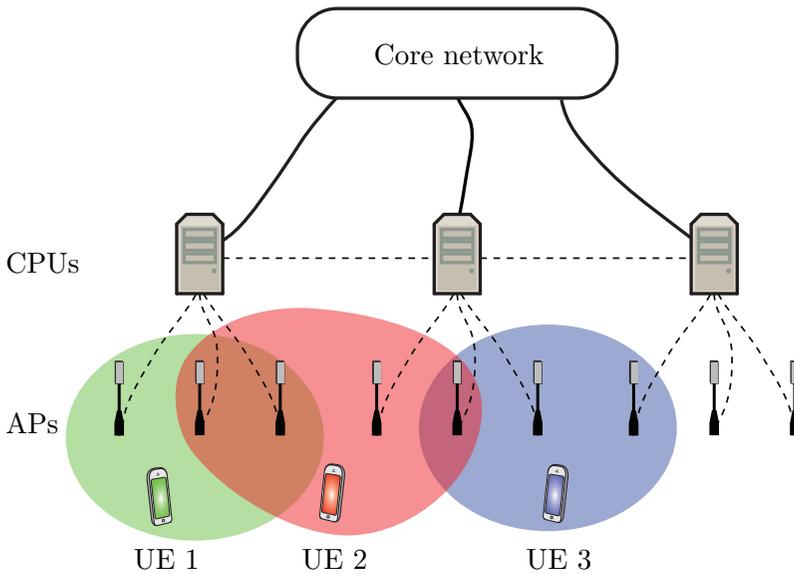
Figure 1.4: Illustrations of the different layers in a cell-free network. Each UE connects to a subset of the APs, which is illustrated by the shaded regions. Each AP is connected to one CPU via fronthaul. The CPUs are interconnected either directly or via the core network.

at a subset of the APs or at other physical locations, and which are connected via fronthaul links. Aligned with the ongoing cloudification of wireless networks [82], [146], known as *cloud radio access network (C-RAN)*, the CPU-related processing tasks can be distributed between the local processors in different ways [25].

Generally speaking, C-RAN is a network deployment architecture where a group of APs is connected to the same CPU, which carry out most of the APs' baseband processing. By sharing computational resources, the total computational capacity can be reduced since it is unlikely that all APs need the maximum capacity simultaneously. One can also make use of general-purpose hardware and open protocols. Recently, the C-RAN abbreviation has started to stand for *centralized RAN*, since the word "cloud" gives the impression that the CPU is owned by another vendor than the wireless network and can be located anywhere in the world. However, to meet the latency constraints of

baseband processing, the CPU is rather an edge-cloud processor located in the same geographical area as the APs. Many different physical-layer technologies can be implemented using the C-RAN architecture. So far, it has mainly been used for cellular networks but it is also the foundation for cell-free networks. Figure 1.4 gives a schematic view of a cell-free network that uses the C-RAN architecture. It is divided into different layers: the core network, the CPU layer, the AP layer, and the UE layer. Each UE is served by a subset of the APs, for example, all the neighboring ones. These subsets are illustrated by the shaded regions in Figure 1.4. For each UE, one of the selected APs is the so-called *Master AP* that is responsible for serving the UE and appointing a CPU where the uplink data decoding and downlink data encoding will be carried out. That CPU delivers the downlink data to all APs that are transmitting to the UE and combines/fuses the uplink received signals obtained at those APs in a final decoding step. A UE can be served by APs connected to different CPUs; there exists a fronthaul link between every pair of APs even if it might go via other entities. The signal processing required for communication can be divided between the APs and CPU in different ways, which will be explored in later sections of this monograph. As the UE moves around, the Master AP assignment, selection of CPU, and selection of cooperating APs may change dynamically.

The word "cell-free" signifies that no cell boundaries exist from a UE perspective during uplink and downlink transmission since all APs that affect a UE will take an active part in the communication. For example, when a UE transmits an uplink data signal then all APs that receive it, with an SNR that is above a threshold, will collaborate in decoding the signal. The partially overlapping shaded regions in Figure 1.4 can be created in that way. The network is jointly serving all the K UEs that are active in the coverage area of the network, even if not all APs might serve every single UE. The differences between cellular and cell-free networks exist at the infrastructure and signal processing side, but can be transparent to the UEs. It should be possible for the same UE to connect to both types of networks without upgrading its software.

To give a first impression of the goal of creating cell-free networks, Figure 1.5(a) shows the downlink data rate achieved by a UE at different
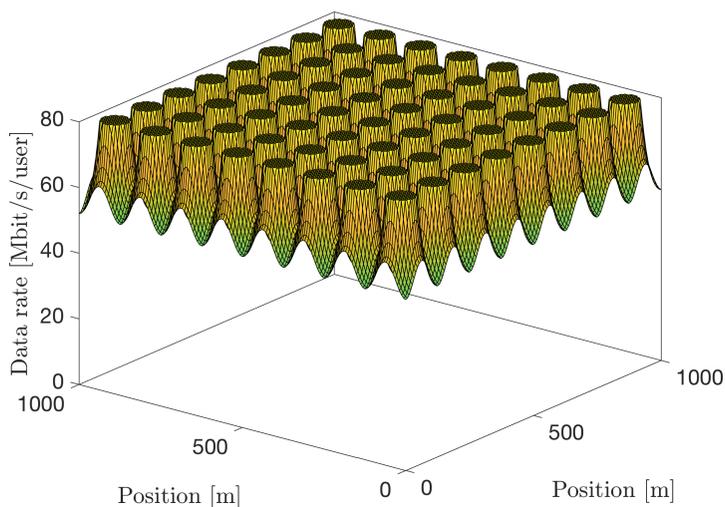
(a) Cell-free network.
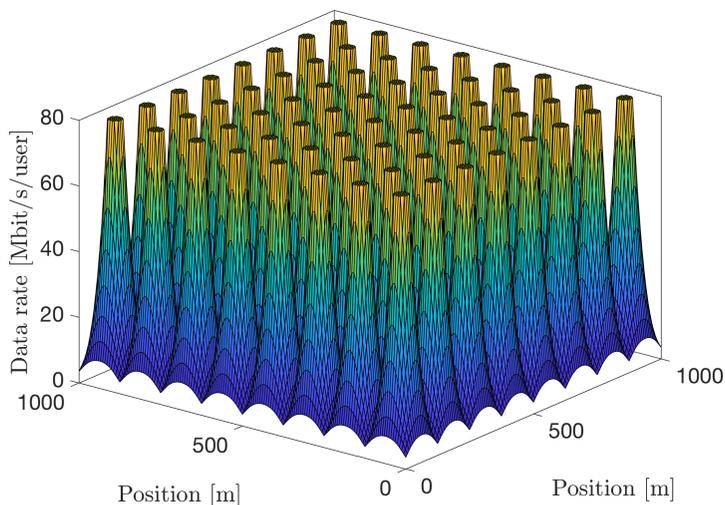
(b) Cellular network with the same AP locations.

Figure 1.5: Example of the downlink data rate achieved by a UE at different locations in a network with 64 APs with omni-directional antennas deployed on a square grid and jointly transmitting to the UE. The propagation parameters are otherwise the same as in Figure 1.2. A cell-free network operation is considered in (a), while a cellular network operation is considered in (b). The key observation is that only the cell-free operation can provide almost uniformly high data rates in the entire network.

locations in a setup that resembles the cellular example in Figure 1.2. For simplicity, an ideal deployment with 64 APs deployed on an 8×8 square grid is considered. The figure shows that the rates vary between 52 and 80 Mbit/s everywhere in the coverage area. One contributing factor is the denser deployment, which greatly reduces the average propagation distance between a UE and the closest AP. However, the main reason is that all the surrounding APs are jointly transmitting to the UE, thereby alleviating the inter-cell interference issue that is one of the main causes of the large rate variations in cellular networks. This is evident when comparing Figure 1.5(a) with Figure 1.5(b), where a cellular network with the same AP locations is considered. The inter-cell interference then gives rise to large rate variations.

1.2 Historical Background

The cellular architecture has played a key role in enabling mobile communications, from the early concepts developed in the 1950s and 1960s [38], [67], [160] to the first commercial deployment in 1979 [102]. The motivating factor of building a cellular network was to make efficient use of the limited frequency spectrum by enabling many concurrent transmissions in the geographical area covered by the network. To control the interference between the transmissions, the coverage area was divided into predefined geographical zones, known as cells, where a fixed AP takes care of the service. In the beginning, a predefined frequency plan was utilized so that adjacent cells use different frequency resources, thereby limiting the inter-cell interference. Over the years, commercial cellular networks have been densified by deploying more APs per area unit [55]. By using steerable multi-antenna panels at each AP, instead of fixed-beam antennas, the interference between adjacent cells can be partially controlled so that the traditional frequency plans can be alleviated. Depending on the deployment scenario (e.g., indoor/outdoor, frequency band, coverage area, and distance from the AP to the closest UE location), different types of AP hardware are utilized [94]. The resulting parts of the cellular networks are sometimes categorized as microcells, picocells, and femtocells. We will use the overarching term *small cells* when referring to such networks [77]. The use of smaller and

smaller cells has been an efficient way to increase the network capacity, in terms of the number of bits per second that can be transferred in a given area. Ideally, the network capacity grows proportionally to the number of APs (with active UEs), but this trend gradually tapers off due to the increasing inter-cell interference [8], [218]. After a certain point, further network densification can actually reduce rather than increase the network capacity. This is particularly the case in the *ultra-dense network* regime [81], [94], [175], where the number of APs is larger than the number of *simultaneously active* UEs. Even if each AP would have a handful of antennas, this is not enough to suppress all the interference in such a dense scenario. A cell-free network is an attempt to move beyond those limits [49], [83], [126], [208], [210]. Before explaining how that can be achieved, we will give a detailed historical background.

As mentioned earlier, a key property of conventional cellular networks is that each UE is assigned to one cell and only served by its AP. This is known as an *interference channel* in information theory and is illustrated in Figure 1.6(a) for the case of three single-antenna transmitters and three single-antenna receivers. Each receive antenna obtains a signal containing the information sent from one desired transmitter (solid line) plus two interfering signals (dashed lines) sent from the undesired transmitters. Even in the absence of noise, identifying the desired signal is like solving an ill-conditioned linear system of equations with three unknowns but only one equation. Hence, the inter-cell interference is unusable in this case; it only limits the performance. When operating such a cellular network, the transmit powers might be adjusted to determine which of the cells will be most affected by the interference. There is no other cooperation between the APs; neither CSI nor transmitted/received signals are shared between cells. These assumptions were challenged by Wyner in [197] from 1994, where the uplink was studied and the benefit of jointly decoding the data from all UEs using the received signals in all cells was explored. In this way, the interference channel is turned into a *multiaccess channel*, where all the receive antennas collaborate. Even if each antenna receives a superposition of multiple signals, there is no unusable interference but the task of the receiver is to extract the information contained in all the received signals. This alternative way of operating the system is
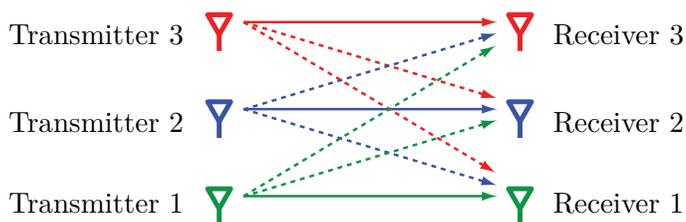
(a) An *interference channel* representing how cellular networks are conventionally operated. Each receiver wants to decode the data sent from its transmitter, subject to the dashed interfering signals from simultaneous transmissions.
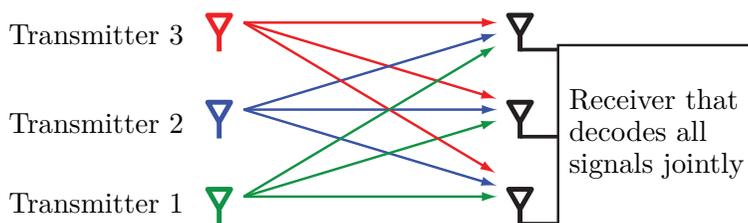
(b) A *multiaccess channel* representing how distributed receive antennas can cooperate to jointly decode the signals from all transmitters. The information contained in all received signals can be utilized. There are no unusable interfering signals. This describes the ideal uplink operation of a cell-free network.
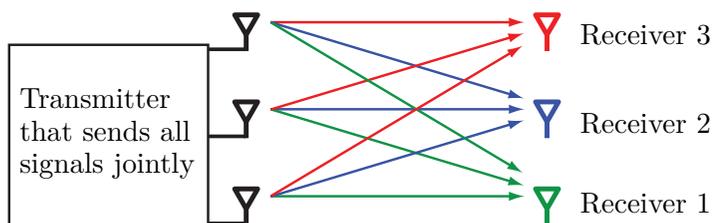
(c) A *broadcast channel* representing how distributed transmit antennas can cooperate to jointly send the signals to all receivers. The signals sent from all antennas can be utilized at each receiver. There are no unusable interfering signals. This describes the ideal downlink operation of a cell-free network.

Figure 1.6: A cellular network is conventionally operated as an interference channel, which is shown in (a). To alleviate inter-cell interference, the uplink of a cell-free network is instead operated as the multiaccess channel shown in (b) and the downlink is operated as the broadcast channel shown in (c).

illustrated in Figure 1.6(b). In this example, the receiver has access to three observations that contain linear combinations of the three desired signals. In the absence of noise, signal detection can be viewed as solving a linear system of equations with three unknowns and three equations, which is a well-conditioned problem. Importantly, the interference is not only canceled by this approach, but the observations made at multiple receive antennas are combined to increase the SNR compared to the case where there was no interference between the transmissions [70]. *Interference is turned from being bad to being good!*

Similarly, Shamai and Zaidel proposed a downlink co-processing framework in [164] from 2001. Using information-theoretic terminology, the cellular downlink was transformed from an interference channel to a *broadcast channel*, where all the geographically distributed transmit antennas collaborate. This case is illustrated in Figure 1.6(c). Each antenna transmits a linear combination of the downlink signals intended for the UEs in all cells, where the linear combination is designed based on the channels to limit inter-cell interference. For example, in the setup shown in Figure 1.6(c) with three geographically distributed transmitters (APs) and three distributed receivers (UEs), zero-forcing (ZF) precoding can be utilized to completely avoid interference. This is not possible in the interference channel in Figure 1.6(a), where each signal is only sent from one transmitter and no precoding can be used.

While the premise of [164], [197] was to add co-processing to an existing cellular network, the idea of building a cell-free network from the outset was pioneered by Zhou, Zhao, Xu, Wang, and Yao in [218] from 2003. Their concept was called *Distributed Wireless Communication System* and resembles the architecture described in Section 1.1 with geographically distributed antennas and processing, and a CPU that controls the system. The paper proposes that a UE should not be served by all the antennas but only by the nearest set of distributed antennas, as illustrated by the shaded regions in Figure 1.4. This is an early step towards a user-centric assignment of network infrastructure, where each UE is served by the user-preferred set of APs instead of by a predefined set. Similar ideas appeared for soft handoff in code-division multiple access (CDMA) systems [192], where UEs at cell edges are jointly served by all the nearest APs.

Many other researchers contributed to this topic during the 2000s and a variety of terminologies have been used to refer to systems where the APs are jointly processing the transmitted and received signals. We will provide some key examples in this paragraph, without attempting to provide an exhaustive list. Non-linear co-processing schemes were developed by Jafar, Foschini, and Goldsmith [88] with the goal of enabling new UEs to be added to a cellular network without affecting the rates of existing UEs. Cooperative downlink processing with multi-antenna APs was studied by Zhang and Dai in [206]. The concept of *Group Cell* was introduced by Zhang, Tao, Zhang, Wang, Li, and Wang in [212] to serve mobile UEs by multiple cells to enable smooth handover during mobility. Multi-cell detection features were also discussed using the group cell name [177]. Coherent coordinated transmission from the APs based on linear ZF precoding and non-linear dirty paper coding was studied by Foschini, Karakayali, and Valenzuela in [66], [95]. The term *Network MIMO* was coined by Venkatesan, Lozano, and Valenzuela in [185] to describe a cellular network where all the APs within the range of a UE share their received signals over a backhaul network, to turn the cellular uplink from an interference channel to a multiaccess channel. Soft handover between distributed antennas in orthogonal frequency-division multiplexing (OFDM) systems was studied by Tölli, Codreanu, and Juntti in [178]. While AP cooperation with infinite-capacity backhaul links was assumed in the above-mentioned works, implementation of joint uplink detection with limited-capacity backhaul was considered by Sanderovich, Somekh, Poor, and Shamai in [155], while the downlink counterpart was studied by Simeone, Somekh, Poor, and Shamai in [173]. Iterative data detection methods, where the APs exchange soft information to reduce the inter-cell interference, were considered by Khattak, Rave, and Fettweis in [99]. Finally, Björnson, Zakhour, Gesbert, and Ottersten showed in [32] that coherent joint transmission can be implemented in time-division duplex (TDD) systems without sharing CSI between the APs, at the cost of increased interference since the AP cannot cancel each others' signals at undesired receivers.

1.2.1   Towards Standardization in 4G

The multi-cell cooperation concepts were considered in the 4G standardization of LTE-Advanced in the late 2000s [144], under the umbrella term of *coordinated multipoint (CoMP)* transmission/reception. The co-processing of data at multiple APs, which is the focus of this monograph, is called *joint processing (JP)* in CoMP [35]. Other CoMP options are coordinated scheduling/precoding where each cell only serves its own UEs, which fall into the category of methods that can be also implemented in conventional cellular networks. Both centralized and decentralized architectures for facilitating JP were explored in the context of CoMP. In the centralized approach, the cooperating APs are connected to a CPU (which might be co-located with an AP) and send their information to it. Hence, the APs can be also viewed as relays that facilitate communication between UEs and the CPU [61]. In the decentralized approach, the cooperating APs only acquire CSI from the UEs [141], but data must still be shared between APs.

Network-Centric Clustering

Since each UE in a conventional cellular network would only be affected by interference from its own cell and a set of neighboring cells, it is only the corresponding cluster of APs that needs to cooperate to alleviate inter-cell interference for this UE. Different ways to implement the AP clustering was explored alongside the development of LTE-Advanced [35]. The starting point for the clustering is that a cellular network already exists and needs to be improved. We will use the example in Figure 1.7(a) to explain the clustering approaches. The first option is *network-centric clustering* where the APs are divided into disjoint clusters [79], [111], [209], each serving a disjoint set of UEs. For example, groups of three neighboring cells can be clustered into a joint region, as illustrated by the colored regions in Figure 1.7(b). Compared to the conventional cellular network in Figure 1.7(a), the cell edges within each cluster are removed, but interference will still occur between clusters. Hence, UEs that are close to a cluster edge might not benefit from the network-centric clustering. The clusters can be changed over time or frequency in an effort to make sure that most of the served UEs
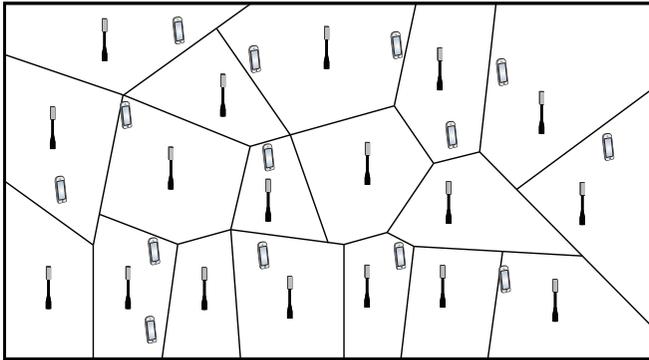
(a) A conventional cellular network, where each UE is only served by one AP.
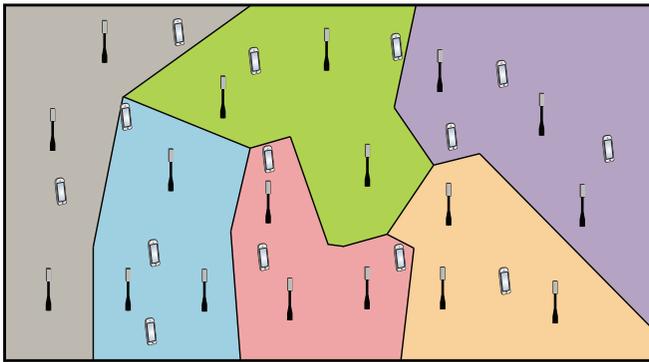
(b) A network-centric implementation of CoMP in a cellular network, where the APs are divided into disjoint clusters. The UEs in a cluster are jointly served by the APs in that cluster.
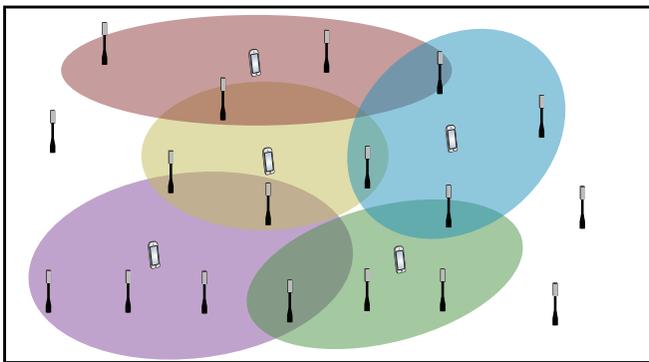
(c) A user-centric implementation of CoMP in a cellular network, where each UE selects a set of preferred APs that will serve it. This is the approach taken also in cell-free networks.

Figure 1.7: Comparison between a conventional cellular network and two ways of implementing multi-cell cooperation.

are in the center of a cluster and not at the edges [46], [93], [110], [141]. The network-centric clustering is conceptually similar to having a conventional cellular network where each cell contains a set of distributed antennas that are controlled by a single AP [54]. Each cell in such a setup corresponds to one cluster in the network-centric clustering.

User-Centric Clustering

Another option is *user-centric clustering* where each UE selects a set of preferred APs [24], [25], [49], [68], [97], [198], [212]. This is illustrated for five UEs in Figure 1.7(c), where each colored region corresponds to the set of APs selected by the corresponding UE. Note that the sets are partially overlapping between neighboring UEs, thus disjoint AP clusters cannot be created to achieve the same result. Irrespective of the UE's location, user-centric clustering will guarantee the control of interference. In this monograph, we will make use of the *dynamic cooperation clustering (DCC)* framework for user-centric clustering, which was introduced by Björnson, Jaldén, Bengtsson, and Ottersten in [24].

If the clusters are well designed, user-centric clustering outperforms network-centric clustering since the latter is essentially a special case of the former. However, both approaches are complicated to add to an existing cellular network since the interfaces between the APs must be standardized to enable cooperation among AP equipment from different vendors. When potential solutions were simulated in the 4G standardization body, the performance gains were often so small that the additional control signaling might remove the gains [35]. An important reason was that the algorithms were jointly designed for frequency-division duplex (FDD) and TDD systems, thus they could not exploit the particular features that only exist in one of these duplexing modes. In particular, CSI for downlink precoding had to be sent around between the APs over low-latency backhaul links to make the system work [138]. It is only in a pure TDD implementation exploiting uplink-downlink channel reciprocity that the CSI necessary for downlink precoding can be obtained at each AP without backhaul signaling [32]. We return to this later in this section.

In Release 10 of LTE-Advanced, only a special case of network-centric clustering was supported [35]: each cluster consists of APs that are

deployed on the same physical site to cover different geographical sectors. Such clustering can only limit the interference between cell sectors, but not between UEs at cell edges. Despite the lack of standardization, the major vendors of AP hardware have made proprietary implementations of CoMP with JP that can only be applied among their own APs. These solutions are often implemented using the C-RAN architecture, which was briefly introduced in Section 1.1. In this cellular context, a set of neighboring APs is connected via a low-latency fronthaul to an edge-cloud processor where the baseband processing is carried out. CoMP algorithms can be conveniently implemented in such a setup. It is not publicly known what CoMP methods are used by different vendors and how well the implementations perform. However, the pCell technology from Artemis [147] is claimed to utilize user-centric clustering.

1.2.2   Cellular Massive MIMO in 5G

Instead of focusing on CoMP, the new feature in the 5G cellular networks is Massive MIMO. This concept was introduced by Marzetta in [113] from 2010 and essentially means that each AP *operates individually* and is equipped with an array of a very large number of active low-gain antennas that can be individually controlled using separate radios (transceiver chains). This stands in contrast to the passive high-gain antennas traditionally used in cellular networks, which might have similar physical dimensions but only a single radio. Massive MIMO has its roots in *space-division multiple access* [7], [151], [176], [196], which was introduced in the 1980s and 1990s to enable multiple UEs to be served by an AP at the same time and frequency. The antenna arrays enable directional transmission to each UE (and directional reception from them), thus UEs located at different locations in the same cell can be served simultaneously with little interference. This technology has later been known as multi-user MIMO.

Benefits

The characteristic feature of Massive MIMO, compared to traditional multi-user MIMO, is that each AP has many more antennas than there are active UEs in the cell. Two important propagation phenomena appear

in those cases [103], [152]: *channel hardening* and *favorable propagation*. The former means that fading channels behave almost as deterministic channels if the antenna signals are processed properly to neutralize the small-scale fading. In principle, the processing makes use of the massive spatial diversity offered by having many antennas. Favorable propagation means that the channels of spatially separated UEs are nearly orthogonal in the spatial domain, since transmission and reception are very spatially directive. We will describe these phenomena in detail in Section 2.6 on p. 66. Motivated by the second phenomenon, it was initially claimed that low-complexity interference-ignoring signal processing methods, such as maximum ratio (MR) processing, are close-to-optimal when each AP is equipped with a large number of antennas. It has later been established that more advanced linear signal processing methods, such as *minimum mean-squared error (MMSE) processing*, are needed to make efficient use of Massive MIMO [23], [76], [124], [156]. In essence, this means that interference must be actively suppressed (one cannot rely on it disappearing automatically when there are many antennas), but the loss in desired signal power is small and there is little need for non-linear methods such as successive interference cancellation [33].

A rigorous framework for analyzing the achievable data rates under imperfect CSI was developed in the Massive MIMO literature and is summarized in recent textbooks, such as [33], [114]. Many tools from this framework will be also utilized in later sections of this monograph.

Limitations

As illustrated in Figure 1.2 earlier in this chapter, Massive MIMO can increase the data rates in a cellular network compared to conventional technology, but large rate variations and inter-cell interference will still remain. Moreover, the 64-antenna panels that have been deployed in 5G cellular networks are not uniform linear arrays (ULAs), as is normally explicitly or implicitly assumed in the Massive MIMO literature [33], [114], but compact planar arrays that can be deployed in the same way as conventional antennas. Since the horizontal width of an array determines its ability to separate UEs located in different azimuth angles with respect to the array (i.e., wider arrays mean better spatial

resolution), the service quality provided by planar arrays is far from what is presented in the literature [12], [31]. In summary, Massive MIMO is a solution to some of the interference problems that are faced in conventional cellular networks. However, a cellular deployment of physically wide horizontal ULAs is practically questionable since it greatly deviates from the form factor of conventional cellular APs. Even if this practical barrier is overcome, the large variations in the distance to the served UEs will still lead to large rate variations of the kind illustrated in Figure 1.2. Hence, a different deployment architecture is required to deliver a more uniform service quality over the coverage area.

1.2.3 Cell-Free Networks Beyond 5G

The *cell-free* terminology was coined by Yang and Marzetta in [201] from 2013, while the name *Cell-free Massive MIMO* first appeared in [127] by Ngo, Ashikhmin, Yang, Larsson, and Marzetta from 2015. While most of the research described earlier adds multi-cell cooperation to an existing cellular network architecture, Cell-free Massive MIMO instead follows in the footsteps of the Distributed Wireless Communication System concept from [218], where a network consisting of distributed cooperating antennas is designed from the outset. The word "massive" refers to an envisioned operating regime with many more APs than UEs [127], and is as an analogy to the conventional Massive MIMO regime in cellular networks; that is, having many more antennas at the infrastructure side than UEs to be served. Interestingly, the envisioned operating regime coincides with that of ultra-dense networks [81], [94], [175], but with the core difference that the APs are cooperating to form a distributed antenna array. The original motivation of Cell-free Massive MIMO was to provide an almost uniformly high service quality in a given geographical area [127], as illustrated in Figure 1.5.

Background

The cell-free architecture, shown in Figure 1.3 and Figure 1.4, was analyzed in the early works [121], [126] with the focus on a distributed operation where the APs perform all the signal processing tasks, except for those that critically require central coordination. The system operates

in TDD mode, which means that the uplink and downlink take place in the same frequency band but are separated in time. Hence, the downlink/uplink channels can be jointly estimated by sending known pilot signals from the UEs to the APs. In this way, each AP obtains local CSI regarding the channels between itself and the different UEs. In the downlink setup studied in [121], [126], each data signal is encoded at a CPU and sent over the fronthaul to the APs, which transmit the signals using MR precoding based on the locally available CSI. Similarly, in the uplink, each AP applies MR combining locally and sends its soft data estimates over the fronthaul to the CPU, which makes the final decoding without having access to any CSI. This concept is well aligned with the cellular joint transmission framework from [32], where the APs only make use of local CSI obtained from uplink pilots in TDD mode. Variations of this type of distributed processing can be found in [16], [28], [40], [63], [139], [200], [211]. One key insight from the more recent works is that the performance can be greatly improved by using MMSE processing instead of MR [28], which is in line with what has also been observed in the Cellular Massive MIMO literature [23], [156]. Hence, even if favorable propagation effects can be observed also in cell-free networks with many distributed antennas [51], it remains important to design the signal processing schemes to actively suppress interference.

The data rates can be also improved by semi-centralized implementations, potentially, at the cost of additional fronthaul signaling. One option is to provide the CPU with statistical CSI so that it can optimize how the uplink data estimates from the APs are combined by taking their relative accuracy into account [4], [28], [122], [130]. For example, an AP that is close to the UE should have more influence than an AP that is further away or that is subject to strong interference. Another option is to let the CPU take care of all the processing while the APs only act as relays [18], [28], [51], [122], [149], [200]. These different options will be analyzed in detail in later sections of this monograph.

The Roots of Cell-Free Massive MIMO

The first papers on Cell-free Massive MIMO assumed all UEs are served by all APs, while the user-centric clustering from the CoMP literature
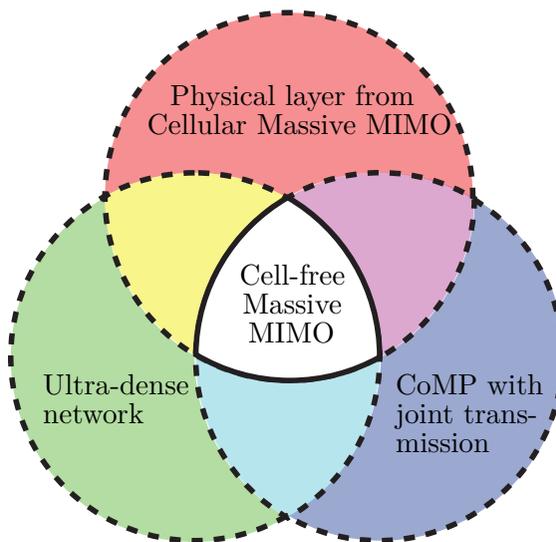
Figure 1.8: Cell-free Massive MIMO can be defined as the intersection between three technology components: The physical layer from Cellular Massive MIMO, the joint transmission concept for distributed APs in the CoMP literature, and the deployment regime of ultra-dense networks.

was first considered in the cell-free context in [40], [41]. A new practical implementation of such clustering was proposed in [84], by first dividing the APs into clusters in a network-centric fashion and then let each UE select a preferred subset of the network-centric clusters. A framework for creating the user-centric clusters in a decentralized fashion was proposed in [29], where the scalability of the different signal processing tasks was also analyzed. Similar user-centric clustering concepts exist in the literature on ultra-dense networks [49].

Many of the concepts described in the Cell-free Massive MIMO literature have previously (or simultaneously) appeared and been analyzed in the cellular literature; for example, in some of the papers mentioned earlier in this section. With this in mind, there are two approaches to defining Cell-free Massive MIMO. The first approach is to specify its unique characteristics. As illustrated by the Venn diagram in Figure 1.8, it can be viewed as the intersection between the physical layer from the Cellular Massive MIMO literature, the joint transmission concept for

distributed APs in the CoMP literature, and the deployment regime of ultra-dense networks. This corresponds to the inner region in the diagram. In other words, we take the best aspects from three technologies, combine them into a single network, and then jointly optimize them to achieve an ultimate embodiment of a wireless network. The second approach is to view Cell-free Massive MIMO as the union of the three circles; that is, an overarching concept focused on cell-free networks but which contains conventional Massive MIMO, conventional CoMP, and conventional ultra-dense networks as three special cases. The presentation of the technical content of this monograph will follow the first approach, thus it is that narrow definition that should be remembered when reading the term "Cell-free Massive MIMO" in later sections. We will focus on describing the foundations of Cell-free Massive MIMO, including the state-of-the-art signal processing and optimization methods. We will focus on how a user-centric viewpoint can be used to identify a scalable implementation, which are two dimensions that are not captured by the Venn diagram. We will compare the achievable performance with that of Cellular Massive MIMO and small cells, which we will extract as two special cases from our analytical formulas. The presentation is not based on a particular set of papers, but is an attempt to summarize the topic as a whole.

1.3 Three Benefits over Cellular Networks

We will end this section by showcasing three major benefits that cell-free networks have compared to conventional cellular networks. More precisely, we compare the setups illustrated in Figure 1.9. The first one is a single-cell setup with a 64-antenna Massive MIMO AP, the second one consists of 64 small cells deployed on a square grid, and the last one is a cell-free network where the same 64 AP locations are used. The comparison of these setups will be made by presenting basic mathematical expressions and simulation results, while a more in-depth analysis of cell-free networks will be provided in later sections.

1.3.1 Benefit 1: Higher SNR With Smaller Variations

The first benefit of the cell-free architecture is that it achieves a higher and more uniform SNR within the coverage area than conventional

cellular networks. To explain this, we assume there is only one active UE in the network and quantify the SNR that the UE achieves in the uplink, when the UE's transmit power is p and the noise power is σ_{ul}^2. In each of the three setups in Figure 1.9, there are 64 antennas. The received power is substantially lower than the transmit power in wireless communications. For the sake of argument, we model the *channel gain* (also known as pathloss or large-scale fading coefficient) for a propagation distance d as (in decibels)

$$\beta(d)\,[\mathrm{dB}] = -30.5 - 36.7\log_{10}\left(\frac{d}{1\,\mathrm{m}}\right). \tag{1.1}$$

The first term says that $30.5\,\mathrm{dB}$ of the power is lost at $1\,\mathrm{m}$ distance while the second term says that another $36.7\,\mathrm{dB}$ of power is lost for every ten-fold increase in the propagation distance. All channels are deterministic and thus known to the transmitters and receivers in this section. A more realistic channel model is provided in Section 2.5 on p. 59 and is then used in the remainder of the monograph.

Massive MIMO Setup

We first consider the single-cell Massive MIMO setup in Figure 1.9(a), where the AP is equipped with $M = 64$ antennas. This represents one cell in a cellular network. We denote by $\mathbf{g} = [g_1 \ldots g_M]^{\mathsf{T}} \in \mathbb{C}^M$ the channel response between the UE and the M antennas. The received uplink signal $\mathbf{y}^{\mathrm{MIMO}} \in \mathbb{C}^M$ at the AP is

$$\mathbf{y}^{\mathrm{MIMO}} = \mathbf{g}s + \mathbf{n} \tag{1.2}$$

where $s \in \mathbb{C}$ is the information signal with transmit power $\mathbb{E}\{|s|^2\} = p$ and $\mathbf{n} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}_M, \sigma_{\mathrm{ul}}^2 \mathbf{I}_M)$ is the receiver noise.

The main task for the AP is to estimate s and this can be done by applying a receive combining vector $\mathbf{v} \in \mathbb{C}^M$ to (1.2), which leads to

$$\hat{s}^{\mathrm{MIMO}} = \mathbf{v}^{\mathsf{H}}\mathbf{y}^{\mathrm{MIMO}} = \mathbf{v}^{\mathsf{H}}\mathbf{g}s + \mathbf{v}^{\mathsf{H}}\mathbf{n}. \tag{1.3}$$

From this expression, it is clear that the SNR is

$$\frac{\mathbb{E}\{|\mathbf{v}^{\mathsf{H}}\mathbf{g}s|^2\}}{\mathbb{E}\{|\mathbf{v}^{\mathsf{H}}\mathbf{n}|^2\}} = \frac{p}{\sigma_{\mathrm{ul}}^2}\frac{|\mathbf{v}^{\mathsf{H}}\mathbf{g}|^2}{\|\mathbf{v}\|^2}. \tag{1.4}$$
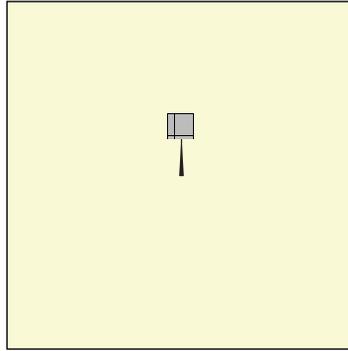
(a) One cell with a 64-antenna AP in a cellular setup.
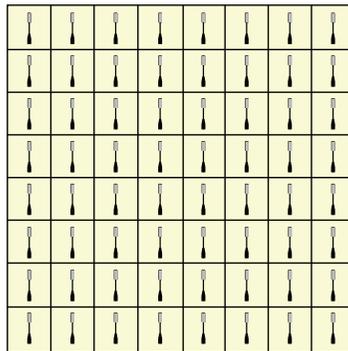
(b) Cellular setup with 64 single-antenna APs.
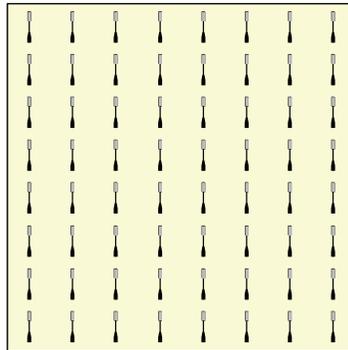
(c) Cell-free setup with the same AP locations as in (b).

Figure 1.9: Three basic setups are compared in Section 1.3: Two cellular networks and one cell-free network (connected to the CPU via fronthaul links, not shown for simplicity).

The AP can select $\mathbf{v}$ based on the channel $\mathbf{g}$ to maximize the SNR. It follows from the Cauchy-Schwartz inequality that (1.4) is maximized when $\mathbf{v}$ and $\mathbf{g}$ are parallel vectors. In particular, the unit-norm MR combining vector $\mathbf{v} = \mathbf{g}/\|\mathbf{g}\|$ can be used to obtain the maximum SNR

$$\mathsf{SNR}^{\mathrm{MIMO}} = \frac{p}{\sigma_{\mathrm{ul}}^2} \|\mathbf{g}\|^2. \tag{1.5}$$

Since all the antennas are co-located in a big array at the AP, there is the same propagation distance d from the UE to all antennas. Hence, $|g_m|^2 = \beta(d)$ for $m = 1, \ldots, M$ using the channel gain model in (1.1). We then obtain

$$\mathsf{SNR}^{\mathrm{MIMO}} = \frac{p}{\sigma_{\mathrm{ul}}^2} M\beta(d) \tag{1.6}$$

which shows that, in a single-cell Massive MIMO system, the SNR is proportional to the number of antennas, M.

Cellular Setup With Small Cells

In the cellular setup in Figure 1.9(b), there are $L = 64$ geographically distributed APs. Each one has a single antenna and the UE will only be served by one of them. We let $h_l \in \mathbb{C}$ denote the channel response between the UE and AP l. In the uplink, the received signal $y_l^{\mathrm{small\text{-}cell}} \in \mathbb{C}$ at AP l is

$$y_l^{\mathrm{small\text{-}cell}} = h_l s + n_l \tag{1.7}$$

where $s \in \mathbb{C}$ denotes the information signal that satisfies $\mathbb{E}\{|s|^2\} = p$ and $n_l \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_{\mathrm{ul}}^2)$ is the receiver noise. The SNR at AP l is

$$\mathsf{SNR}_l^{\mathrm{small\text{-}cell}} = \frac{\mathbb{E}\{|h_l s|^2\}}{\mathbb{E}\{|n_l|^2\}} = \frac{p}{\sigma_{\mathrm{ul}}^2} |h_l|^2. \tag{1.8}$$

The UE needs to choose only one of the APs since this is a conventional cellular network with no cooperation among APs. The UE will naturally select the one providing the largest SNR. Hence, the SNR experienced by the UE becomes

$$\begin{aligned} \mathsf{SNR}^{\mathrm{small\text{-}cell}} &= \max_{l \in \{1, \ldots, L\}} \mathsf{SNR}_l^{\mathrm{small\text{-}cell}} \\ &= \frac{p}{\sigma_{\mathrm{ul}}^2} \max_{l \in \{1, \ldots, L\}} |h_l|^2. \end{aligned} \tag{1.9}$$

If we let d_l denote the distance between the UE and AP l, then $|h_l|^2 = \beta(d_l)$ and the SNR in (1.9) can be rewritten as

$$\mathsf{SNR}^{\text{small-cell}} = \frac{p}{\sigma_{\text{ul}}^2} \max_{l \in \{1,\dots,L\}} \beta(d_l). \qquad (1.10)$$

Cell-Free Setup

In the cell-free setup in Figure 1.9(c), we have the same L APs as in the previous small-cell setup, but the APs are now cooperating to serve the UE. We can write the received signals in (1.7) jointly as

$$\mathbf{y}^{\text{cell-free}} = \mathbf{h}s + \mathbf{n} \qquad (1.11)$$

where $\mathbf{h} = [h_1 \ \dots \ h_L]^{\mathsf{T}}$ and $\mathbf{n} = [n_1 \ \dots \ n_L]^{\mathsf{T}}$. Similar to the single-cell Massive MIMO case above, a receive combining vector $\mathbf{v} \in \mathbb{C}^L$ can be applied to (1.11) in an effort to estimate s. This leads to

$$\hat{s}^{\text{cell-free}} = \mathbf{v}^{\mathsf{H}} \mathbf{y}^{\text{cell-free}} = \mathbf{v}^{\mathsf{H}} \mathbf{h}s + \mathbf{v}^{\mathsf{H}} \mathbf{n}. \qquad (1.12)$$

Since this equation has the same structure as (1.3), it follows that MR combining with $\mathbf{v} = \mathbf{h}/\|\mathbf{h}\|$ provides the maximum SNR:

$$\mathsf{SNR}^{\text{cell-free}} = \frac{p}{\sigma_{\text{ul}}^2} \|\mathbf{h}\|^2 = \frac{p}{\sigma_{\text{ul}}^2} \sum_{l=1}^{L} |h_l|^2. \qquad (1.13)$$

If we compare this expression with that for the small-cell network in (1.9), we observe that the cell-free network obtains an SNR proportional to $\sum_{l=1}^{L} |h_l|^2$, while the small-cell setup only contains the largest term in that sum. Hence, the cell-free network will always obtain a larger SNR, but the difference will be small if there is one term that is much larger than the sum of the others.

If we instead compare the cell-free setup with the single-cell Massive MIMO setup, the main difference is due to the channels $\mathbf{h}$ and $\mathbf{g}$. The SNRs are proportional to $\|\mathbf{h}\|^2$ and $\|\mathbf{g}\|^2$, respectively. We cannot conclude from the mathematical expressions which of these squared norms is the largest. It will depend on the UE location. Therefore, we need to continue the comparison using simulations. Recall that d_l denotes the distance between AP l and the UE, thus we can also write

(1.13) as

$$\mathsf{SNR}^{\text{cell-free}} = \frac{p}{\sigma_{\text{ul}}^2} \sum_{l=1}^{L} \beta(d_l). \tag{1.14}$$

Numerical Comparison

We will now compare the three setups in Figure 1.9 by simulation when the total coverage area is $400\,\text{m} \times 400\,\text{m}$. We will drop one UE uniformly at random in the area and compute the uplink SNRs as described above, assuming the transmit power is $p = 10\,\text{dBm}$ and the noise power is $\sigma_{\text{ul}}^2 = -96\,\text{dBm}$, which are reasonable values when the bandwidth is $10\,\text{MHz}$. When computing the propagation distances, we assume the APs are deployed $10\,\text{m}$ above the UEs.

Figure 1.10 shows the cumulative distribution function (CDF) of the SNR achieved by the UE at different random locations. In the single-cell Massive MIMO case, there are $50\,\text{dB}$ SNR variations, where the largest values are achieved when the UE is right underneath the AP and the smallest values are achieved when the UE is in the corner. The SNR variations are much smaller for the cell-free network, since the distances to the closest AP is generally much shorter than in the Massive MIMO case. Moreover, the SNR is higher at the vast majority of UE locations. If we look at the 95% likely SNR, indicated by the dashed line where the CDF value is 0.05, there is an $18\,\text{dB}$ difference. More precisely, the cell-free network guarantees an SNR of $24.5\,\text{dB}$ (or higher) at 95% of all UE locations, while Massive MIMO only guarantees $6.5\,\text{dB}$. It is only in the upper end of the CDF curves (representing the most fortunate UE locations) that Massive MIMO is the preferred option. This represents the case when the UEs are very close to the 64-antenna Massive MIMO array, while a UE can only be close to a few AP antennas at a time in the cell-free network.

As expected from the analytical expressions, the cell-free network always achieves a higher SNR than the corresponding small-cell setup. The difference is negligible in the upper end of the CDF curves, when the UE is very close to only one of the APs so there is a single dominant term in (1.14), while there is a $4\,\text{dB}$ gap in the 95% likely SNR. Based on this example, we can conclude that distributed antennas are preferred
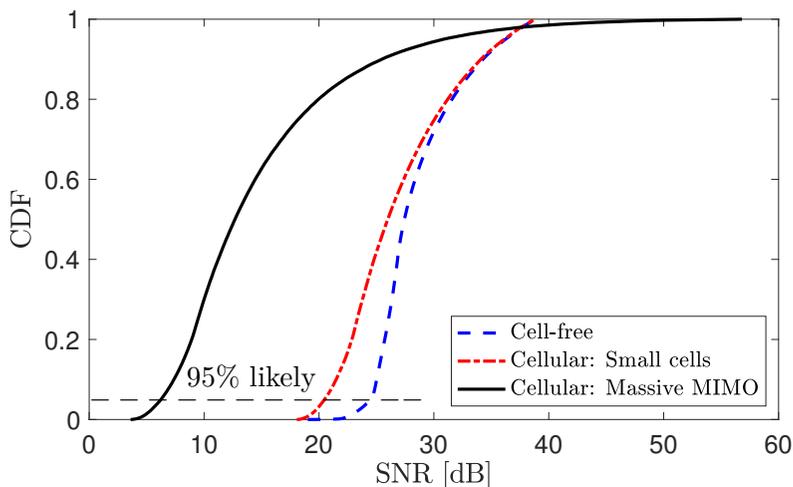
Figure 1.10: The SNR achieved by a UE in each of the setups illustrated in Figure 1.9. The UE location is selected uniformly at random in the area, which gives rise to the CDFs.

over large co-located arrays, but the cell-free architecture only has a minor benefit compared to the cellular small-cell network having the same AP locations. To observe a more convincing practical benefit of the cell-free approach, we need to consider a setup with multiple UEs so that there is interference between the concurrent transmissions.

1.3.2   Benefit 2: Better Ability to Manage Interference

We will now demonstrate that cell-free networks have the ability to manage interference, which is what small-cell networks are lacking. For the sake of argument, we once again consider the uplink of the three setups shown in Figure 1.9 but now with $K = 8$ UEs. We let p denote the transmit power used by each UE, while σ_{ul}^2 denotes the noise power.

Massive MIMO Setup

In the single-cell Massive MIMO setup in Figure 1.9(a), we let $\mathbf{g}_k \in \mathbb{C}^M$ denote the channel from UE k to the AP. Similar to (1.2), the received

uplink signal becomes

$$\mathbf{y}^{\mathrm{MIMO}} = \sum_{i=1}^{K} \mathbf{g}_i s_i + \mathbf{n} \tag{1.15}$$

where $s_i \in \mathbb{C}$ is the information signal transmitted by UE i (with $\mathbb{E}\{|s_i|^2\} = p$) and $\mathbf{n} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}_M, \sigma_{\mathrm{ul}}^2 \mathbf{I}_M)$ is the receiver noise. The AP applies the receive combining vector $\mathbf{v}_k \in \mathbb{C}^M$ to the received signal in (1.15) in an effort to obtain the estimate

$$\hat{s}_k^{\mathrm{MIMO}} = \mathbf{v}_k^{\mathrm{H}} \mathbf{y}^{\mathrm{MIMO}} = \sum_{i=1}^{K} \mathbf{v}_k^{\mathrm{H}} \mathbf{g}_i s_i + \mathbf{v}_k^{\mathrm{H}} \mathbf{n} \tag{1.16}$$

of the signal s_k from UE k. The corresponding SINR is

$$\mathsf{SINR}_k^{\mathrm{MIMO}} = \frac{\mathbb{E}\{|\mathbf{v}_k^{\mathrm{H}} \mathbf{g}_k s_k|^2\}}{\mathbb{E}\left\{\left|\sum_{\substack{i=1\\i\neq k}}^{K} \mathbf{v}_k^{\mathrm{H}} \mathbf{g}_i s_i + \mathbf{v}_k^{\mathrm{H}} \mathbf{n}\right|^2\right\}} = \frac{|\mathbf{v}_k^{\mathrm{H}} \mathbf{g}_k|^2 p}{\mathbf{v}_k^{\mathrm{H}} \left(p \sum_{\substack{i=1\\i\neq k}}^{K} \mathbf{g}_i \mathbf{g}_i^{\mathrm{H}} + \sigma_{\mathrm{ul}}^2 \mathbf{I}_M\right) \mathbf{v}_k}$$

$$\leq p \mathbf{g}_k^{\mathrm{H}} \left(p \sum_{\substack{i=1\\i\neq k}}^{K} \mathbf{g}_i \mathbf{g}_i^{\mathrm{H}} \sigma_{\mathrm{ul}}^2 \mathbf{I}_M\right)^{-1} \mathbf{g}_k \tag{1.17}$$

where the upper bound is achieved by [33, Lemma B.10]

$$\mathbf{v}_k = \left(p \sum_{\substack{i=1\\i\neq k}}^{K} \mathbf{g}_i \mathbf{g}_i^{\mathrm{H}} + \sigma_{\mathrm{ul}}^2 \mathbf{I}_M\right)^{-1} \mathbf{g}_k. \tag{1.18}$$

We will provide a more detailed derivation later in this monograph. For now, the important thing is that the maximum uplink SINR of UE k is given by (1.17).

Cellular Setup With Small Cells

In the small-cell setup in Figure 1.9(b), we let $h_{kl} \in \mathbb{C}$ denote the channel response between UE k and AP l. Similar to (1.7), the received uplink signal at AP l becomes

$$y_l^{\mathrm{small\text{-}cell}} = \sum_{i=1}^{K} h_{il} s_i + n_l \tag{1.19}$$

where $s_i \in \mathbb{C}$ denotes the information signal from UE i (with $\mathbb{E}\{|s_i|^2\} = p$) and $n_l \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_{\mathrm{ul}}^2)$ is the receiver noise. The SINR at AP l with respect to the signal from UE k is

$$\mathsf{SINR}_{kl}^{\text{small-cell}} = \frac{\mathbb{E}\{|h_{kl}s_k|^2\}}{\mathbb{E}\left\{\left|\sum_{\substack{i=1 \\ i \neq k}}^{K} h_{il}s_i + n_l\right|^2\right\}} = \frac{p|h_{kl}|^2}{p\sum_{\substack{i=1 \\ i \neq k}}^{K}|h_{il}|^2 + \sigma_{\mathrm{ul}}^2}. \tag{1.20}$$

Each UE selects to receive service from the AP that provides the largest SINR. Hence, the SINR of UE k is

$$\mathsf{SINR}_k^{\text{small-cell}} = \max_{l \in \{1,\dots,L\}} \mathsf{SINR}_{kl}^{\text{small-cell}}. \tag{1.21}$$

The preferred AP might not be the one with the largest SNR due to the interference. It can happen that one AP serves multiple UEs.

Cell-Free Setup

In the cell-free setup in Figure 1.9(c), the L APs from the small-cell setup are cooperating in detecting the information sent from the K UEs. We can write the received signals in (1.19) jointly as

$$\mathbf{y}^{\text{cell-free}} = \sum_{i=1}^{K} \mathbf{h}_i s_i + \mathbf{n} \tag{1.22}$$

where $\mathbf{h}_i = [h_{i1} \ \dots \ h_{iL}]^{\mathsf{T}}$ and $\mathbf{n} = [n_1 \ \dots \ n_L]^{\mathsf{T}}$. Similar to the single-cell Massive MIMO case, a receive combining vector $\mathbf{v}_k \in \mathbb{C}^L$ is applied to (1.22) to detect the signal from UE k. This leads to the estimate

$$\hat{s}_k^{\text{cell-free}} = \mathbf{v}_k^{\mathsf{H}}\mathbf{y}^{\text{cell-free}} = \sum_{i=1}^{K} \mathbf{v}_k^{\mathsf{H}}\mathbf{h}_i s_i + \mathbf{v}_k^{\mathsf{H}}\mathbf{n} \tag{1.23}$$

of s_k. The corresponding SINR is

$$\mathsf{SINR}_k^{\text{cell-free}} = \frac{\mathbb{E}\{|\mathbf{v}_k^{\mathsf{H}}\mathbf{h}_k s_k|^2\}}{\mathbb{E}\left\{\left|\sum_{\substack{i=1 \\ i \neq k}}^{K} \mathbf{v}_k^{\mathsf{H}}\mathbf{h}_i s_i + \mathbf{v}_k^{\mathsf{H}}\mathbf{n}\right|^2\right\}} = \frac{|\mathbf{v}_k^{\mathsf{H}}\mathbf{h}_k|^2 p}{\mathbf{v}_k^{\mathsf{H}}\left(p\sum_{\substack{i=1 \\ i \neq k}}^{K} \mathbf{h}_i\mathbf{h}_i^{\mathsf{H}} + \sigma_{\mathrm{ul}}^2\mathbf{I}_M\right)\mathbf{v}_k}$$

$$\leq p\mathbf{h}_k^{\mathsf{H}}\left(p\sum_{\substack{i=1 \\ i \neq k}}^{K} \mathbf{h}_i\mathbf{h}_i^{\mathsf{H}} + \sigma_{\mathrm{ul}}^2\mathbf{I}_M\right)^{-1}\mathbf{h}_k \tag{1.24}$$

where the upper bound is achieved by [33, Lemma B.10]

$$\mathbf{v}_k = \left(p\sum_{\substack{i=1 \\ i \neq k}}^{K} \mathbf{h}_i\mathbf{h}_i^{\text{H}} + \sigma_{\text{ul}}^2\mathbf{I}_M \right)^{-1} \mathbf{h}_k. \qquad (1.25)$$

Compared to the single UE case studied earlier, it is harder to utilize the SINR expressions derived in this section to deduce which setup will provide the best performance. Intuitively, the cell-free setup will provide higher SINR than the small-cell setup since we are using the optimal combining vector, while one suboptimal option is to let $\mathbf{v}_k$ contain 1 at the position representing the AP with the highest local SINR and 0 elsewhere. That suboptimal selection would lead to the same SINR as in the small-cell setup. To compare the cell-free setup with the single-cell Massive MIMO setup, we need to run simulations since the SINR expressions in (1.17) and (1.24) have a similar form, but contain channel vectors that are generated differently.

Numerical Comparison

We will now simulate the performance of this multi-user setup using the channel gain model in (1.1) and the same parameter values as in Figure 1.10. More precisely, if the propagation distance is d, then the channel is generated as $\sqrt{\beta(d)}e^{j\phi}$ where ϕ is an independent random variable uniformly distributed between 0 and 2π. This variable models the random phase shift between the transmitter and receiver. This phase was omitted in the previous simulation since the result was determined only by the norms of the channels. However, it is important to include the phases when considering multi-user interference, which is also determined by the directions of the channel vectors.

Figure 1.11 shows the CDF of the SINR achieved by a randomly selected UE in a random realization of the $K = 8$ uniformly distributed UE locations. As compared to Figure 1.10, all the curves are moved to the left in Figure 1.11 due to the interference among the UEs. The Massive MIMO case is barely affected by the interference, which demonstrates that this technology has the ability to separate the UEs' channels spatially using the large array of co-located antennas. However,
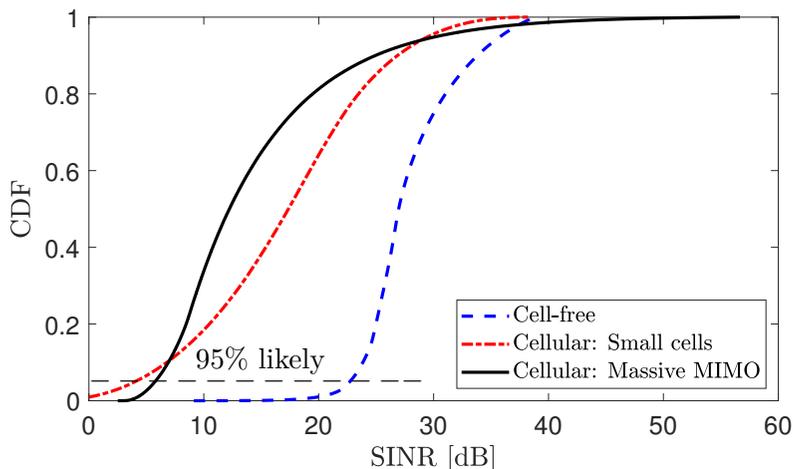
Figure 1.11: The SINR achieved by an arbitrary UE in each of the setups illustrated in Figure 1.9. There are $K = 8$ UEs that are distributed uniformly at random in the area. This gives rise to the CDF.

due to the large variations in distances to the AP, there are $50\,\mathrm{dB}$ variations in the SINR between different UE locations. The cell-free network is also barely affected by the interference, but one can see a tiny lower tail that corresponds to the random event that two UEs are randomly deployed at almost the same location.

The major difference from the single-UE case is that the small-cell curve is moved far to the left and the 95%-likely SINR is even lower than with Massive MIMO. The reason is that each AP only has a single antenna and thus cannot suppress inter-cell interference. The cell-free setup is greatly outperforming the small-cell setup in this multi-user setup. This is what will occur in practice since mobile networks are deployed to serve multiple UEs in the same geographical area.

1.3.3   Benefit 3: Coherent Transmission Increases the SNR

The previous two benefits were exemplified in the uplink but there are also counterparts in the downlink, which lead to similar results but for partially different reasons. One important difference is that the received power in the uplink increases with the number of receive antennas (i.e.,

a larger fraction of the transmit power is collected), thus it is always beneficial to have more antennas. Consider now a downlink scenario where we can deploy any number of antennas, but constrain the total downlink transmit power to be constant (to not change the energy consumption). We then need to determine how the power should be divided between the APs to maximize the SNR. Suppose a UE is in the vicinity of two APs but one has a substantially better channel. It might then seem logical that all the transmit power should be assigned to the AP with the better channel, but we will show that this is not the optimal strategy.

Suppose, for the sake of argument, that AP 1 has the channel response $h_1 = \sqrt{\alpha}$ to the UE, while AP 2 has the channel response $h_2 = \sqrt{\alpha/2}$. If we compare the channel gains $|h_1|^2 = \alpha$ and $|h_2|^2 = \alpha/2$, it is clear that AP 1 has the best channel. Let ρ denote the total downlink transmit power and σ_{dl}^2 denote the receiver noise power. If only AP 1 transmits to the UE, then the SNR at the receiver is

$$\frac{\rho\alpha}{\sigma_{\mathrm{dl}}^2}. \tag{1.26}$$

However, if AP 1 instead transmits with power $2\rho/3$ and AP 2 transmits with power $\rho/3$, which also corresponds to a total power of ρ, then the SNR is

$$\frac{1}{\sigma_{\mathrm{dl}}^2}\left(\sqrt{\frac{2\rho}{3}}h_1 + \sqrt{\frac{\rho}{3}}h_2\right)^2 = 1.5\frac{\rho\alpha}{\sigma_{\mathrm{dl}}^2}. \tag{1.27}$$

Hence, the SNR is higher when we transmit from both APs. This is a consequence of the coherent combination (i.e., constructive interference) of the signals from the two APs. The power gain is reminiscent of the beamforming gain from co-located arrays, where the transmit power is stronger in some angular directions than in other ones, but the physical interpretation is somewhat different. The coherent combination of signals that are transmitted from different geographical points does not give rise to beam patterns but rather local signal amplification in a region around the receiver. Moreover, all the antennas in a co-located array will experience (roughly) the same channel gain so it is logical that they should be jointly utilized for coherent transmission. In contrast, distributed antennas can experience very different channel gains but are anyway useful for coherent transmission.

We will now illustrate that the signal focusing obtained by a distributed array does not give rise to signal beams. Figure 1.12 shows the SNR variations when transmitting from $M = 40$ antennas that are equally spaced along the perimeter of a square. The adjacent antennas are one-wavelength-spaced and transmit with equal power. The received power decays as the square of the propagation distance (as would be the case in free space). If a narrowband signal is considered, the received signals will be phase-shifted (time-delayed) by the ratio between the propagation distance and the signal's wavelength. For the sake of argument, the distances in Figure 1.12 are therefore measured as fractions of the wavelength (each side of the square is ten wavelengths). To achieve a coherent combination at the point of the receiver, each antenna must phase-shift (time-delay) its signal before transmission to make sure that all the M signals are reaching the receiver perfectly synchronized.

In Figure 1.12(a), we show the SNRs measured at different locations when focusing all the signals into a single point. The SNR values are normalized so that they are equal to 1 at the point-of-interest (i.e., the location of the receiver) and smaller elsewhere. We observe that the SNR is much larger at that point than on all the surrounding points, where the 40 signals are not coherently combined. There are some points near the edges of the simulation area where the SNR is also strong but this is not due to a coherent combination of multiple signal components. Instead, it is because these points are close to some of the transmit antennas. Figure 1.12(b) shows the same results but from above. The figure reveals that the SNR is strong in a circular region around the point-of-interest. The diameter of this region is roughly half-a-wavelength. In summary, the signal focusing from distributed arrays will not give rise to angular beams (as in Cellular Massive MIMO) but local signal focusing around the receiver in a region that is smaller than the wavelength. When considering a three-dimensional propagation environment, the SNR will be large within a sphere around the point-of-interest with the diameter being half-a-wavelength. When transmitting multiple signals, we can focus each one at a different point and if these points are several wavelengths apart, the mutual interference will be small according to Figure 1.12.
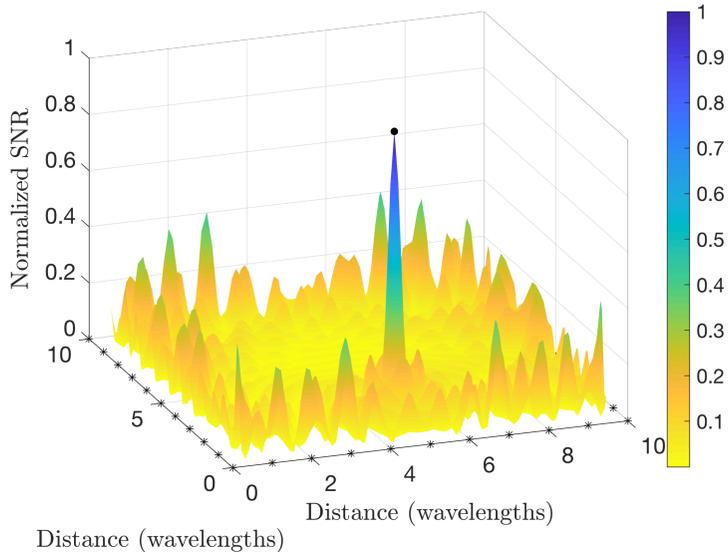
(a) View with normalized SNR (between 0 and 1) on the vertical axis.
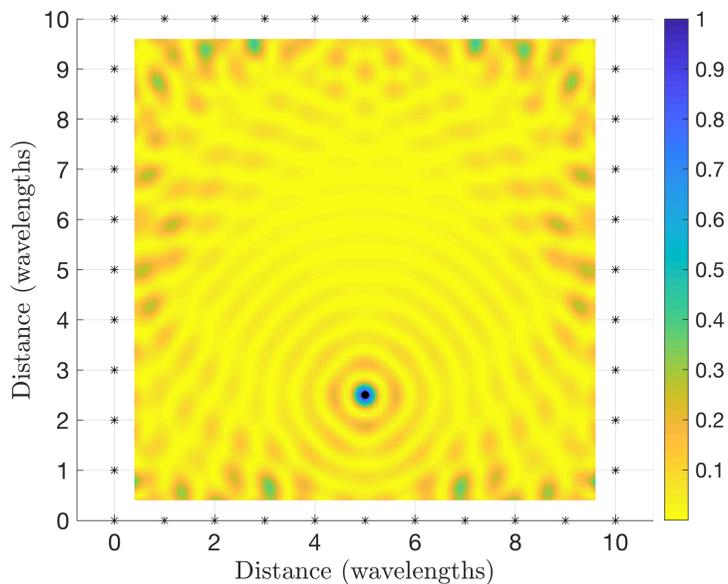
(b) Same as in (a) but viewed from above.

Figure 1.12: The received signal power at different locations when transmitting from one-wavelength-spaced antennas along the walls (each marked with a star). The antennas transmit with equal power and the signals are phase-shifted to achieve coherent combination at the point where the normalized SNR is 1.

1.4   Summary of the Key Points in Section 1

- Traditional wireless networks use the cellular architecture. The cellular approach was conceived for providing wide-area coverage to low-rate voice services. Each AP is surrounded by UEs at very different distances, having widely different SNRs. This architecture is badly suited for providing ubiquitous access to high-rate data services.

- The cell-free architecture turns the situation around: each UE is surrounded by APs. Each AP has relatively simple hardware and cooperates with surrounding APs to jointly serve the UEs in their area of influence. A cell-free network is user-centric if each UE is served by its nearest APs.

- The name *Cell-free Massive MIMO* signifies that it is the combination of three previously known components: The physical layer of Massive MIMO, the vision of creating ultra-dense networks with many more APs than UEs, and the coordinated multipoint methods for achieving a cell-free network. The main novelty lies in how to co-design these components to achieve a user-centric operation that is sufficiently scalable to enable large-scale deployments.

- The first key benefit of the cell-free architecture is the smaller SNR variations compared to cellular networks with a sparse deployment of APs and Massive MIMO.

- The second key benefit is the ability to manage interference by joint processing at multiple APs, which is not done in cellular networks with an equally dense AP deployment.

- The third key benefit is that coherent transmission increases the SNR. It is better to involve APs with weaker channels in the transmission than only using the AP with the best channel.

Acknowledgements

Appendices

A

Notation and Abbreviations

Mathematical Notation

Upper-case boldface letters are used to denote matrices (e.g., $\mathbf{X}, \mathbf{Y}$), while column vectors are denoted with lower-case boldface letters (e.g., $\mathbf{x}, \mathbf{y}$). Scalars are denoted by lower/upper-case italic letters (e.g., x, y, X, Y) and sets by calligraphic letters (e.g., $\mathcal{X}, \mathcal{Y}$).

The following mathematical notations are used:

$\mathbb{C}^{N \times M}$	The set of complex-valued $N \times M$ matrices
$\mathbb{R}^{N \times M}$	The set of real-valued $N \times M$ matrices
$\mathbb{C}^N, \mathbb{R}^N$	Short forms of $\mathbb{C}^{N \times 1}$ and $\mathbb{R}^{N \times 1}$ for vectors
$\mathbb{R}^N_{\geq 0}$	The set of non-negative members of $\mathbb{R}^N$
$x \in \mathcal{S}$	x is a member of the set $\mathcal{S}$
$x \notin \mathcal{S}$	x is not a member of the set $\mathcal{S}$
$\{x \in \mathcal{S} : P\}$	The subset of $\mathcal{S}$ containing all members that satisfy a property P
$[\mathbf{x}]_i$	The ith element of a vector $\mathbf{x}$
$[\mathbf{X}]_{ij}$	The (i, j)th element of a matrix $\mathbf{X}$
$[\mathbf{X}]_{:,1}$	The first column of a matrix $\mathbf{X}$
$\mathrm{diag}(\cdot)$	$\mathrm{diag}(x_1, \ldots, x_N)$ is a diagonal matrix with

	the scalars $x_1, \ldots, x_N$ on the diagonal,
	$\text{diag}(\mathbf{X}_1, \ldots, \mathbf{X}_N)$ is a block diagonal matrix with
	the matrices $\mathbf{X}_1, \ldots, \mathbf{X}_N$ on the diagonal
$\mathbf{X}^\star$	The complex conjugate of $\mathbf{X}$
$\mathbf{X}^\mathsf{T}$	The transpose of $\mathbf{X}$
$\mathbf{X}^\mathsf{H}$	The conjugate transpose of $\mathbf{X}$
$\mathbf{X}^{-1}$	The inverse of a square matrix $\mathbf{X}$
$\mathbf{X}^{\frac{1}{2}}$	The square-root of a square matrix $\mathbf{X}$
$\Re(x)$	Real part of x
j	The imaginary unit
$|x|$	Absolute value (or magnitude) of a scalar variable x
e	Euler's number ($e \approx 2.718281828$)
$\log_a(x)$	Logarithm of x using the base $a > 0$
$\sin(x)$	The sine function of x
$\cos(x)$	The cosine function of x
$\mathrm{tr}(\mathbf{X})$	Trace of a square matrix $\mathbf{X}$
$\mathcal{N}(\mathbf{x}, \mathbf{R})$	The real Gaussian distribution
	with mean $\mathbf{x}$ and covariance matrix $\mathbf{R}$
$\mathcal{N}_\mathbb{C}(\mathbf{0}, \mathbf{R})$	The circularly symmetric complex Gaussian
	distribution with zero mean and correlation matrix $\mathbf{R}$,
	where circular symmetry means that if $\mathbf{y} \sim \mathcal{N}_\mathbb{C}(\mathbf{0}, \mathbf{R})$
	then $e^{\mathrm{j}\phi}\mathbf{y} \sim \mathcal{N}_\mathbb{C}(\mathbf{0}, \mathbf{R})$ for any given ϕ
$\mathbb{E}\{x\}$	The expectation of a random variable x
$\mathbb{V}\{x\}$	The variance of a random variable x
$\|\mathbf{x}\|$	The L_2-norm $\|\mathbf{x}\| = \sqrt{\sum_i |[\mathbf{x}]_i|^2}$ of a vector $\mathbf{x}$
$\mathbf{I}_M$	The $M \times M$ identity matrix
$\mathbf{1}_N$	The $N \times 1$ matrix (i.e., vector) with only ones
$\mathbf{1}_{N \times M}$	The $N \times M$ matrix with only ones
$\mathbf{0}_M$	The $M \times 1$ matrix (i.e., vector) with only zeros
$\mathbf{0}_{N \times M}$	The $N \times M$ matrix with only zeros

Abbreviations

The following acronyms and abbreviations are used in this monograph:

ADC	Analog-to-Digital Converter
AP	Access Point
ASD	Angular Standard Deviation
AWGN	Additive White Gaussian Noise
CDF	Cumulative Distribution Function
CDMA	Code-Division Multiple Access
CoMP	Coordinated Multipoint
C-RAN	Cloud Radio Access Network
CPU	Central Processing Unit
CSI	Channel State Information
DCC	Dynamic Cooperation Clustering
FDD	Frequency-Division Duplex
FIR	Finite Impulse Response
FPA	Fractional Power Allocation
FPC	Fractional Power Control
i.i.d.	Independent and Identically Distributed
JP	Joint Processing
L-MMSE	Local MMSE
LoS	Line-of-Sight
LP-MMSE	Local P-MMSE
LSFD	Large-Scale Fading Decoding
MIMO	Multiple-Input Multiple-Output
MMF	Max-Min Fairness
MMSE	Minimum Mean-Squared Error
MR	Maximum Ratio
MSE	Mean-Squared Error
n-opt	Nearly Optimal
NLoS	Non-Line-of-Sight
NMSE	Normalized MSE
OFDM	Orthogonal Frequency-Division Multiplexing
opt	Optimal
P-RZF	Partial Regularized Zero-Forcing
P-MMSE	Partial MMSE
PDF	Probability Density Function
RF	Radio Frequency
SE	Spectral Efficiency

SINR	Signal-to-Interference-plus-Noise Ratio
SIR	Signal-to-Interference Ratio
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
TDD	Time-Division Duplex
UatF	Use-and-then-Forget
UE	User Equipment
ULA	Uniform Linear Array
ZF	Zero-Forcing

B

Useful Lemmas

This appendix contains a few classical results related to matrices, which are utilized to prove the results in other parts of the monograph.

Lemma B.1 (Matrix inversion lemma). Consider the matrices $\mathbf{A} \in \mathbb{C}^{N_1 \times N_1}$, $\mathbf{B} \in \mathbb{C}^{N_1 \times N_2}$, $\mathbf{C} \in \mathbb{C}^{N_2 \times N_2}$, and $\mathbf{D} \in \mathbb{C}^{N_2 \times N_1}$. The following identity holds if all the involved inverses exist:

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{DA}^{-1}. \tag{B.1}$$

Lemma B.2. For matrices $\mathbf{A} \in \mathbb{C}^{N_1 \times N_2}$ and $\mathbf{B} \in \mathbb{C}^{N_2 \times N_1}$, it holds that

$$\left(\mathbf{I}_{N_1} + \mathbf{AB}\right)^{-1} \mathbf{A} = \mathbf{A} \left(\mathbf{I}_{N_2} + \mathbf{BA}\right)^{-1} \tag{B.2}$$

$$\mathrm{tr}\,(\mathbf{AB}) = \mathrm{tr}\,(\mathbf{BA}). \tag{B.3}$$

Lemma B.3. For the non-zero positive semi-definite matrix $\mathbf{A} \in \mathbb{C}^{N \times N}$ and positive definite matrix $\mathbf{B} \in \mathbb{C}^{N \times N}$, their inner product is strictly positive:

$$\mathrm{tr}\,(\mathbf{AB}) > 0. \tag{B.4}$$

Proof. Consider the eigendecomposition of $\mathbf{A} = \mathbf{U\Lambda U}^{\mathrm{H}}$ where $\mathbf{\Lambda} = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_N)$ with $\lambda_n \geq 0$, for $n = 1, \ldots, N$. Let $\mathbf{u}_n$ denote the nth column of the eigenvalue matrix $\mathbf{U}$. Then, we have

$$\mathrm{tr}\,(\mathbf{AB}) = \sum_{n=1}^{N} \lambda_n \mathbf{u}_n^{\mathrm{H}} \mathbf{B} \mathbf{u}_n \tag{B.5}$$

that is strictly greater than zero since $\mathbf{B}$ is a positive definite matrix and at least one eigenvalue of $\mathbf{A}$ is positive. $\qquad\square$

Lemma B.4. For the positive semi-definite matrices $\mathbf{A} \in \mathbb{C}^{N \times N}$, $\mathbf{C} \in \mathbb{C}^{N \times N}$, and positive definite matrix $\mathbf{B} \in \mathbb{C}^{N \times N}$, the following inequality holds:

$$\text{tr}\left(\mathbf{A}\left(\mathbf{B} + \mathbf{C}\right)^{-1}\right) \leq \text{tr}\left(\mathbf{A}\mathbf{B}^{-1}\right) \tag{B.6}$$

where the equality holds only when $\mathbf{C}\mathbf{B}^{-1}\mathbf{A} = \mathbf{0}_{N \times N}$.

Proof. Consider the eigendecomposition of $\mathbf{C} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\text{H}} = \mathbf{U}_1\boldsymbol{\Lambda}_1\mathbf{U}_1^{\text{H}}$, where $\mathbf{U} \in \mathbb{C}^{N \times N}$ is the unitary matrix of eigenvectors and $\boldsymbol{\Lambda} = \text{diag}\left(\lambda_1, \ldots, \lambda_N\right)$ with the eigenvalues $\lambda_n \geq 0$, for $n = 1, \ldots, N$. $\mathbf{U}_1 \in \mathbb{C}^{N \times r}$ and $\boldsymbol{\Lambda}_1 \in \mathbb{C}^{r \times r}$ are the partitions of $\mathbf{U}$ and $\boldsymbol{\Lambda}$, respectively corresponding to the positive eigenvalues. Applying the matrix inversion lemma (Lemma B.1) to the inverse of $\mathbf{B} + \mathbf{U}_1\boldsymbol{\Lambda}_1\mathbf{U}_1^{\text{H}}$, we can express the left side of the inequality in (B.6) as

$$\text{tr}\left(\mathbf{A}\left(\mathbf{B} + \mathbf{C}\right)^{-1}\right)$$
$$= \text{tr}\left(\mathbf{A}\mathbf{B}^{-1}\right) - \text{tr}\left(\mathbf{A}\mathbf{B}^{-1}\mathbf{U}_1\left(\mathbf{U}_1^{\text{H}}\mathbf{B}^{-1}\mathbf{U}_1 + \boldsymbol{\Lambda}_1^{-1}\right)^{-1}\mathbf{U}_1^{\text{H}}\mathbf{B}^{-1}\right) \tag{B.7}$$

that is strictly less than $\text{tr}\left(\mathbf{A}\mathbf{B}^{-1}\right)$ if $\mathbf{U}_1^{\text{H}}\mathbf{B}^{-1}\mathbf{A}\mathbf{B}^{-1}\mathbf{U}_1$ is non-zero by Lemma B.3 noting that the matrix $\left(\mathbf{U}_1^{\text{H}}\mathbf{B}^{-1}\mathbf{U}_1 + \boldsymbol{\Lambda}_1^{-1}\right)^{-1}$ is positive definite. If $\mathbf{U}_1^{\text{H}}\mathbf{B}^{-1}\mathbf{A}\mathbf{B}^{-1}\mathbf{U}_1 = \mathbf{0}_{r \times r}$ that is equivalent to $\mathbf{C}\mathbf{B}^{-1}\mathbf{A} = \mathbf{0}_{N \times N}$, then both sides of (B.6) are equal. $\qquad\square$

Lemma B.5. Consider the vector $\mathbf{a} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}_N, \mathbf{A})$, with covariance matrix $\mathbf{A} \in \mathbb{C}^{N \times N}$, and any deterministic matrix $\mathbf{B} \in \mathbb{C}^{N \times N}$. It holds that

$$\mathbb{E}\{|\mathbf{a}^{\text{H}}\mathbf{B}\mathbf{a}|^2\} = |\text{tr}(\mathbf{B}\mathbf{A})|^2 + \text{tr}(\mathbf{B}\mathbf{A}\mathbf{B}^{\text{H}}\mathbf{A}). \tag{B.8}$$

Proof. Note that $\mathbf{a} = \mathbf{A}^{\frac{1}{2}}\mathbf{w}$ for $\mathbf{w} = [w_1 \ldots w_N]^{\text{T}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}_N, \mathbf{I}_N)$, thus we can write

$$\mathbb{E}\{|\mathbf{a}^{\text{H}}\mathbf{B}\mathbf{a}|^2\} = \mathbb{E}\{|\mathbf{w}^{\text{H}}(\mathbf{A}^{\text{H}})^{\frac{1}{2}}\mathbf{B}\mathbf{A}^{\frac{1}{2}}\mathbf{w}|^2\} = \mathbb{E}\{|\mathbf{w}^{\text{H}}\mathbf{C}\mathbf{w}|^2\} \tag{B.9}$$

where we defined $\mathbf{C} = (\mathbf{A}^{\text{H}})^{\frac{1}{2}}\mathbf{B}\mathbf{A}^{\frac{1}{2}}$. Let c_{n_1,n_2} denote the element in $\mathbf{C}$ on row n_1 and column n_2. Using this notation, we can expand (B.9) as

$$\mathbb{E}\{|\mathbf{w}^{\text{H}}\mathbf{C}\mathbf{w}|^2\} = \sum_{n_1=1}^{N}\sum_{n_2=1}^{N}\sum_{m_1=1}^{N}\sum_{m_2=1}^{N}\mathbb{E}\{w_{n_1}^{\star}c_{n_1,n_2}w_{n_2}w_{m_1}c_{m_1,m_2}^{\star}w_{m_2}^{\star}\}$$

$$\stackrel{(a)}{=} \sum_{n=1}^{N} \mathbb{E}\{|w_n|^4\}|c_{n,n}|^2 + \sum_{n=1}^{N} \sum_{\substack{m=1 \\ m \neq n}}^{N} \mathbb{E}\{|w_n|^2\}\mathbb{E}\{|w_m|^2\}c_{n,n}c_{m,m}^{\star}$$

$$+ \sum_{n_1=1}^{N} \sum_{\substack{n_2=1 \\ n_2 \neq n_1}}^{N} \mathbb{E}\{|w_{n_1}|^2\}\mathbb{E}\{|w_{n_2}|^2\}|c_{n_1,n_2}|^2$$

$$\stackrel{(b)}{=} \sum_{n=1}^{N} 2|c_{n,n}|^2 + \sum_{n=1}^{N} \sum_{\substack{m=1 \\ m \neq n}}^{N} c_{n,n}c_{m,m}^{\star} + \sum_{n_1=1}^{N} \sum_{\substack{n_2=1 \\ n_2 \neq n_1}}^{N} |c_{n_1,n_2}|^2$$

$$= \sum_{n=1}^{N} \sum_{m=1}^{N} c_{n,n}c_{m,m}^{\star} + \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} |c_{n_1,n_2}|^2$$

$$\stackrel{(c)}{=} |\mathrm{tr}(\mathbf{C})|^2 + \mathrm{tr}(\mathbf{C}\mathbf{C}^{\mathrm{H}}) \tag{B.10}$$

where (a) utilizes that circular symmetry implies that $\mathbb{E}\{w_{n_1}^{\star} w_{n_2} w_{m_1} w_{m_2}^{\star}\}$ is only non-zero when the terms with conjugates have matching indices to the terms without conjugates. The first expression is given by $n_1 = n_2 = m_1 = m_2$, the second term is given by $n_1 = n_2$ and $m_1 = m_2$ with $n_1 \neq m_1$, and the third term is given by $n_1 = m_1$ and $n_2 = m_2$ with $n_1 \neq n_2$. In (b), we utilize that $\mathbb{E}\{|w_n|^2\} = 1$ and $\mathbb{E}\{|w_n|^4\} = 2$. In (c), we write the sums of elements in $\mathbf{C}$ using the trace. The resulting expression is equivalent to (B.8), which is shown by replacing $\mathbf{C}$ with $\mathbf{A}$ and $\mathbf{B}$ and utilizing the fact that $\mathrm{tr}(\mathbf{C}_1\mathbf{C}_2) = \mathrm{tr}(\mathbf{C}_2\mathbf{C}_1)$ for any matrices $\mathbf{C}_1, \mathbf{C}_2$ such that $\mathbf{C}_1$ and $\mathbf{C}_2^{\mathsf{T}}$ have the same dimensions. $\qquad\square$

C

Collection of Proofs

This appendix contains proofs of lemmas, theorems, and corollaries that were deemed to long to be placed in the main body of the monograph.

C.1 Proofs from Section 4

We report below the proofs from Section 4.

C.1.1 Proof of Lemma 4.2

To prove this result, it is enough to show that the Hessian matrix $D^2\mathsf{NMSE}\left(\boldsymbol{\lambda}\right)$ is negative definite for $\lambda_n \geq 0$. $D^2\mathsf{NMSE}\left(\boldsymbol{\lambda}\right)$ is given as

$$D^2\mathsf{NMSE}\left(\boldsymbol{\lambda}\right) = \frac{1}{N\beta}\mathrm{diag}\left(\frac{-2\eta\tau_p\sigma_{\mathrm{ul}}^2}{\left(\eta\tau_p\lambda_1 + \sigma_{\mathrm{ul}}^2\right)^3}, \ldots, \frac{-2\eta\tau_p\sigma_{\mathrm{ul}}^2}{\left(\eta\tau_p\lambda_N + \sigma_{\mathrm{ul}}^2\right)^3}\right) \quad \text{(C.1)}$$

and is negative definite since it is diagonal with strictly negative entries.

C.1.2 Proof of Lemma 4.3

Note that all the elements of $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}'$ are identical except the $(r-1)$th and the rth ones. Hence, the difference between $\mathsf{NMSE}\left(\boldsymbol{\lambda}\right)$ and $\mathsf{NMSE}\left(\boldsymbol{\lambda}'\right)$ results from the summation terms in (4.29) for $n \in \{r-1, r\}$. Using this property, $\mathsf{NMSE}\left(\boldsymbol{\lambda}\right) - \mathsf{NMSE}\left(\boldsymbol{\lambda}'\right)$ is

$$\frac{\eta\tau_p}{N\beta}\left(\frac{(\lambda_{r-1} + \lambda_r)^2}{\eta\tau_p(\lambda_{r-1} + \lambda_r) + \sigma_{\mathrm{ul}}^2} - \frac{\lambda_{r-1}^2}{\eta\tau_p\lambda_{r-1} + \sigma_{\mathrm{ul}}^2} - \frac{\lambda_r^2}{\eta\tau_p\lambda_r + \sigma_{\mathrm{ul}}^2}\right)$$

$$\overset{(a)}{=} \frac{1}{N\beta}\left(\frac{(x+y)^2}{x+y+c} - \frac{x^2}{x+c} - \frac{y^2}{y+c}\right)$$

$$= \frac{(x+y)^2\left(xy + c\left(x+y+c\right)\right) - \left(x^2\left(y+c\right) + y^2\left(x+c\right)\right)\left(x+y+c\right)}{N\beta\left(x+y+c\right)\left(x+c\right)\left(y+c\right)}$$

$$\overset{(b)}{=} \frac{(x+y)^2\,xy + 2cxy\,(x+y+c) - xy\,(x+y)\,(x+y+c)}{N\beta\,(x+y+c)\,(x+c)\,(y+c)}$$

$$\overset{(c)}{=} \frac{2cxy\,(x+y+c) - cxy\,(x+y)}{N\beta\,(x+y+c)\,(x+c)\,(y+c)} \overset{(d)}{>} 0 \tag{C.2}$$

where we have introduced $x = \lambda_{r-1}$, $y = \lambda_r$, and the constant $c = \sigma_{\mathrm{ul}}^2/(\eta\tau_p)$ in (a) for simplicity. In (b) and (c), we have canceled the common terms $(x^2 + y^2)c(x + y + c)$ and $(x + y)^2\,xy$, respectively, in the numerator. Finally, the result in (d) is obtained from the fact that $x > 0$, $y > 0$, and $c > 0$.

C.2 Proofs from Section 5

We report below the proofs from Section 5.

C.2.1 Proof of Theorem 5.1

The processed signal in (5.3) can be treated as the discrete memoryless interference channel in Lemma 3.5 on p. 86 with a random channel response $h = \mathbf{v}_k^{\mathrm{H}}\mathbf{D}_k\widehat{\mathbf{h}}_k$, the input $x = s_k$, the output $y = \mathbf{v}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{y}^{\mathrm{ul}}$, and the realization $u = \{\mathbf{D}_k\widehat{\mathbf{h}}_i : i = 1, \ldots, K\}$ that affects the conditional variance of the interference. The effective noise $\mathbf{v}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{n}$ may not be Gaussian and all the interference and noise are included in υ with $n = 0$ in Lemma 3.5 on p. 86. The input power is $p = \mathbb{E}\{|s_k|^2\} = p_k$. The term υ is given by

$$\upsilon = \sum_{\substack{i=1 \\ i\neq k}}^{K} \mathbf{v}_k^{\mathrm{H}}\mathbf{D}_k\widehat{\mathbf{h}}_i s_i + \sum_{i=1}^{K} \mathbf{v}_k^{\mathrm{H}}\mathbf{D}_k\widetilde{\mathbf{h}}_i s_i + \mathbf{v}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{n}. \tag{C.3}$$

To prove the theorem, we need to show that the requirements in Lemma 3.5 on p. 86 are satisfied and then compute the conditional variance $p_\upsilon(h, u) = \mathbb{E}\{|\upsilon|^2\,|h, u\}$. First, we notice that the realizations of h and u are known at the CPU and υ has conditional zero-mean given (h, u), i.e., $\mathbb{E}\{\upsilon|h, u\} = 0$ since the symbols $\{s_i : i = 1, \ldots, K\}$ and the noise vector $\mathbf{n}$ are independent of the channel estimates and zero-mean estimation errors. The conditional variance given (h, u) is

$$p_\upsilon(h, u) = \mathbb{E}\left\{|\upsilon|^2\,|h, u\right\} = \mathbb{E}\left\{|\upsilon|^2\,\Big|\,\left\{\mathbf{D}_k\widehat{\mathbf{h}}_i\right\}\right\}$$

$$
\begin{aligned}
&= \sum_{\substack{i=1 \\ i \neq k}}^{K} p_i \left| \mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \widehat{\mathbf{h}}_i \right|^2 + \sum_{i=1}^{K} p_i \mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \mathbb{E} \left\{ \widetilde{\mathbf{h}}_i \widetilde{\mathbf{h}}_i^{\mathrm{H}} \right\} \mathbf{D}_k \mathbf{v}_k \\
&\quad + \mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \mathbb{E} \{ \mathbf{n} \mathbf{n}^{\mathrm{H}} \} \mathbf{D}_k \mathbf{v}_k \\
&= \sum_{\substack{i=1 \\ i \neq k}}^{K} p_i \left| \mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \widehat{\mathbf{h}}_i \right|^2 + \sum_{i=1}^{K} p_i \mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \mathbf{C}_i \mathbf{D}_k \mathbf{v}_k + \sigma_{\mathrm{ul}}^2 \mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \mathbf{D}_k \mathbf{v}_k
\end{aligned}
$$

$$(C.4)$$

which is equal to the denominator of (5.5). In the derivation, we have used the fact that the individual terms of υ are mutually uncorrelated and the combining vector $\mathbf{v}_k$ depends only on the channel estimates, which are independent of the estimation errors. As a final requirement for using Lemma 3.5, we note that the input signal $x = s_k$ is conditionally uncorrelated with υ given (h, u), i.e., $\mathbb{E}\{s_k^{\star} \upsilon | \{\mathbf{D}_k \widehat{\mathbf{h}}_i\}\} = 0$ due to the independence of the different symbols and the zero-mean channel estimation errors.

As a last step, we note that only the fraction τ_u / τ_c of the samples is used for uplink data transmission, which results in the lower bound on the capacity that is stated in the theorem and measured in bit/s/Hz.

C.2.2   Proof of Theorem 5.2

By adding and subtracting the term $\mathbb{E}\{\mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \mathbf{h}_k\} s_k$, the received signal in (5.3) can alternatively be expressed as

$$\widehat{s}_k = \mathbb{E}\{\mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \mathbf{h}_k\} s_k + \upsilon_k \tag{C.5}$$

where

$$\upsilon_k = \left(\mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \mathbf{h}_k - \mathbb{E}\{\mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \mathbf{h}_k\} \right) s_k + \sum_{\substack{i=1 \\ i \neq k}}^{K} \mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \mathbf{h}_i s_i + \mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \mathbf{n}.$$

This is a deterministic channel (since $\mathbb{E}\{\mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \mathbf{h}_k\}$ is constant) with the additive interference plus noise term υ_k, which has zero mean since $\{s_i : i = 1, \ldots, K\}$ and $\mathbf{n}$ have zero mean. Although υ_k contains the desired signal s_k, it is uncorrelated with it since

$$\mathbb{E}\left\{ s_k^{\star} \upsilon_k \right\} = \underbrace{\mathbb{E}\left\{ (\mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \mathbf{h}_k - \mathbb{E}\{\mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \mathbf{h}_k\}) \right\}}_{=0} \mathbb{E}\{|s_k|^2\} = 0. \tag{C.6}$$

Therefore, we can apply Lemma 3.3 on p. 83 with $h = \mathbb{E}\{\mathbf{v}_k^H \mathbf{D}_k \mathbf{h}_k\}$, $x = s_k$, $p = p_k$, $v = v_k$, and $\sigma^2 = 0$. By noting that the signals of different UEs are independent and that the noise contributions at different APs are independent, we have that

$$\mathbb{E}\{|v_k|^2\} = \sum_{i=1}^{K} p_i \mathbb{E}\left\{|\mathbf{v}_k^H \mathbf{D}_k \mathbf{h}_i|^2\right\} - p_k \left|\mathbb{E}\left\{\mathbf{v}_k^H \mathbf{D}_k \mathbf{h}_k\right\}\right|^2 + \sigma_{ul}^2 \mathbb{E}\left\{\|\mathbf{D}_k \mathbf{v}_k\|^2\right\}. \tag{C.7}$$

The SE expression presented in the theorem then follows from Lemma 3.3. As a last step, we note that only the fraction τ_u/τ_c of the samples is used for uplink data transmission, which results in the lower bound on the capacity that is stated in the theorem and measured in bit/s/Hz.

C.2.3 Proof of Theorem 5.4

Since the CPU does not have knowledge of the channel estimates, it needs to treat the average channel gain $\mathbf{a}_k^H \mathbb{E}\{\mathbf{g}_{kk}\}$ as the true deterministic channel. Hence, the signal model is

$$\widehat{s}_k = \mathbf{a}_k^H \mathbb{E}\{\mathbf{g}_{kk}\} s_k + v_k \tag{C.8}$$

which is a deterministic channel with the additive interference-plus-noise term

$$v_k = (\mathbf{a}_k^H \mathbf{g}_{kk} - \mathbf{a}_k^H \mathbb{E}\{\mathbf{g}_{kk}\}) s_k + \sum_{\substack{i=1 \\ i \neq k}}^{K} \mathbf{a}_k^H \mathbf{g}_{ki} s_i + n_k'. \tag{C.9}$$

The term v_k has zero mean and is uncorrelated with the signal term in (C.8) since

$$\underbrace{\mathbb{E}\left\{\mathbf{a}_k^H \mathbf{g}_{kk} - \mathbf{a}_k^H \mathbb{E}\left\{\mathbf{g}_{kk}\right\}\right\}}_{=0} \mathbb{E}\left\{|s_k|^2\right\} = 0. \tag{C.10}$$

Therefore, we can apply Lemma 3.3 on p. 83 with $h = \mathbf{a}_k^H \mathbb{E}\{\mathbf{g}_{kk}\}$, $x = s_k$, $p = p_k$, $v = v_k$, and $\sigma^2 = 0$. By noting that the signals of different UEs are independent and that the received noise at different APs are independent, we have that

$$\mathbb{E}\{|v_k|^2\} = \sum_{i=1}^{K} p_i \mathbb{E}\{|\mathbf{a}_k^H \mathbf{g}_{ki}|^2\} - p_k \left|\mathbf{a}_k^H \mathbb{E}\{\mathbf{g}_{kk}\}\right|^2 + \mathbf{a}_k^H \mathbf{F}_k \mathbf{a}_k. \tag{C.11}$$

The SE expression presented in the theorem then follows from Lemma 3.3. As a last step, we note that only the fraction τ_u/τ_c of the samples is used for uplink data transmission, which results in the lower bound on the capacity that is stated in the theorem and measured in bit/s/Hz.

C.2.4   Proof of Corollary 5.6

The proof consists of a direct computation of the three types of expectations that appear in (5.26). We begin with

$$
\begin{aligned}
\left[\mathbb{E}\left\{\mathbf{g}_{ki}\right\}\right]_l &= \mathbb{E}\left\{\mathbf{v}_{kl}^{\mathsf{H}}\mathbf{D}_{kl}\mathbf{h}_{il}\right\} = \operatorname{tr}\left(\mathbf{D}_{kl}\mathbb{E}\left\{\widehat{\mathbf{h}}_{il}\widehat{\mathbf{h}}_{kl}^{\mathsf{H}}\right\}\right) \\
&= \begin{cases} \sqrt{\eta_k\eta_i}\tau_p \operatorname{tr}\left(\mathbf{D}_{kl}\mathbf{R}_{il}\boldsymbol{\Psi}_{t_kl}^{-1}\mathbf{R}_{kl}\right) & i \in \mathcal{P}_k \\ 0 & i \notin \mathcal{P}_k \end{cases}
\end{aligned}
\tag{C.12}
$$

where the second equality follows from the second identity of Lemma B.2 on p. 279 and the fact that $\widetilde{\mathbf{h}}_{il}$ and $\widehat{\mathbf{h}}_{kl}$ are independent. The third equality follows from (4.19) and gives the expression in (5.33). Similarly,

$$
\left[\mathbf{F}_k\right]_{ll} = \sigma_{\mathrm{ul}}^2\mathbb{E}\left\{\left\|\mathbf{D}_{kl}\mathbf{v}_{kl}\right\|^2\right\} = \sigma_{\mathrm{ul}}^2\operatorname{tr}\left(\mathbf{D}_{kl}\mathbb{E}\left\{\widehat{\mathbf{h}}_{kl}\widehat{\mathbf{h}}_{kl}^{\mathsf{H}}\right\}\right) = \sigma_{\mathrm{ul}}^2\left[\mathbb{E}\left\{\mathbf{g}_{kk}\right\}\right]_l
\tag{C.13}
$$

where we used the second identity from Lemma B.2 and then identify the expression of $\left[\mathbb{E}\left\{\mathbf{g}_{kk}\right\}\right]_l$. This gives us the expression in (5.35).

It remains to compute the elements of $\mathbb{E}\left\{\mathbf{g}_{ki}\mathbf{g}_{ki}^{\mathsf{H}}\right\}$. We observe that $\mathbb{E}\left\{\left[\mathbf{g}_{ki}\right]_l\left[\mathbf{g}_{ki}^{\star}\right]_r\right\} = \mathbb{E}\left\{\left[\mathbf{g}_{ki}\right]_l\right\}\mathbb{E}\left\{\left[\mathbf{g}_{ki}^{\star}\right]_r\right\}$ for $r \neq l$ due to the independence of the channels of different APs. Hence, it only remains to compute

$$
\left[\mathbb{E}\left\{\mathbf{g}_{ki}\mathbf{g}_{ki}^{\mathsf{H}}\right\}\right]_{ll} = \mathbb{E}\left\{\widehat{\mathbf{h}}_{kl}^{\mathsf{H}}\mathbf{D}_{kl}\mathbf{h}_{il}\mathbf{h}_{il}^{\mathsf{H}}\mathbf{D}_{kl}\widehat{\mathbf{h}}_{kl}\right\} = \operatorname{tr}\left(\mathbf{D}_{kl}\mathbb{E}\left\{\mathbf{h}_{il}\mathbf{h}_{il}^{\mathsf{H}}\mathbf{D}_{kl}\widehat{\mathbf{h}}_{kl}\widehat{\mathbf{h}}_{kl}^{\mathsf{H}}\right\}\right)
\tag{C.14}
$$

where we utilized the second identity from Lemma B.2. If $i \notin \mathcal{P}_k$, we can utilize the independence of $\widehat{\mathbf{h}}_{kl}$ and $\mathbf{h}_{il}$ to obtain

$$
\begin{aligned}
\operatorname{tr}\left(\mathbf{D}_{kl}\mathbb{E}\left\{\mathbf{h}_{il}\mathbf{h}_{il}^{\mathsf{H}}\mathbf{D}_{kl}\widehat{\mathbf{h}}_{kl}\widehat{\mathbf{h}}_{kl}^{\mathsf{H}}\right\}\right) &= \operatorname{tr}\left(\mathbf{D}_{kl}\mathbb{E}\left\{\mathbf{h}_{il}\mathbf{h}_{il}^{\mathsf{H}}\right\}\mathbf{D}_{kl}\mathbb{E}\left\{\widehat{\mathbf{h}}_{kl}\widehat{\mathbf{h}}_{kl}^{\mathsf{H}}\right\}\right) \\
&= \eta_k\tau_p\operatorname{tr}\left(\mathbf{D}_{kl}\mathbf{R}_{il}\mathbf{R}_{kl}\boldsymbol{\Psi}_{t_kl}^{-1}\mathbf{R}_{kl}\right)
\end{aligned}
\tag{C.15}
$$

where we (for brevity) omitted one $\mathbf{D}_{kl}$ term that does not affect the result.

If $i \in \mathcal{P}_k$, we notice that

$$\text{tr}\left(\mathbf{D}_{kl}\mathbb{E}\left\{\mathbf{h}_{il}\mathbf{h}_{il}^{\text{H}}\mathbf{D}_{kl}\widehat{\mathbf{h}}_{kl}\widehat{\mathbf{h}}_{kl}^{\text{H}}\right\}\right)$$
$$= \text{tr}\left(\mathbf{D}_{kl}\mathbb{E}\left\{\widehat{\mathbf{h}}_{il}\widehat{\mathbf{h}}_{il}^{\text{H}}\mathbf{D}_{kl}\widehat{\mathbf{h}}_{kl}\widehat{\mathbf{h}}_{kl}^{\text{H}}\right\}\right) + \text{tr}\left(\mathbf{D}_{kl}\mathbb{E}\left\{\widetilde{\mathbf{h}}_{il}\widetilde{\mathbf{h}}_{il}^{\text{H}}\mathbf{D}_{kl}\widehat{\mathbf{h}}_{kl}\widehat{\mathbf{h}}_{kl}^{\text{H}}\right\}\right)$$
$$\tag{C.16}$$

where the equality follows from separating $\mathbf{h}_{il}$ into its estimate and estimation error. The second term becomes $\eta_k\tau_p\text{tr}(\mathbf{D}_{kl}\mathbf{C}_{il}\mathbf{R}_{kl}\boldsymbol{\Psi}_{t_kl}^{-1}\mathbf{R}_{kl})$ by utilizing the independence of estimates and estimation error and omitting one $\mathbf{D}_{kl}$ term. The first term is computed by utilizing the result from (4.18) to rewrite the estimate as $\widehat{\mathbf{h}}_{il} = \sqrt{\frac{\eta_i}{\eta_k}}\mathbf{R}_{il}\mathbf{R}_{kl}^{-1}\widehat{\mathbf{h}}_{kl}$:

$$\text{tr}\left(\mathbf{D}_{kl}\mathbb{E}\left\{\widehat{\mathbf{h}}_{il}\widehat{\mathbf{h}}_{il}^{\text{H}}\mathbf{D}_{kl}\widehat{\mathbf{h}}_{kl}\widehat{\mathbf{h}}_{kl}^{\text{H}}\right\}\right)$$
$$= \frac{\eta_i}{\eta_k}\text{tr}\left(\mathbf{D}_{kl}\mathbb{E}\left\{\mathbf{R}_{il}(\mathbf{R}_{kl})^{-1}\widehat{\mathbf{h}}_{kl}\widehat{\mathbf{h}}_{kl}^{\text{H}}(\mathbf{R}_{kl})^{-1}\mathbf{R}_{il}\mathbf{D}_{kl}\widehat{\mathbf{h}}_{kl}\widehat{\mathbf{h}}_{kl}^{\text{H}}\right\}\right)$$
$$= \frac{\eta_i}{\eta_k}\mathbb{E}\{|\widehat{\mathbf{h}}_{kl}^{\text{H}}\mathbf{D}_{kl}\mathbf{R}_{il}(\mathbf{R}_{kl})^{-1}\widehat{\mathbf{h}}_{kl}|^2\}$$
$$= \eta_k\eta_i\tau_p^2\left|\text{tr}\left(\mathbf{D}_{kl}\mathbf{R}_{il}\boldsymbol{\Psi}_{t_kl}^{-1}\mathbf{R}_{kl}\right)\right|^2 + \eta_k\tau_p\text{tr}\left(\mathbf{D}_{kl}(\mathbf{R}_{il} - \mathbf{C}_{il})\mathbf{R}_{kl}\boldsymbol{\Psi}_{t_kl}^{-1}\mathbf{R}_{kl}\right)$$
$$\tag{C.17}$$

where the last step follows from Lemma B.5 on p. 280 and some algebra. By adding these two terms together, we finally obtain (5.34). Note that the proof holds even if $\mathbf{R}_{kl}$ is non-invertible because $\mathbf{R}_{kl}^{-1}\widehat{\mathbf{h}}_{kl} = \sqrt{\eta_k\tau_p}\mathbf{R}_{kl}^{-1}\mathbf{R}_{kl}\boldsymbol{\Psi}_{t_kl}^{-1}\mathbf{y}_{t_kl}^{\text{pilot}} = \sqrt{\eta_k\tau_p}\boldsymbol{\Psi}_{t_kl}^{-1}\mathbf{y}_{t_kl}^{\text{pilot}}$, where there is no inversion. We are only using the inversion to shorten the notation.

C.3 Proofs from Section 6

We report below the proofs from Section 6.

C.3.1 Proof of Theorem 6.1

Since the UE k only has knowledge of the average of the effective channel, $\mathbb{E}\{\mathbf{h}_k^{\text{H}}\mathbf{D}_k\mathbf{w}_k\}$, the received signal in (6.7) at UE k can be expressed as

$$y_k^{\text{dl}} = \mathbb{E}\{\mathbf{h}_k^{\text{H}}\mathbf{D}_k\mathbf{w}_k\}\varsigma_k + \upsilon_k + n_k \tag{C.18}$$

which is a deterministic channel with the additive noise n_k and the additive interference term

$$\upsilon_k = \left(\mathbf{h}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{w}_k - \mathbb{E}\left\{\mathbf{h}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{w}_k\right\}\right)\varsigma_k + \sum_{\substack{i=1 \\ i \neq k}}^{K}\mathbf{h}_k^{\mathrm{H}}\mathbf{D}_i\mathbf{w}_i\varsigma_i. \qquad (\text{C.19})$$

The υ_k term has zero mean (since ς_i has zero mean) and although it contains the desired signal ς_k, it is uncorrelated with it since

$$\mathbb{E}\left\{\varsigma_k^{\star}\upsilon_k\right\} = \underbrace{\mathbb{E}\left\{\left(\mathbf{h}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{w}_k - \mathbb{E}\left\{\mathbf{h}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{w}_k\right\}\right)\right\}}_{=0}\mathbb{E}\{|\varsigma_k|^2\} = 0. \qquad (\text{C.20})$$

Therefore, we can apply Lemma 3.3 on p. 83 with $h = \mathbb{E}\left\{\mathbf{h}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{w}_k\right\}$, $x = \varsigma_k$, $p = 1$, $\upsilon = \upsilon_k$, and $\sigma^2 = \sigma_{\mathrm{dl}}^2$. By noting that the signals of different UEs are independent, we have that

$$\mathbb{E}\{|\upsilon_k|^2\} = \sum_{i=1}^{K}\mathbb{E}\left\{|\mathbf{h}_k^{\mathrm{H}}\mathbf{D}_i\mathbf{w}_i|^2\right\} - |\mathbb{E}\left\{\mathbf{h}_k^{\mathrm{H}}\mathbf{D}_k\mathbf{w}_k\right\}|^2. \qquad (\text{C.21})$$

The SE expression presented in the theorem then follows from Lemma 3.3 on p. 83. As a last step, we note that only the fraction τ_d/τ_c of the samples is used for downlink data transmission, which results in the lower bound on the capacity that is stated in the theorem and measured in bit/s/Hz.

C.3.2   Proof of Theorem 6.2

Let $\gamma_k = \mathsf{SINR}_k^{(\mathrm{ul,c-UatF})}$ denote the value of the effective SINR in (5.9) for the uplink powers $\{p_i : i = 1, \ldots, K\}$ and combining vectors $\{\mathbf{D}_i\mathbf{v}_i : i = 1, \ldots, K\}$. We want to show that $\gamma_k = \mathsf{SINR}_k^{(\mathrm{dl,c})}$ is achievable in the downlink when (6.11) is satisfied for all i. Plugging (6.11) into (6.10) yields the following SINR constraints:

$$\gamma_k = \frac{\rho_k\left|\mathbb{E}\left\{\mathbf{h}_k^{\mathrm{H}}\dfrac{\mathbf{D}_k\mathbf{v}_k}{\sqrt{\mathbb{E}\{\|\mathbf{D}_k\mathbf{v}_k\|^2\}}}\right\}\right|^2}{\displaystyle\sum_{i=1}^{K}\rho_i\mathbb{E}\left\{\left|\mathbf{h}_k^{\mathrm{H}}\dfrac{\mathbf{D}_i\mathbf{v}_i}{\sqrt{\mathbb{E}\{\|\mathbf{D}_i\mathbf{v}_i\|^2\}}}\right|^2\right\} - \rho_k\left|\mathbb{E}\left\{\mathbf{h}_k^{\mathrm{H}}\dfrac{\mathbf{D}_k\mathbf{v}_k}{\sqrt{\mathbb{E}\{\|\mathbf{D}_k\mathbf{v}_k\|^2\}}}\right\}\right|^2 + \sigma_{\mathrm{dl}}^2}$$
$$(\text{C.22})$$

for $k = 1, \ldots, K$. We define the diagonal matrix $\boldsymbol{\Gamma} \in \mathbb{R}^{K \times K}$ with the kth diagonal element being

$$[\boldsymbol{\Gamma}]_{kk} = \frac{1}{\gamma_k} \left| \mathbb{E} \left\{ \mathbf{h}_k^{\mathrm{H}} \frac{\mathbf{D}_k \mathbf{v}_k}{\sqrt{\mathbb{E}\{\mathbf{v}_k^{\mathrm{H}} \mathbf{D}_k \mathbf{v}_k\}}} \right\} \right|^2 \tag{C.23}$$

and let $\boldsymbol{\Sigma} \in \mathbb{R}^{K \times K}$ be the matrix whose (k, i)th element is

$$[\boldsymbol{\Sigma}]_{ki} = \mathbb{E} \left\{ \left| \mathbf{h}_k^{\mathrm{H}} \frac{\mathbf{D}_i \mathbf{v}_i}{\sqrt{\mathbb{E}\{\mathbf{v}_i^{\mathrm{H}} \mathbf{D}_i \mathbf{v}_i\}}} \right|^2 \right\} - \begin{cases} 0 & i \neq k \\ \gamma_k [\boldsymbol{\Gamma}]_{kk} & i = k. \end{cases} \tag{C.24}$$

Using these matrices, the SINR constraint in (C.22) can be expressed as

$$[\boldsymbol{\Gamma}]_{kk} = \frac{\sum_{i=1}^{K} \rho_i [\boldsymbol{\Sigma}]_{ki} + \sigma_{\mathrm{dl}}^2}{\rho_k}. \tag{C.25}$$

By rearranging this equation, we obtain $\sigma_{\mathrm{dl}}^2 = \rho_k [\boldsymbol{\Gamma}]_{kk} - \sum_{i=1}^{K} \rho_i [\boldsymbol{\Sigma}]_{ki}$. The K constraints can be written in matrix form as $\mathbf{1}_K \sigma_{\mathrm{dl}}^2 = (\boldsymbol{\Gamma} - \boldsymbol{\Sigma}) \boldsymbol{\rho}$ with $\boldsymbol{\rho} = [\rho_1 \ldots \rho_K]^{\mathrm{T}}$ being the downlink transmit power vector. The SINR constraints in (C.22) are thus satisfied if

$$\boldsymbol{\rho} = (\boldsymbol{\Gamma} - \boldsymbol{\Sigma})^{-1} \mathbf{1}_K \sigma_{\mathrm{dl}}^2. \tag{C.26}$$

This is a feasible power if $\boldsymbol{\Gamma} - \boldsymbol{\Sigma}$ is invertible, which always holds when $\mathbf{p} = [p_1 \ldots p_K]^{\mathrm{T}}$ is feasible. To show this, we notice that the K uplink SINR conditions can be expressed in a similar form where $\boldsymbol{\Sigma}$ is replaced by $\boldsymbol{\Sigma}^{\mathrm{T}}$ such that $\mathbf{p} = (\boldsymbol{\Gamma} - \boldsymbol{\Sigma}^{\mathrm{T}})^{-1} \mathbf{1}_K \sigma_{\mathrm{ul}}^2$. Since the eigenvalues of $\boldsymbol{\Gamma} - \boldsymbol{\Sigma}$ and $\boldsymbol{\Gamma} - \boldsymbol{\Sigma}^{\mathrm{T}}$ are the same and the uplink SINR conditions are satisfied by assumption, we can always select the downlink powers according to (C.26). Substituting $\mathbf{1}_K = \frac{1}{\sigma_{\mathrm{ul}}^2} (\boldsymbol{\Gamma} - \boldsymbol{\Sigma}^{\mathrm{T}}) \mathbf{p}$ into (C.26) yields

$$\boldsymbol{\rho} = \frac{\sigma_{\mathrm{dl}}^2}{\sigma_{\mathrm{ul}}^2} (\boldsymbol{\Gamma} - \boldsymbol{\Sigma})^{-1} (\boldsymbol{\Gamma} - \boldsymbol{\Sigma}^{\mathrm{T}}) \mathbf{p}. \tag{C.27}$$

The total transmit power condition now follows from direct computation by noting that $\mathbf{1}_K^{\mathrm{T}} (\boldsymbol{\Gamma} - \boldsymbol{\Sigma})^{-1} \mathbf{1}_K = \mathbf{1}_K^{\mathrm{T}} (\boldsymbol{\Gamma} - \boldsymbol{\Sigma}^{\mathrm{T}})^{-1} \mathbf{1}_K$.

To complete the proof, we need to show the power allocation coefficients obtained by (C.27) are positive. Please see [33, Proof of Theorem 4.8] for the final technical details.

References

[1] 3GPP, *Further advancements for E-UTRA physical layer aspects (Release 9)*. 3GPP TS 36.814, Mar. 2017.

[2] A. Abdallah and M. M. Mansour, "Efficient angle-domain processing for FDD-based cell-free massive MIMO systems," pp. 2188–2203, *IEEE Trans. Commun.*, vol. 68, no. 4, 2020.

[3] F. Adachi, M. T. Feeney, J. D. Parsons, and A. G. Williamson, "Crosscorrelation between the envelopes of 900 MHz signals received at a mobile radio base station site," pp. 506–512, *IEE F, Commun., Radar and Signal Process.*, vol. 133, no. 6, 1986.

[4] A. Adhikary, A. Ashikhmin, and T. L. Marzetta, "Uplink interference reduction in large-scale antenna systems," pp. 2194–2206, *IEEE Trans. Commun.*, vol. 65, no. 5, 2017.

[5] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing–the large-scale array regime," pp. 6441–6463, *IEEE Trans. Inf. Theory*, vol. 59, no. 10, 2013.

[6] M. Alonzo, S. Buzzi, A. Zappone, and C. D'Elia, "Energy-efficient power control in cell-free and user-centric massive MIMO at millimeter wave," pp. 651–663, *IEEE Trans. Green Commun. Net.*, vol. 3, no. 3, 2019.

[7] S. Anderson, M. Millnert, M. Viberg, and B. Wahlberg, "An adaptive array for mobile communication systems," pp. 230–236, *IEEE Trans. Veh. Technol.*, vol. 40, no. 1, 1991.

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

[9]   V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region," pp. 3032–3050, *IEEE Trans. Inf. Theory*, vol. 55, no. 7, 2011.

[10] V. S. Annapureddy and V. V. Veeravalli, "Sum capacity of MIMO interference channels in the low interference regime," pp. 2565–2581, *IEEE Trans. Inf. Theory*, vol. 57, no. 5, 2011.

[11] A. Ashikhmin and T. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1137–1141, 2012.

[12] M. Z. Aslam, Y. Corre, E. Björnson, and E. G. Larsson, "Performance of a dense urban massive MIMO network from a simulated ray-based channel," *EURASIP J. Wireless Commun. Netw.*, no. 106, 2019.

[13] M. Attarifar, A. Abbasfar, and A. Lozano, "Random vs structured pilot assignment in cell-free Massive MIMO wireless networks," in *IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2018.

[14] M. Attarifar, A. Abbasfar, and A. Lozano, "Subset MMSE receivers for cell-free networks," pp. 4183–4194, *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, 2020.

[15] M. Bashar, A. Akbari, K. Cumanan, H. Q. Ngo, A. G. Burr, P. Xiao, M. Debbah, and J. Kittler, "Exploiting deep learning in limited-fronthaul cell-free Massive MIMO uplink," pp. 1678–1697, *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, 2020.

[16] M. Bashar, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo, "On the uplink max-min SINR of cell-free Massive MIMO systems," pp. 2021–2036, *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, 2019.

[17] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, M. Debbah, and P. Xiao, "Max–min rate of cell-free Massive MIMO uplink with optimal uniform quantization," pp. 6796–6815, *IEEE Trans. Commun.*, vol. 67, no. 10, 2019.

[18] M. Bashar, H. Q. Ngo, A. G. Burr, D. Maryopi, K. Cumanan, and E. G. Larsson, "On the performance of backhaul constrained cell-free Massive MIMO with linear receivers," in *Asilomar Conf. Signals Syst. Comput.*, pp. 624–628, Nov. 2018.

[19] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," in *Handbook of Antennas in Wireless Communications*, L. C. Godara, Ed., CRC Press, 2001.

[20] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," pp. 2619–2691, *IEEE Trans. Inf. Theory*, vol. 44, no. 6, 1998.

[21] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure," pp. 142–148, *IEEE Signal Process. Mag.*, vol. 31, no. 4, 2014.

[22] E. Björnson and P. Giselsson, "Two applications of deep learning in the physical layer of communication systems," pp. 134–140, *IEEE Signal Process. Mag.*, vol. 37, no. 5, 2020.

[23] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has unlimited capacity," pp. 574–590, *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, 2018.

[24] E. Björnson, N. Jaldén, M. Bengtsson, and B. Ottersten, "Optimality properties, distributed strategies, and measurement-based evaluation of coordinated multicell OFDMA transmission," pp. 6086–6101, *IEEE Trans. Signal Process.*, vol. 59, no. 12, 2011.

[25] E. Björnson and E. Jorswieck, "Optimal resource allocation in coordinated multi-cell systems," pp. 113–381, *Foundations and Trends® in Communications and Information Theory*, vol. 9, no. 2-3, 2013.

[26] E. Björnson, E. Jorswieck, M. Debbah, and B. Ottersten, "Multi-objective signal processing optimization: The way to balance conflicting metrics in 5G systems," pp. 14–23, *IEEE Signal Process. Mag.*, vol. 31, no. 6, 2014.

[27] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," pp. 114–123, *IEEE Commun. Mag.*, vol. 54, no. 2, Feb. 2016.

[28] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," pp. 182–186, *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, Jan. 2020.

[29] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," pp. 4247–4261, *IEEE Trans. Commun.*, vol. 68, no. 7, 2020.

[30] E. Björnson, L. Sanguinetti, and M. Debbah, "Massive MIMO with imperfect channel covariance information," in *Asilomar Conf. Signals Syst. Comput.*, pp. 974–978, Nov. 2016.

[31] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, "Massive MIMO is a reality—What is next? Five promising research directions for antenna arrays," pp. 3–20, *Digital Signal Processing*, vol. 94, Nov. 2019.

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

[33] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," pp. 154–655, *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, 2017.

[34] H. Boche and M. Schubert, "A general duality theory for uplink and downlink beamforming," in *IEEE Veh. Technol. Conf. (VTC-Fall)*, pp. 87–91, 2002.

[35] M. Boldi, A. Tölli, M. Olsson, E. Hardouin, T. Svensson, F. Boccardi, L. Thiele, and V. Jungnickel, "Coordinated multipoint (CoMP) systems," in *Mobile and Wireless Communications for IMT-Advanced and Beyond*, A. Osseiran, J. Monserrat, and W. Mohr, Eds., Wiley, 2011, pp. 121–155.

[36] M. N. Boroujerdi, A. Abbasfar, and M. Ghanbari, "Cell free Massive MIMO with limited capacity fronthaul," pp. 633–648, *Wireless Pers Commun*, vol. 104, 2019.

[37] S. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge University Press, 2004.

[38] K. Bullington, "Frequency economy in mobile radio bands," pp. 42–62, *The Bell System Technical Journal*, vol. 32, no. 1, Jan. 1953.

[39] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," Research Laboratory of Electronics, Massachusetts Institute of Technology, Tech. Rep. 216, 1952.

[40] S. Buzzi and C. D'Andrea, "Cell-free massive MIMO: User-centric approach," pp. 706–709, *IEEE Commun. Lett.*, vol. 6, no. 6, 2017.

[41] S. Buzzi and C. D'Andrea, "User-centric communications versus cell-free massive MIMO for 5G cellular networks," in *ITG Workshop on Smart Antennas (WSA)*, 2017.

[42] S. Buzzi, C. D'Andrea, M. Fresia, Y. .-. Zhang, and S. Feng, "Pilot assignment in cell-free massive MIMO based on the Hungarian algorithm," pp. 34–37, *IEEE Wireless Communications Letters*, vol. 10, no. 1, 2021.

[43] S. Buzzi, C. D'Andrea, A. Zappone, and C. D'Elia, "User-centric 5G cellular networks: Resource allocation and comparison with the cell-free massive MIMO approach," pp. 1250–1264, *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, 2020.

[44] G. Caire, "On the ergodic rate lower bounds with applications to massive MIMO," pp. 3258–3268, *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, 2018.

[45] E. D. Carvalho and D.T.M. Slock, "Cramer-Rao bounds for semi-blind, blind and training sequence based channel estimation," in *IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, pp. 129–132, 1997.

[46] E. Cayirci and C. Ersoy, "Application of 3G PCS technologies to rapidly deployable mobile networks," pp. 20–27, *IEEE Netw.*, vol. 16, no. 5, 2002.

[47] S. Chakraborty, E. Björnson, and L. Sanguinetti, "Centralized and distributed power allocation for max-min fairness in cell-free massive MIMO," in *Asilomar Conf. Signals Syst. Comput.*, pp. 576–580, 2019.

[48] S. Chakraborty, Ö. T. Demir, E. Björnson, and P. Giselsson, "Efficient downlink power allocation algorithms for cell-free massive MIMO systems," *IEEE Open J. Commun. Society*, to appear.

[49] S. Chen, F. Qin, B. Hu, X. Li, and Z. Chen, "User-centric ultra-dense networks for 5G: Challenges, methodologies, and directions," pp. 78–85, *IEEE Wireless Commun.*, vol. 23, no. 2, 2016.

[50] S. Chen, J. Zhang, E. Björnson, J. Zhang, and B. Ai, "Structured massive access for scalable cell-free massive MIMO systems," *IEEE J. Sel. Areas Commun.*, to appear.

[51] Z. Chen and E. Björnson, "Channel hardening and favorable propagation in cell-free massive MIMO with stochastic geometry," pp. 5205–5219, *IEEE Trans. Commun.*, vol. 17, no. 11, 2018.

[52] Z. Chen, E. Björnson, and E. G. Larsson, "Dynamic resource allocation in co-located and cell-free massive MIMO," pp. 209–220, *IEEE Trans. Green Commun. Net.*, vol. 4, no. 1, 2020.

[53] H. V. Cheng and E. G. Larsson, "Some fundamental limits on frequency synchronization in massive MIMO," in *Asilomar Conf. Signals Syst. Comput.*, pp. 1213–1217, 2013.

[54] W. Choi and J. G. Andrews, "Downlink performance and capacity of distributed antenna systems in a multicell environment," pp. 69–73, *IEEE Trans. Wireless Commun.*, vol. 6, no. 1, Jan. 2007.

[55] M. Cooper, "The myth of spectrum scarcity," DYNA llc, Tech. Rep., Mar. 2010. [Online]. Available: https://ecfsapi.fcc.gov/file/7020396128.pdf.

[56] T. M. Cover and J. A. Thomas, *Elements of information theory.* Wiley, 1991.

[57] Ö. T. Demir and E. Björnson, "The Bussgang decomposition of nonlinear systems: Basic theory and MIMO extensions [lecture notes]," pp. 131–136, *IEEE Signal Processing Magazine*, vol. 38, no. 1, 2021.

[58] Ö. T. Demir and E. Björnson, "Joint power control and LSFD for wireless-powered cell-free massive MIMO," *IEEE Trans. Wireless Commun.*, to appear.

[59]   Ö. T. Demir, E. Björnson, and L. Sanguinetti, "Cell-free massive MIMO with large-scale fading decoding and dynamic cooperation clustering," *Submitted for publication*, 2020.

[60]   Ericsson, *Ericsson mobility report*, Jun. 2020. [Online]. Available: http://www.ericsson.com/mobility-report.

[61]   I. Estella Aguerri, A. Zaidi, G. Caire, and S. Shamai, "On the capacity of cloud radio access networks with oblivious relaying," pp. 4575–4596, *IEEE Trans. Inf. Theory*, vol. 65, no. 7, 2019.

[62]   B. Etzlinger and H. Wymeersch, "Synchronization and localization in wireless networks," pp. 1–106, *Foundations and Trends® in Signal Processing*, vol. 12, no. 1, 2018.

[63]   W. Fan, J. Zhang, E. Björnson, S. Chen, and Z. Zhong, "Performance analysis of cell-free massive MIMO over spatially correlated fading channels," in *IEEE Int. Conf. Commun. (ICC)*, 2019.

[64]   M. Farooq, H. Q. Ngo, E.-K. Hong, and L.-N. Tran, "Utility maximization for large-scale cell-free massive MIMO downlink," *CoRR*, vol. abs/2009.07167, 2020. [Online]. Available: http://arxiv.org/abs/2009.07167.

[65]   G. Femenias and F. Riera-Palou, "Cell-free millimeter-wave massive MIMO systems with limited fronthaul capacity," pp. 44 596–44 612, *IEEE Access*, vol. 7, 2019.

[66]   G. J. Foschini, K. Karakayali, and R. A. Valenzuela, "Coordinating multiple antenna cellular networks to achieve enormous spectral efficiency," pp. 548–555, *IEE Proceedings - Communications*, vol. 153, no. 4, Aug. 2006.

[67]   R. H. Frefkiel, "A high-capacity mobile radiotelephone system model using a coordinated small-zone approach," pp. 173–177, *IEEE Trans. Veh. Technol.*, vol. 19, no. 2, 1970.

[68]   I. D. Garcia, N. Kusashima, K. Sakaguchi, and K. Araki, "Dynamic cooperation set clustering on base station cooperation cellular networks," in *IEEE Ann. Int. Symp. Pers., Indoor, and Mobile Radio Commun. (PIMRC)*, pp. 2127–2132, 2010.

[69]   L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource Allocation and Cross-Layer Control in Wireless Networks," pp. 1–144, *Foundations and Trends in Networking*, vol. 1, no. 1, 2006.

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

[71] M. Grant and S. Boyd, *CVX: Matlab software for disciplined convex programming, version 2.1*, http://cvxr.com/cvx, Mar. 2014.

[72] M. Guillaud, D. Slock, and R. Knopp, "A practical method for wireless channel reciprocity exploitation through relative calibration," in *Int. Symp. Signal Process. App. (ISSPA)*, pp. 403–406, 2005.

[73] S. Haghighatshoar and G. Caire, "Massive MIMO pilot decontamination and channel interpolation via wideband sparse channel estimation," pp. 8316–8332, *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, 2017.

[74] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," pp. 951–963, *IEEE Trans. Inf. Theory*, vol. 49, no. 4, 2003.

[75] Y.-W. P. Hong, C. W. Tan, L. Zheng, C. Hsieh, and C. Lee, "A unified framework for wireless max-min utility optimization with general monotonic constraints," in *IEEE Int. Conf. Comp. Commun. (INFOCOM)*, pp. 2076–2084, 2014.

[76] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?," pp. 160–171, *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, 2013.

[77] J. Hoydis, M. Kobayashi, and M. Debbah, "Green small-cell networks," pp. 37–43, *IEEE Veh. Technol. Mag.*, vol. 6, no. 1, 2011.

[78] X. Hu, C. Zhong, X. Chen, W. Xu, H. Lin, and Z. Zhang, "Cell-free Massive MIMO systems with low resolution ADCs," pp. 6844–6857, *IEEE Trans. Commun.*, vol. 67, no. 10, 2019.

[79] H. Huang, M. Trivellato, A. Hottinen, M. Shafi, P. Smith, and R. Valenzuela, "Increasing downlink cellular throughput with limited network MIMO coordination," pp. 2983–2989, *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, 2009.

[80] H. Huh, G. Caire, H. Papadopoulos, and S. Ramprashad, "Achieving "massive MIMO" spectral efficiency with a not-so-large number of antennas," pp. 3226–3239, *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, 2012.

[81] I. Hwang, B. Song, and S. S. Soliman, "A holistic view on hyperdense heterogeneous and small cell networks," pp. 20–27, *IEEE Commun. Mag.*, vol. 51, no. 6, 2013.

[82] C. I, J. Huang, R. Duan, C. Cui, J. Jiang, and L. Li, "Recent progress on C-RAN centralization and cloudification," pp. 1030–1039, *IEEE Access*, vol. 2, 2014.

[83] G. Interdonato, E. Björnson, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Ubiquitous cell-free massive MIMO communications," *EURASIP J. Wirel. Commun. Netw.*, vol. 2019, no. 197, 2019.

[84] G. Interdonato, P. Frenger, and E. G. Larsson, "Scalability aspects of cell-free massive MIMO," in *IEEE Int. Conf. Commun. (ICC)*, 2019.

[85] G. Interdonato, M. Karlsson, E. Björnson, and E. G. Larsson, "Local partial zero-forcing precoding for cell-free massive MIMO," pp. 4758–4774, *IEEE Transactions on Wireless Communications*, vol. 19, no. 7, 2020.

[86] G. Interdonato, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Downlink training in cell-free massive MIMO: A blessing in disguise," pp. 5153–5169, *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, 2019.

[87] G. Interdonato, H. Q. Ngo, E. G. Larsson, and P. Frenger, "On the performance of Cell-free Massive MIMO with short-term power constraints," in *IEEE Int. Workshop Computer Aided Modelling and Design of Commun. Links and Networks (CAMAD)*, pp. 225–230, 2016.

[88] S. A. Jafar, G. J. Foschini, and A. J. Goldsmith, "Phantom-Net: Exploring optimal multicellular multiple antenna systems," *EURASIP J. Wirel. Commun. Netw.*, no. 691857, 2004.

[89] S. Jeong, A. Farhang, F. Gao, and M. F. Flanagan, "Frequency synchronisation for massive MIMO: A survey," pp. 2639–2645, *IET Communications*, vol. 14, no. 16, 2020.

[90] Z. Jiang, A. F. Molisch, G. Caire, and Z. Niu, "Achievable rates of FDD Massive MIMO systems with spatial channel correlation," pp. 2868–2882, *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, 2015.

[91] Y. Jin, J. Zhang, S. Jin, and B. Ai, "Channel estimation for cell-free mmwave massive MIMO through deep learning," pp. 10 325–10 329, *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, 2019.

[92] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," pp. 2640–2651, *IEEE Trans. Commun.*, vol. 10, no. 8, 2011.

[93] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfrojd, and T. Svensson, "The role of small cells, coordinated multipoint, and massive MIMO in 5G," pp. 44–51, *IEEE Commun. Mag.*, vol. 52, no. 5, 2014.

[94] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-dense networks: A survey," pp. 2522–2545, *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, 2016.

[95] M. Karakayali, G. Foschini, and R. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," pp. 56–61, *IEEE Wireless Commun.*, vol. 13, no. 4, 2006.

[96] S. Kashyap, C. Mollén, E. Björnson, and E. G. Larsson, "Frequency-domain interpolation of the zero-forcing matrix in massive MIMO-OFDM," in *IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, 2016.

[97] S. Kaviani, O. Simeone, W. Krzymien, and S. Shamai, "Linear precoding and equalization for network MIMO with partial cooperation," pp. 2083–2095, *IEEE Trans. Veh. Technol.*, vol. 61, no. 5, 2012.

[98] S. M. Kay, *Fundamentals of statistical signal processing: Estimation theory.* Prentice Hall, 1993.

[99] S. Khattak, W. Rave, and G. Fettweis, "Distributed iterative multiuser detection through base station cooperation," *EURASIP J. Wirel. Commun. Netw.*, no. 390489, 2008.

[100] S. Kim, J. W. Choi, and B. Shim, "Downlink pilot precoding and compressed channel feedback for FDD-based cell-free systems," pp. 3658–3672, *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, 2020.

[101] S. Kim and B. Shim, "FDD-based cell-free massive MIMO systems," in *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2018.

[102] K. Kinoshita *et al.*, "History of mobile communications with a look back at NTT DOCOMO R&D and outlook for the future," pp. 4–15, *NTT DOCOMO Technical Journal*, vol. 20, 2018.

[103] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," pp. 186–195, *IEEE Commun. Mag.*, vol. 52, no. 2, 2014.

[104] X. Li, E. Björnson, S. Zhou, and J. Wang, "Massive MIMO with multi-antenna users: When are additional user antennas beneficial?" In *IEEE ICT*, 2016.

[105] H. Liu, J. Zhang, X. Zhang, A. Kurniawan, T. Juhana, and B. Ai, "Tabu-search-based pilot assignment for cell-free massive MIMO systems," pp. 2286–2290, *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, 2020.

[106] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," pp. 57–73, *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, 2008.

[107] J. Ma and L. Ping, "Data-aided channel estimation in large antenna systems," pp. 3111–3124, *IEEE Trans. Signal Process.*, vol. 62, no. 12, 2014.

[108] T. C. Mai, H. Q. Ngo, and T. Q. Duong, "Downlink spectral efficiency of cell-free massive MIMO systems with multi-antenna users," pp. 4803–4815, *IEEE Trans. Commun.*, vol. 68, no. 8, 2020.

[109] T. C. Mai, H. Q. Ngo, M. Egan, and T. Q. Duong, "Pilot power control for cell-free massive MIMO," pp. 11 264–11 268, *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, 2018.

[110] P. Marsch, S. Brück, A. Garavaglia, M. Schulist, R. Weber, and A. Dekorsy, "Clustering," in *Coordinated multi-point in mobile communications: From theory to practice*, P. Marsch and G. Fettweis, Eds., Cambridge, 2011, ch. 7, pp. 139–159.

[111] P. Marsch and G. Fettweis, "On multicell cooperative transmission in backhaul-constrained cellular systems," pp. 253–269, *Ann. Telecommun.*, vol. 63, 2008.

[112] D. Maryopi, M. Bashar, and A. Burr, "On the uplink throughput of zero forcing in cell-free Massive MIMO with coarse quantization," pp. 7220–7224, *IEEE Trans. Veh. Technol.*, vol. 68, no. 7, 2019.

[113] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," pp. 3590–3600, *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, 2010.

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

[115] H. Masoumi and M. J. Emadi, "Joint pilot and data power control in cell-free Massive MIMO system," in *Int. Conf. Millimeter-Wave and Terahertz Technologies (MMWaTT)*, pp. 34–37, 2018.

[116] H. Masoumi and M. J. Emadi, "Performance analysis of cell-free Massive MIMO system with limited fronthaul capacity and hardware impairments," pp. 1038–1053, *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, 2020.

[117] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," pp. 933–946, *IEEE Trans. Inf. Theory*, vol. 46, no. 3, 2000.

[118] A. F. Molisch, *Wireless communications*. John Wiley & Sons, 2007.

[119] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," pp. 620–643, *IEEE Trans. Inf. Theory*, vol. 55, no. 2, 2009.

[120] R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," pp. 773–786, *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, 2014.

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

[122] E. Nayebi, A. Ashikhmin, T. L. Marzetta, and B. D. Rao, "Performance of cell-free massive MIMO systems with MMSE and LSFD receivers," in *Asilomar Conf. Signals Syst. Comput.*, pp. 203–207, 2016.

[123] D. Neumann, M. Joham, and W. Utschick, "Suppression of pilot-contamination in massive MIMO systems," in *IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, pp. 11–15, 2014.

[124] D. Neumann, M. Joham, and W. Utschick, "On MSE based receiver design for Massive MIMO," in *Int. ITG Conf. Systems Commun. Coding (SCC)*, 2017.

[125] D. Neumann, M. Joham, and W. Utschick, "Covariance matrix estimation in massive MIMO," pp. 863–867, *IEEE Signal Process. Lett.*, vol. 25, no. 6, Jun. 2018.

[126] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free Massive MIMO versus small cells," pp. 1834–1850, *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, 2017.

[127] H. Q. Ngo, A. E. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO: Uniformly great service for everyone," in *IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, pp. 201–205, 2015.

[128] H. Q. Ngo and E. Larsson, "EVD-based channel estimations for multicell multiuser MIMO with very large antenna arrays," in *IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, 2012.

[129] H. Q. Ngo and E. G. Larsson, "No downlink pilots are needed in TDD Massive MIMO," pp. 2921–2935, *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, 2017.

[130] H. Q. Ngo, H. Tataria, M. Matthaiou, S. Jin, and E. G. Larsson, "On the performance of cell-free massive MIMO in Ricean fading," in *Asilomar Conf. Signals Syst. Comput.*, pp. 980–984, Nov. 2018.

[131] H. Q. Ngo, L. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," pp. 25–39, *IEEE Trans. Green Commun. Net.*, vol. 2, no. 1, 2018.

[132] L. D. Nguyen, T. Q. Duong, H. Q. Ngo, and K. Tourki, "Energy efficiency in cell-free massive MIMO with zero-forcing precoding design," pp. 1871–1874, *IEEE Commun. Lett.*, vol. 21, no. 8, 2017.

[133] J. Niemela and J. Lempiainen, "Mitigation of pilot pollution through base station antenna configuration in WCDMA," in *IEEE VTC-Fall*, vol. 6, pp. 4270–4274, 2004.

[134] R. Nikbakht, A. Jonsson, and A. Lozano, "Unsupervised learning for C-RAN power control and power allocation," *IEEE Commun. Lett.*, to appear.

[135] R. Nikbakht and A. Lozano, "Uplink fractional power control for cell-free wireless networks," in *IEEE Int. Conf. Commun. (ICC)*, 2019.

[136] R. Nikbakht, R. Mosayebi, and A. Lozano, "Uplink fractional power control and downlink power allocation for cell-free networks," pp. 774–777, *IEEE Wireless Commun. Lett.*, vol. 9, no. 6, 2020.

[137] K. Nishimori, K. Cho, Y. Takatori, and T. Hori, "Automatic calibration method using transmitting signals of an adaptive array for TDD systems," pp. 1636–1640, *IEEE Trans. Veh. Technol.*, vol. 50, no. 6, 2001.

[138] A. Osseiran, J. F. Monserrat, and P. Marsch, "5G mobile and wireless communications technology," in ch. 9, Coordinated Multi-Point Transmission in 5G, Cambridge: Cambridge University Press, 2016.

[139] Ö. Özdogan, E. Björnson, and J. Zhang, "Cell-free massive MIMO with Rician fading: Estimation schemes and spectral efficiency," in *Asilomar Conf. Signals Syst. Comput.*, pp. 975–979, Nov. 2018.

[140] Ö. Özdogan, E. Björnson, and J. Zhang, "Performance of cell-free massive MIMO with Rician fading and phase shifts," pp. 5299–5315, *IEEE Trans. Commun.*, vol. 18, no. 11, 2019.

[141]	A. Papadogiannis, E. Hardouin, and D. Gesbert, "Decentralising multicell cooperative processing: A novel robust framework," *EURASIP J. Wirel. Commun. Netw.*, no. 890685, 2009.

[142]	P. Parida, H. S. Dhillon, and A. F. Molisch, "Downlink performance analysis of cell-free massive MIMO with finite fronthaul capacity," in *IEEE Veh. Technol. Conf. (VTC-Spring)*, 2018.

[143]	S. Park, O. Simeone, O. Sahin, and S. Shamai, "Robust and efficient distributed compression for cloud radio access networks," pp. 692–703, *IEEE Trans. Veh. Technol.*, vol. 62, no. 2, 2013.

[144]	S. Parkvall, E. Dahlman, A. Furuskär, Y. Jading, M. Olsson, S. Wänstedt, and K. Zangi, "LTE-advanced - evolving LTE towards IMT-advanced," in *IEEE Veh. Technol. Conf. (VTC-Fall)*, 2008.

[145]	K. I. Pedersen, P. E. Mogensen, and B. H. Fleury, "Power azimuth spectrum in outdoor environments," pp. 1583–1584, *Electronics Lett.*, vol. 33, no. 18, 1997.

[146]	M. Peng, Y. Sun, X. Li, Z. Mao, and C. Wang, "Recent advances in cloud radio access networks: System architectures, key techniques, and open issues," pp. 2282–2308, *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, 2016.

[147]	S. Perlman and A. Forenza, *An introduction to pCell*, 2015. [Online]. Available: http://www.rearden.com/artemis/An-Introduction-to-pCell-White-Paper-150224.pdf.

[148]	A. Pizzo, T. L. Marzetta, and L. Sanguinetti, "Spatially-stationary model for holographic MIMO small-scale fading," pp. 1964–1979, *IEEE J. Sel. Areas Commun.*, vol. 38, no. 9, 2020.

[149]	F. Riera-Palou, G. Femenias, A. G. Armada, and A. Pérez-Neira, "Clustered cell-free massive MIMO," in *IEEE Glob. Comm. Workshops (GC Wkshps)*, 2018.

[150]	R. Rogalin, O. Y. Bursalioglu, H. Papadopoulos, G. Caire, A. F. Molisch, A. Michaloliakos, V. Balan, and K. Psounis, "Scalable synchronization and reciprocity calibration for distributed multiuser MIMO," pp. 1815–1831, *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, 2014.

[151]	R. H. Roy and B. Ottersten, *Spatial division multiple access wireless communication systems*, US Patent, 1991.

[152] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," pp. 40–60, *IEEE Signal Process. Mag.*, vol. 30, no. 1, 2013.

[153] R. Sabbagh, C. Pan, and J. Wang, "Pilot allocation and sum-rate analysis in cell-free massive MIMO systems," in *IEEE Int. Conf. Commun. (ICC)*, May 2018.

[154] J. Salz and J. H. Winters, "Effect of fading correlation on adaptive arrays in digital mobile radio," pp. 1049–1057, *IEEE Trans. Veh. Technol.*, vol. 43, no. 4, 1994.

[155] A. Sanderovich, O. Somekh, H. V. Poor, and S. Shamai, "Uplink macro diversity of limited backhaul cellular network," pp. 3457–3478, *IEEE Trans. Inf. Theory*, vol. 55, no. 8, Aug. 2009.

[156] L. Sanguinetti, E. Björnson, and J. Hoydis, "Towards massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination," pp. 232–257, *IEEE Trans. Commun.*, vol. 68, no. 1, Jan. 2020.

[157] L. Sanguinetti, M. Morelli, and L. Marchetti, "A random access algorithm for LTE systems," *Trans. Emerging Telecommun. Technol.*, vol. 24, no. 1, 2013.

[158] F. Sapienza and S. Kim, "Dominant pilot recovery in IS-95 CDMA systems using repeaters," pp. 134–137, *IEICE Trans. Commun.*, vol. 84, no. 1, 2001.

[159] T. Schenk, *RF imperfections in high-rate wireless systems: Impact and digital compensation.* Springer, 2008.

[160] H. J. Schulte and W. A. Cornell, "Multi-area mobile telephone system," pp. 49–53, *IEEE Trans. Veh. Technol.*, vol. 9, no. 1, 1960.

[161] S. Sesia, I. Toufik, and M. Baker, *LTE - the UMTS long term evolution: from theory to practice.* Chichester: Wiley, 2009.

[162] Z. H. Shaik, E. Björnson, and E. G. Larsson, "Cell-free massive MIMO with radio stripes and sequential uplink processing," in *IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, 2020.

[163] Z. H. Shaik, E. Björnson, and E. G. Larsson, "MMSE-optimal sequential processing for cell-free massive MIMO with radio stripes," *Submitted for publication*, 2021.

[164] S. Shamai and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *IEEE Veh. Technol. Conf. (VTC-Spring)*, vol. 3, pp. 1745–1749, 2001.

[165] X. Shang, B. Chen, G. Kramer, and H. V. Poor, "Noisy-interference sum-rate capacity of parallel Gaussian interference channels," pp. 210–226, *IEEE Trans. Inf. Theory*, vol. 57, no. 1, 2011.

[166] X. Shang, G. Kramer, B. Chen, and H. V. Poor, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels," pp. 689–699, *IEEE Trans. Inf. Theory*, vol. 55, no. 2, 2009.

[167] C. E. Shannon, "A mathematical theory of communication," pp. 379–423, 623–656, *Bell System Technical Journal*, vol. 27, 1948.

[168] C. E. Shannon, "Communication in the presence of noise," pp. 10–21, *IRE*, vol. 37, no. 1, 1949.

[169] C. Shepard, H. Yu, N. Anand, L. Li, T. Marzetta, R. Yang, and L. Zhong, "Argos: Practical many-antenna base stations," in *ACM Int. Conf. Mobile Comp. Netw. (MobiCom)*, 2012.

[170] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," pp. 4331–4340, *IEEE Trans. Signal Process.*, vol. 59, no. 9, 2011.

[171] H. Shirani-Mehr, G. Caire, and M. Neely, "MIMO downlink scheduling with non-perfect channel state knowledge," pp. 2055–2066, *IEEE Trans. Commun.*, vol. 58, no. 7, 2010.

[172] D. Shiu, G. Foschini, M. Gans, and J. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," pp. 502–513, *IEEE Trans. Commun.*, vol. 48, no. 3, 2000.

[173] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai, "Downlink multicell processing with limited-backhaul capacity," *EURASIP J. Adv. Signal Process.*, no. 840814, 2009.

[174] A. Simonsson and A. Furuskär, "Uplink power control in LTE - overview and performance, subtitle: Principles and benefits of utilizing rather than compensating for SINR variations," in *IEEE Veh. Technol. Conf. (VTC)*, 2008.

[175] S. Stefanatos and A. Alexiou, "Access point density and bandwidth partitioning in ultra dense wireless networks," pp. 3376–3384, *IEEE Trans. Commun.*, vol. 62, no. 9, 2014.

[176] S. C. Swales, M. A. Beach, D. J. Edwards, and J. P. McGeehan, "The performance enhancement of multibeam adaptive base-station antennas for cellular land mobile radio systems," pp. 56–67, *IEEE Trans. Veh. Technol.*, vol. 39, no. 1, 1990.

[177] X. Tao, J. Xu, X. Xu, C. Tang, and P. Zhang, "Group cell FuTURE B3G TDD system," in *IEEE Ann. Int. Symp. Pers., Indoor, and Mobile Radio Commun. (PIMRC)*, vol. 2, pp. 967–971, 2005.

[178] A. Tölli, M. Codreanu, and M. Juntti, "Cooperative MIMO-OFDM cellular system with soft handover between distributed base station antennas," pp. 1428–1440, *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, 2008.

[179] T. Trump and B. Ottersten, "Estimation of nominal direction of arrival and angular spread using an array of sensors," pp. 57–69, *Signal Processing*, vol. 50, no. 1-2, 1996.

[180] D. Tse and P. Viswanath, *Fundamentals of wireless communications*. Cambridge University Press, 2005.

[181] R. Tütüncü, K. Toh, and M. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," pp. 189–217, *Mathematical Programming*, vol. 95, no. 2, 2003.

[182] K. Upadhya and S. A. Vorobyov, "Covariance matrix estimation for massive MIMO," pp. 546–550, *IEEE Signal Process. Lett.*, vol. 25, no. 4, Apr. 2018.

[183] K. Upadhya, S. A. Vorobyov, and M. Vehkapera, "Superimposed pilots are superior for mitigating pilot contamination in massive MIMO," pp. 2917–2932, *IEEE Trans. Signal Process.*, vol. 65, no. 11, 2017.

[184] K. Upadhya, S. A. Vorobyov, and M. Vehkaperä, "Downlink performance of superimposed pilots in massive MIMO systems," pp. 6630–6644, *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, 2018.

[185] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network MIMO: Overcoming intercell interference in indoor wireless systems," in *Asilomar Conf. Signals Syst. Comput.*, pp. 83–87, 2007.

[186] D. Verenzuela, E. Björnson, and L. Sanguinetti, "Spectral and energy efficiency of superimposed pilots in uplink Massive MIMO," pp. 7099–7115, *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, 2018.

[187] D. Verenzuela, E. Björnson, X. Wang, M. Arnold, and S. ten Brink, "Massive-MIMO iterative channel estimation and decoding (MICED) in the uplink," pp. 854–870, *IEEE Trans. Commun.*, vol. 68, no. 2, 2020.

[188] J. Vieira, F. Rusek, O. Edfors, S. Malkowsky, L. Liu, and F. Tufvesson, "Reciprocity calibration for Massive MIMO: Proposal, modeling, and validation," pp. 3042–3056, *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, 2017.

[189] J. Vieira, R. Rusek, and F. Tufvesson, "Reciprocity calibration methods for Massive MIMO based on antenna coupling," in *IEEE Glob. Comm. Conf. (GLOBECOM)*, pp. 3708–3712, 2014.

[190] J. Vinogradova, E. Björnson, and E. G. Larsson, "On the separability of signal and interference-plus-noise subspaces in blind pilot decontamination," in *IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, 2016.

[191] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," pp. 1912–1921, *IEEE Trans. Inf. Theory*, vol. 49, no. 8, 2003.

[192] A. J. Viterbi, A. M. Viterbi, K. S. Gilhousen, and E. Zehavi, "Soft handoff extends CDMA cell coverage and increases reverse link capacity," pp. 1281–1288, *IEEE J. Sel. Areas Commun.*, vol. 12, no. 8, Oct. 1994.

[193] B. Vojčić and W. Jang, "Transmitter precoding in synchronous multiuser communications," pp. 1346–1355, *IEEE Trans. Commun.*, vol. 46, no. 10, 1998.

[194] P. Weeraddana, M. Codreanu, M. Latva-aho, A. Ephremides, and C. Fischione, "Weighted sum-rate maximization in wireless networks: A review," pp. 1–163, *Foundations and Trends in Networking*, vol. 6, no. 1-2, 2012.

[195] J. F. Whitehead, "Signal-level-based dynamic power control for co-channel interference management," in *IEEE Vehicular Technology Conference (VTC)*, pp. 499–502, 1993.

[196] J. H. Winters, "Optimum combining for indoor radio systems with multiple users," pp. 1222–1230, *IEEE Trans. Commun.*, vol. 35, no. 11, 1987.

[197] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," pp. 1713–1727, *IEEE Trans. Inf. Theory*, vol. 40, no. 6, 1994.

[198] X. Xu, D. Wang, X. Tao, and T. Svensson, "Resource pooling for frameless network architecture with adaptive resource allocation," pp. 1–12, *Science China Information Sciences*, vol. 56, 2013.

[199] H. Yan, A. Ashikhmin, and H. Yang, "Optimally supporting IoT with cell-free massive MIMO," in *IEEE Global Communications Conference (GLOBECOM)*, 2020.

[200] H. Yang and E. G. Larsson, "Can massive MIMO support uplink intensive applications?" In *IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2019.

[201] H. Yang and T. L. Marzetta, "Capacity performance of multicell large-scale antenna systems," in *IEEE Allerton Conf. Commun., Control and Comp.*, pp. 668–675, 2013.

[202] H. Yin, L. Cottatellucci, D. Gesbert, R. R. Müller, and G. He, "Robust pilot decontamination based on joint angle and power domain discrimination," pp. 2990–3003, *IEEE Trans. Signal Process.*, vol. 64, no. 11, 2016.

[203] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," pp. 264–273, *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, 2013.

[204] P. Zetterberg, "Experimental investigation of TDD reciprocity-based zero-forcing transmit precoding," *EURASIP J. Adv. Signal Process.*, no. 137541, 2011.

[205] P. Zetterberg and B. Ottersten, "The spectrum efficiency of a base station antenna array system for spatially selective transmission," pp. 651–660, *IEEE Trans. Veh. Technol.*, vol. 44, no. 3, 1995.

[206] H. Zhang and H. Dai, "Cochannel interference mitigation and cooperative processing in downlink multicell multiuser MIMO networks," pp. 222–235, *EURASIP J. Wirel. Commun. Netw.*, vol. 2, 2004.

[207] H. Zhang, N. Mehta, A. Molisch, J. Zhang, and H. Dai, "Asynchronous interference mitigation in cooperative base station systems," pp. 155–165, *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, 2008.

[208] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, "Prospective multiple antenna technologies for beyond 5G," pp. 1637–1660, *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, 2020.

[209] J. Zhang, R. Chen, J. G. Andrews, A. Ghosh, and R. Heath, "Networked MIMO with clustered linear precoding," pp. 1910–1921, *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, 2009.

[210] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, and L. Hanzo, "Cell-free massive MIMO: A new next-generation paradigm," pp. 99 878–99 888, *IEEE Access*, vol. 7, 2019.

[211] J. Zhang, Y. Wei, E. Björnson, Y. Han, and S. Jin, "Performance analysis and power control of cell-free massive MIMO systems with hardware impairments," pp. 55 302–55 314, *IEEE Access*, vol. 6, 2018.

[212] P. Zhang, X. Tao, J. Zhang, Y. Wang, L. Li, and Y. Wang, "A vision from the future: Beyond 3G TDD," pp. 38–44, *IEEE Commun. Mag.*, vol. 43, no. 1, 2005.

[213] Y. Zhang, H. Cao, P. Zhong, C. Qi, and L. Yang, "Location-based greedy pilot assignment for cell-free massive MIMO systems," in *Int. Conf. Comp. Commun. (ICCC)*, pp. 392–396, 2018.

[214] Y. Zhang, M. Zhou, H. Cao, L. Yang, and H. Zhu, "On the performance of cell-free massive MIMO with mixed-ADC under Rician fading channels," pp. 43–47, *IEEE Commun. Lett.*, vol. 24, no. 1, 2020.

[215]   Y. Zhao, I. G. Niemegeers, and S. H. De Groot, "Power allocation in cell-free massive MIMO: A deep learning method," pp. 87 185–87 200, *IEEE Access*, vol. 8, 2020.

[216]   J. Zheng, J. Zhang, L. Zhang, X. Zhang, and B. Ai, "Efficient receiver design for uplink cell-free Massive MIMO with hardware impairments," pp. 4537–4541, *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, 2020.

[217]   L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," pp. 359–383, *IEEE Trans. Inf. Theory*, vol. 48, no. 2, 2002.

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

[219]   X. Zhu, Z. Wang, L. Dai, and C. Qian, "Smart pilot assignment for Massive MIMO," pp. 1644–1647, *IEEE Commun. Lett.*, vol. 19, no. 9, 2015.