

Article

# A Left-Null-Space-Based Massive Access Method for Cell-Free Massive MIMO IoT Systems

Yang Wang  and Die Hu \* 

Department of Communication Science and Engineering, Fudan University, Shanghai 200082, China; 20210720080@fudan.edu.cn

\* Correspondence: hudie@fudan.edu.cn

**Abstract:** Cell-free massive multiple-input multiple-output (MIMO) is a promising technology for the Internet of Things (IoT) because it can increase connectivity and provide considerable energy efficiency (EE) and spectral efficiency (SE). However, pilot contamination caused by pilot reuse severely limits the performance of the system. In this paper, we propose a left-null-space-based massive access method that can significantly reduce interference among users. The proposed method includes three stages: initial orthogonal access, left-null-space-based opportunistic access and data detection of all accessed users. The simulation results show that the proposed method can achieve much higher spectral efficiency than the existing massive access methods.

**Keywords:** cell-free massive multi-input multi-output; MIMO; massive access; interference; pilot contamination; left-null-space; spectral efficiency

## 1. Introduction

The Internet of Things (IoT) is envisioned as a promising technology, where highly connected users can share information and coordinate decisions [1]. It has led to revolutionary applications in transportation, smart healthcare, environmental monitoring, smart homes and other scenarios [2]. However, as the number of users connected to the IoT rapidly increases, hyperconnectivity, low latency, high energy efficiency (EE) and spectral efficiency (SE) become new challenges to solve [3]. Cell-free massive multiple-input multiple-output (MIMO) has been proposed in [4] and can meet these requirements. By deploying many access points (APs) in a geographic coverage area and connecting them to a central processing unit (CPU), cell-free massive MIMO can accommodate more users due to coverage probability improvement and interference management [5]. Since the boundaries have disappeared, cell-free massive MIMO can offer much more uniform connectivity, huge energy efficiency and spectral efficiency [6–8].

Channel state information (CSI) is important for cell-free massive MIMO systems [9]. However, pilot resources need to be reused among users due to the limitation of coherence duration time, which can lead to the so-called pilot contamination [10]. This phenomenon will significantly reduce channel estimation accuracy and degrade system performance [11]. There are some strategies that have been proposed to limit these negative effects, such as a pilot-precoding-based scheme [12], a channel-estimation-based scheme [13] and a pilot-assignment-based scheme [14]. Although those pilot contamination suppression schemes have achieved certain effects, there is still room for improvement to achieve higher performance of the system, especially in massive access scenarios.

Based on the user density in the network, massive access can be divided into random access and structured access [15]. In [16], random access is classified into two types: contention-based and contention-free. Contention-based random access allows each accessing user to pick a preamble at random from a predefined set and then transmit it to the APs. This method performs well in highly crowded scenarios [17]. In contrast, contention-free



**Citation:** Wang, Y.; Hu, D. A Left-Null-Space-Based Massive Access Method for Cell-Free Massive MIMO IoT Systems. *Sensors* **2023**, *23*, 5285. <https://doi.org/10.3390/s23115285>

Academic Editor: Stefano Savazzi

Received: 6 April 2023

Revised: 16 May 2023

Accepted: 30 May 2023

Published: 2 June 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

random access assigns a dedicated preamble to each user. Therefore, it can eliminate access delays. In [18], the authors improved the random access performance by using the strongest-user collision resolution (SUCR), where only the users with the strongest channel gains are allowed to access the network. In [19], the authors viewed a set of contaminated pilot signals as a graph code on which iterative belief propagation could be performed. This makes it possible to reduce pilot contamination. In [20], the system performance was improved by the averaging of the pilot collision events across the transmission slots for random access. However, if two users in close proximity of each other use the same pilot sequence, the random access usually causes strong pilot contamination and then deteriorates the performance of channel estimation.

A proper pilot assignment can also suppress pilot contamination effects and thus improve the system's performance in the massive access scenario. Structured access allocates a dedicated pilot resource to every user [21]; thus, contention-free random access can be regarded as a special type of structured access. Structured access is preferable in scenarios where the number of pilots is smaller than the number of users, except in extremely crowded scenarios, similar to the IoT networks [22]. The reason is that the strong pilot contamination might significantly reduce the system's performance in extremely crowded scenarios. In [15], a structured access framework that focuses on improving pilot assignment schemes was proposed to suppress pilot contamination. A greedy pilot assignment was proposed in [4] that iteratively refines the pilot assignment for the lowest downlink rate user. The authors in [15] proposed a user-group pilot assignment that aims to assign mutually orthogonal pilots to the users served by similar subsets of APs. A K-means-type pilot assignment was proposed in [23] that aimed at maximizing the minimum distance among the pilot sharing users. In [24], a weighted graph was constructed to represent the relationship between the potential strength of pilot contamination between users, which can mitigate pilot contamination by assigning different pilots to connected users with a large weight in a greedy way. The authors of [25] introduced a fractional pilot reuse scheme, where a fraction of users within each cell reuse the same pilot subset across the whole system, while the rest are allocated orthogonal subsets depending on a reuse parameter. However, since all users are allowed to access the network, the pilot assignment method will be less effective due to the limited resources in the crowded scenario.

To reduce inter-user interference, we propose a left-null-space-based massive access method that takes advantage of the resources in the space domain. The idea of reducing interference based on null space has been widely used in the past, especially for the purpose of spatial multiplexing. In [26], a partial nulling scheme was adopted to obtain diversity gains by utilizing degrees of freedom. In addition, Ref. [27] proposed a null-space expansion scheme to mitigate inter-user interference by exploiting the excess degrees of freedom to extend the null-space dimension, which are provided by the massive base station (BS) antenna array. However, most works only focus on the downlink transmission. Specifically, with the given channel matrix, the existing works designed the precoder based on the null space of the channel matrix to mitigate interference. They usually do not consider the channel estimation problem. In this paper, we consider the uplink transmission and focus on the access problem. We need to obtain accurate channel estimation in the case of multi-user interference, which is different from existing null-space-based work. The proposed method includes three stages. In the first initial orthogonal access stage, the users with the strongest channel large-scale fading coefficients are selected for orthogonal access to the network, where the number of these users is the same as that of the pilot sequences. In the second left-null-space-based opportunistic access stage, the users who fall in the left null space of the users in the first stage are selected to access the network. In the last stage, we detect the data of all accessed users. The simulation results demonstrate that, compared with the existing massive access methods, the proposed massive access method can achieve much higher spectral efficiency, especially in crowded scenarios.

The remainder of this paper is organized as follows. Section 2 develops a cell-free massive MIMO IoT system model. Section 3 introduces two traditional massive access

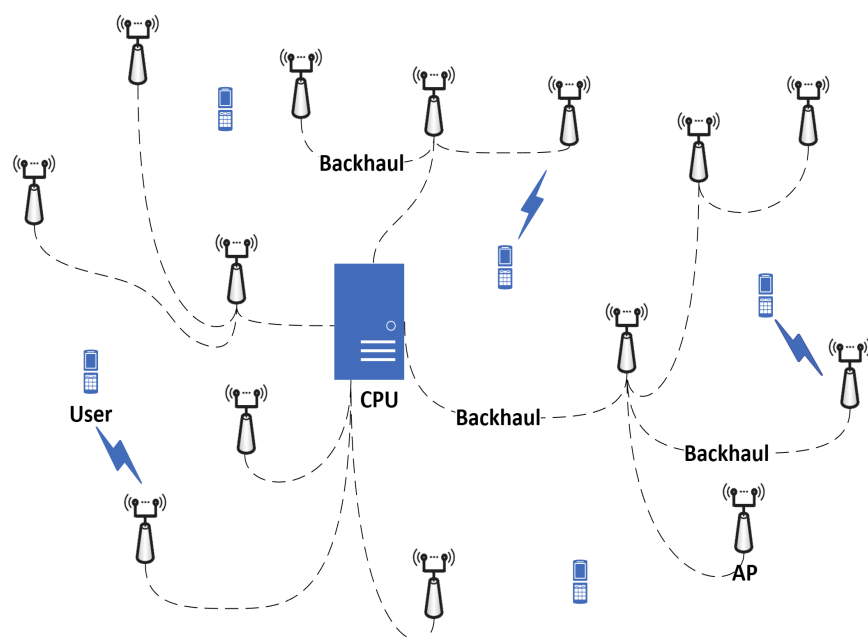
methods. Section 4 presents the proposed method, which includes three stages: initial orthogonal access, left-null-space-based opportunistic access and data detection of all accessed users. The performance of three methods is numerically evaluated in Section 5. Finally, the major conclusions and implications are drawn in Section 6.

*Notations:* The superscripts  $T, H$  denote transpose and conjugate transpose, respectively;  $\mathcal{N}(\mathbf{H}^T) = \{\mathbf{y} \in \mathbb{C}^{N_R \times 1} : \mathbf{y}^T \mathbf{H} = \mathbf{0}^T \in \mathbb{C}^{1 \times N_T}\}$  represents the left null space of a matrix  $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ ;  $\mathbf{I}_n$  stands for the  $n \times n$  identity matrix;  $N_{\mathbb{C}}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$  denotes the multivariate circularly symmetric complex Gaussian distribution with zero mean and variance  $\sigma^2 \mathbf{I}_n$ ;  $\|\mathbf{x}\|_2$  and  $\mathbb{E}\{\mathbf{x}\}$  denote the Euclidean norm and the expected value of  $\mathbf{x}$ , respectively.

## 2. System Model

Cell-free massive MIMO was proposed in [4] and can offer a high coverage probability. According to the varying degrees of cooperation between the APs and the CPU, there are four levels of receiver cooperation for cell-free massive MIMO, which can be called Level 4, Level 3, Level 2 and Level 1. Level 4 is a fully centralized network where all APs forward the signals to the CPU without any processing, and all channel estimation and data detection are performed at the CPU. Level 3 relies on the large-scale fading decoding (LSFD) strategy, where every AP locally estimates the channels, obtains the local estimates of the data and then passes them to the CPU for final decoding. Level 2 is a direct simplification of Level 3; the only difference between the two levels is the data detection at the CPU, where Level 2 simply takes the average of the local estimates and Level 3 uses an optimized weighting coefficient vector to maximize the SE. Level 1 is a small-cell network where all the channel estimates and data detection are performed locally at the APs. In this case, there is nothing to exchange between the CPU and the APs. The performances of the aforementioned four levels of receiver cooperation are analyzed and compared in [28], where it is shown that Level 4 can achieve the highest SE. Therefore, in this paper, we focus on Level 4.

Consider a cell-free massive MIMO system consisting of  $L$  APs equipped with  $N$  antennas. The APs are connected to a CPU via a backhaul network [4], as illustrated in Figure 1. The users waiting for connection to the system are equipped with a single antenna. All  $L$  APs directly send the received pilot sequences and data signals to the CPU without processing, and the CPU performs channel estimation and data detection.



**Figure 1.** A cell-free massive MIMO IoT system.

The received  $n$ th vector at the CPU,  $\mathbf{y}(n) \in \mathbb{C}^{LN \times 1}$ , can be expressed as

$$\mathbf{y}(n) = \sum_{k=1}^K \sqrt{p_k} \mathbf{h}_k x_k(n) + \mathbf{z}(n) \quad (1)$$

where  $K$  is the number of accessed users,  $p_k$  denotes the transmit power of the user  $k$ ,  $x_k(n)$  denotes the  $n$ th symbol transmitted by the user  $k$ ,  $\mathbf{z}(n) \in \mathbb{C}^{LN \times 1}$  is the complex Gaussian noise vector with zero mean and covariance matrix  $\sigma^2 \mathbf{I}_{LN}$ ,  $\mathbf{h}_k = [\mathbf{h}_{1k}^T, \dots, \mathbf{h}_{Lk}^T]^T \in \mathbb{C}^{LN \times 1}$  denotes the channel vector between the user  $k$  and all APs,  $\mathbf{h}_{lk} \in \mathbb{C}^{N \times 1}$  is the channel vector between the user  $k$  and the AP  $l$ , which is given by

$$\mathbf{h}_{lk} = \beta_{lk}^{1/2} \tilde{\mathbf{h}}_{lk} \quad (2)$$

$\beta_{lk}$  denotes the large-scale fading coefficient that depends on the location and the propagation environment, and  $\tilde{\mathbf{h}}_{lk} \sim N_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_N)$  represents the small-scale fading coefficient vector. We assume that  $\mathbf{h}_k$  is constant in a time block of  $\tau_c$  channel uses [11], where  $\tau_p$  channel uses are specifically for pilots and the remaining  $\tau_c - \tau_p$  channel uses are for payload data.

After collecting  $J$  consecutive received vectors, we can obtain

$$\mathbf{Y} = \sum_{k=1}^K \sqrt{p_k} \mathbf{h}_k \mathbf{x}_k + \mathbf{Z} \quad (3)$$

where  $\mathbf{Y} = [\mathbf{y}(1), \dots, \mathbf{y}(J)] \in \mathbb{C}^{LN \times J}$ ,  $\mathbf{x}_k = [x_k(1), \dots, x_k(J)] \in \mathbb{C}^{1 \times J}$  and  $\mathbf{Z} = [\mathbf{z}(1), \dots, \mathbf{z}(J)] \in \mathbb{C}^{LN \times J}$ .

### 3. Traditional Massive Access Methods

Assume that there are  $\tau_p$  mutually orthogonal  $\tau_p$ -length pilot sequences  $\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_{\tau_p} \in \mathbb{C}^{1 \times \tau_p}$  with  $\|\boldsymbol{\phi}_m\|_2^2 = \tau_p$ , where  $m = 1, \dots, \tau_p$ . Consider a highly crowded scenario where the total number of the users  $U$  is much larger than the number of available pilot sequences  $\tau_p$ . Since different users have to reuse the same pilot sequence, how to reduce interference and improve the system's performance becomes very important. Ref. [23] proposed a K-means-clustering-based pilot assignment method for cell-free massive MIMO systems, which can reduce the effect of the pilot contamination. In this method, the CPU first generates  $U/\tau_p$  centroids through the K-means clustering algorithm, which can be performed offline. Specifically,  $J_p$  points and  $U/\tau_p$  centroids are first arbitrarily generated in the coverage area, where  $J_p \gg U$ . Then, each point is assigned to its nearest centroid. After that, the centroid positions are updated by averaging the points belonging to the respective clusters. Next, all  $J_p$  points are reassigned to the new centroids. The above assignment is repeated, and the steps are updated until  $U/\tau_p$  stable clusters are formed. Each cluster has a centroid and comprises at most  $\tau_p$  users. After the centroids are determined, all  $U$  users access the network in sequence, where each user finds its closest centroid and joins the corresponding cluster. For any user, if the closest centroid has been selected by other  $\tau_p$  users, then it tries to join the next closest centroid, and repeats this procedure until it succeeds in joining a cluster. After all users find their clusters,  $\tau_p$  orthogonal pilot sequences are arbitrarily assigned to the users in the same cluster.

Different from the method where all users are allowed to access the system in [23], another massive access method is based on user selection. That is, according to the users' channel conditions, such as large-scale fading coefficients, the CPU selects  $K$  users from  $U$  users to access the system [18]. More specifically, let  $\bar{\beta}_k$  denote the average large-scale fading coefficient from the user  $k$  to all APs, where  $\bar{\beta}_k \triangleq \frac{1}{L} \sum_{l=1}^L \beta_{lk}$ .  $K$  users with the largest  $\bar{\beta}_k$  are allowed to access the network. In order to avoid strong interference among the users with good channel conditions, one can assign  $\tau_p$  pilot sequences to the  $\tau_p$  users with the best channel conditions and arbitrarily assign the pilot sequences to the remaining  $K - \tau_p$  users.

After pilot assignment for the accessed users, the CPU performs channel estimation for each user based on the received pilot sequences. Since the number of accessed users  $K$  is more than  $\tau_p$ , the same pilot sequence may be assigned to several users. Let  $\mathcal{S}_k$  represent the set of users sharing the same pilot sequence with user  $k$ , including user  $k$  itself. According to (3), during the training phase, the received pilot matrix at the CPU can be obtained as

$$\mathbf{Y}^p = \sum_{k=1}^K \sqrt{p_k} \mathbf{h}_k \boldsymbol{\phi}_{m_k} + \mathbf{Z}^p \quad (4)$$

where  $m_k \in \{1, \dots, \tau_p\}$  is the index of the pilot sequence assigned to the user  $k$ .

Given (4), the least square (LS) estimate of  $\mathbf{h}_k$  can be obtained as

$$\begin{aligned} \hat{\mathbf{h}}_k &= \frac{1}{\tau_p \sqrt{p_k}} \mathbf{Y}^p \boldsymbol{\phi}_{m_k}^H \\ &= \frac{1}{\tau_p \sqrt{p_k}} (\sqrt{p_k} \mathbf{h}_k \boldsymbol{\phi}_{m_k} + \sum_{i \in \mathcal{S}_k, i \neq k} \sqrt{p_i} \mathbf{h}_i \boldsymbol{\phi}_{m_k} + \sum_{j \notin \mathcal{S}_k, j=1}^K \sqrt{p_j} \mathbf{h}_j \boldsymbol{\phi}_{m_j} + \mathbf{Z}^p) \boldsymbol{\phi}_{m_k}^H \\ &= \mathbf{h}_k + \sum_{i \in \mathcal{S}_k, i \neq k} \mathbf{h}_i + \bar{\mathbf{z}}_k \end{aligned} \quad (5)$$

where  $\bar{\mathbf{z}}_k = \frac{1}{\tau_p \sqrt{p_k}} \mathbf{Z}^p \boldsymbol{\phi}_{m_k}^H$ . The third equation follows from the facts that  $\boldsymbol{\phi}_{m_k} \boldsymbol{\phi}_{m_k}^H = \tau_p$  and  $\boldsymbol{\phi}_{m_j} \boldsymbol{\phi}_{m_k}^H = 0$  for  $m_j \neq m_k$ . Note that the second term  $\sum_{i \in \mathcal{S}_k, i \neq k} \mathbf{h}_i$  in the third equation may cause so-called *pilot contamination* and degrade the system performance. Here we only consider LS channel estimator because it has low complexity and requires no prior statistical information [29].

Given  $\hat{\mathbf{h}}_k$ , the achievable spectral efficiency of the user  $k$  can be calculated by [28]

$$\text{SE}_k = \left(1 - \frac{\tau_p}{\tau_c}\right) \mathbb{E} \left\{ \log_2(1 + \text{SINR}_k) \right\} \quad (6)$$

where the signal-to-interference-and-noise ratio (SINR) is given by

$$\text{SINR}_k = \frac{p_k |\mathbf{v}_k^H \mathbf{h}_k|^2}{\sum_{i=1, i \neq k}^K p_i |\mathbf{v}_k^H \mathbf{h}_i|^2 + \sigma^2 \mathbf{v}_k^H \mathbf{v}_k} \quad (7)$$

and  $\mathbf{v}_k \in \mathbb{C}^{LN \times 1}$  is the combining vector for the user  $k$ . Any received combining vector  $\mathbf{v}_k$  can be adopted in (7). For example, one can let  $\mathbf{v}_k = \hat{\mathbf{h}}_k$ , i.e., use the simplest maximum ratio (MR) combination.

#### 4. The Proposed Massive Access Method

In this section, we propose a null-space-based massive access method for cell-free massive MIMO systems that can significantly reduce interference between users. Similar to [18], we consider a crowded system with  $U$  users and finally select  $K$  users to access the system. The proposed massive access method consists of three stages, which are summarized as follows.

##### 1. Stage 1: Initial orthogonal access

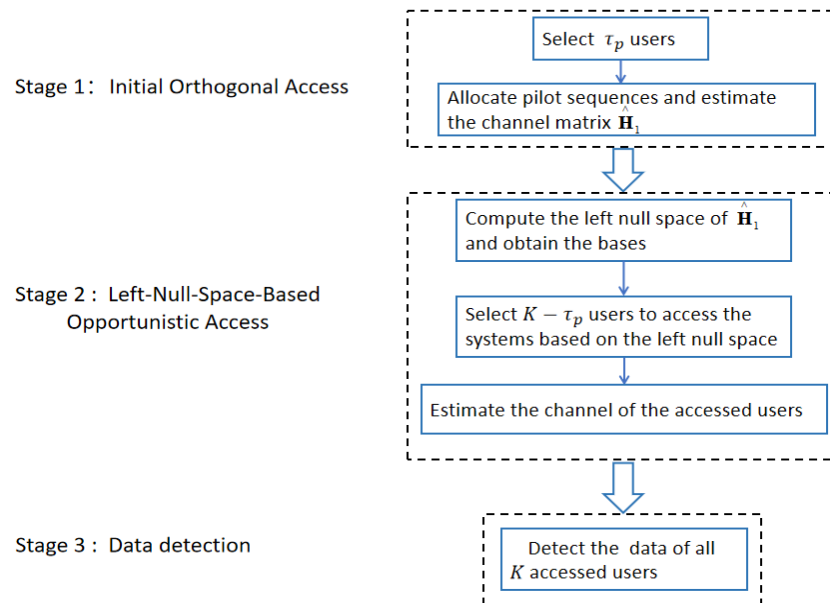
- (1) Step 1: Select  $\tau_p$  users from all users and allow them to access the system.
- (2) Step 2: Allocate  $\tau_p$  orthogonal pilot sequences to the accessed users and estimate their channel vectors.

##### 2. Stage 2: Left-null-space-based opportunistic access

- (1) Step 1: Compute the accessed users' left null space according to the results of the channel estimation in Stage 1 and obtain the bases of the left null space.

- (2) Step 2: Based on the left null space, select  $K - \tau_p$  users from the remaining  $U - \tau_p$  users to access the network.
  - (3) Step 3: Estimate the channel of the new selected  $K - \tau_p$  users.
3. Stage 3: Detect the data of all  $K$ -accessed users.

Figure 2 is a flow diagram of the proposed method. In the proposed method,  $\tau_p$  users in the first stage can access the system without interference, and  $K - \tau_p$  users in the second stage can access the system with less interference. The data of all  $K$  users are detected in the last stage. More details about the three stages are given in Section 4.1, Section 4.2 and Section 4.3, respectively.



**Figure 2.** The proposed massive access method.

#### 4.1. Stage 1: Initial Orthogonal Access

In this stage, the CPU allows  $\tau_p$  users to access the system according to the number of pilot sequences  $\tau_p$ . Different schemes can be adopted to select these users. For example, the simplest way is to randomly select  $\tau_p$  users from all users. However, such a random selection scheme cannot guarantee the performance of the channel estimation because it may select some users with poor channel conditions. In order to ensure the accuracy of the channel estimation, we adopt the similar user selection method described in Section 3. That is, we select  $\tau_p$  users based on the strongest channel large-scale fading coefficients and then allocate  $\tau_p$  orthogonal pilot sequences to them.

Due to the orthogonality of the pilot sequences, according to (5), the channel estimation of the user  $k$  can be given by

$$\hat{\mathbf{h}}_k = \mathbf{h}_k + \bar{\mathbf{z}}_k \quad (8)$$

where  $k = 1, \dots, \tau_p$ . It can be seen that there is no interference between the users. Furthermore, since all the selected users have good channel conditions, the accuracy of the channel estimation can be guaranteed.

#### 4.2. Stage 2: Left-Null-Space-Based Opportunistic Access

To exploit the resources in the spatial domain, we propose an algorithm based on the orthogonal property of the left null space. By taking advantage of the orthogonality between different bases of the left null space, the CPU can distinguish different users in the spatial domain even though they share the same pilot sequence. This is equivalent to

adding an additional spatial domain dimension to the original time and frequency domain dimensions, which can significantly reduce the pilot contamination effect.

In this stage, all the remaining  $U - \tau_p$  users first randomly choose their pilot sequence from the predefined set and then transmit them to the APs. After all of the APs forward the received signals to the CPU, the received pilot matrix at the CPU,  $\mathbf{Y}_2^p \in \mathbb{C}^{LN \times \tau_p}$ , can be obtained as

$$\begin{aligned} \mathbf{Y}_2^p &= \sum_{i=1}^{U-\tau_p} \sqrt{p_i} \mathbf{h}_i \boldsymbol{\phi}_{m_i} + \mathbf{Z}_2^p \\ &= \sum_{m=1}^{\tau_p} \sum_{i \in \mathcal{P}_m} \sqrt{p_i} \mathbf{h}_i \boldsymbol{\phi}_m + \mathbf{Z}_2^p \end{aligned} \tag{9}$$

where  $\mathcal{P}_m$  is the set of users sharing the same pilot sequence  $m$  and  $\mathbf{Z}_2^p \in \mathbb{C}^{LN \times \tau_p}$ .

To select  $K - \tau_p$  users with the least interference from the whole  $U - \tau_p$  users, the CPU then correlates the received signal with the normalized pilot sequence  $\boldsymbol{\phi}_m^H / \sqrt{\tau_p}$  to obtain  $\bar{\mathbf{h}}_m \in \mathbb{C}^{LN \times 1}$ , which is given by

$$\begin{aligned} \bar{\mathbf{h}}_m &\triangleq \frac{1}{\sqrt{\tau_p}} \mathbf{Y}_2^p \boldsymbol{\phi}_m^H \\ &= \sum_{i \in \mathcal{P}_m} \sqrt{p_i \tau_p} \mathbf{h}_i + \bar{\mathbf{z}}_m \end{aligned} \tag{10}$$

where  $\bar{\mathbf{z}}_m = \frac{1}{\sqrt{\tau_p}} \mathbf{Z}_2^p \boldsymbol{\phi}_m^H$ . By doing so, users are divided into  $\tau_p$  groups, where each group of users shares the same pilot sequence.

After that, the CPU computes the left null space  $\mathcal{N}(\hat{\mathbf{H}}_1^T)$  according to the channel estimations obtained in Stage 1, where  $\hat{\mathbf{H}}_1 = [\hat{\mathbf{h}}_1, \dots, \hat{\mathbf{h}}_{\tau_p}] \in \mathbb{C}^{LN \times \tau_p}$ . QR factorization can be used to obtain  $\mathcal{N}(\hat{\mathbf{H}}_1^T)$ , i.e., the right  $LN - \tau_p$  column of the Q matrix after QR factorization is  $\mathcal{N}(\hat{\mathbf{H}}_1^T)$ . Let  $\mathbf{e}_n$  denote the  $n$ th column of the  $\mathcal{N}(\hat{\mathbf{H}}_1^T)$  and define  $\mathbf{W}_n = \mathbf{e}_n \mathbf{e}_n^H$ , where  $n = 1, \dots, LN - \tau_p$ .  $\mathbf{e}_n$  represents the base of the left null space. Different bases are orthogonal to they are all orthogonal to  $\hat{\mathbf{H}}_1$ . each other and Given  $\mathcal{N}(\hat{\mathbf{H}}_1^T)$ , the CPU next searches for the users who fall in the left null space. Specifically, the CPU computes the following projection as

$$\begin{aligned} \bar{\mathbf{g}}_{n,m} &= \mathbf{W}_n \bar{\mathbf{h}}_m \\ &= \sum_{i \in \mathcal{P}_m} \sqrt{p_i \tau_p} \mathbf{W}_n \mathbf{h}_i + \mathbf{W}_n \bar{\mathbf{z}}_m \\ &= \sqrt{p_j \tau_p} \mathbf{W}_n \mathbf{h}_j + \sum_{i \in \mathcal{P}_m, i \neq j} \sqrt{p_i \tau_p} \mathbf{W}_n \mathbf{h}_i + \mathbf{W}_n \bar{\mathbf{z}}_m \end{aligned} \tag{11}$$

If the energy of  $\bar{\mathbf{g}}_{n,m}$  is greater than a certain threshold  $\varepsilon$ , i.e.,

$$\mathbf{E}_{n,m} > \varepsilon \tag{12}$$

where  $\mathbf{E}_{n,m} = \|\bar{\mathbf{g}}_{n,m}\|_2^2$ , we consider that there is a user in the left null space. Here we assume that there is at most one user whose channel vector matches the base  $\mathbf{e}_n$ . Once (12) happens, the CPU puts this user, whose index is denoted by  $j_{(n,m)}$ , into the list  $\mathcal{T}_s$ , where  $\mathcal{T}_s$  denotes the set of index values of the users that are allowed to access the system. Note that by adjusting the threshold  $\varepsilon$ , the CPU can control the number of accessed users in Stage 2.

When (12) occurs, (11) can be further written as

$$\bar{\mathbf{g}}_{n,j_{(n,m)}} = \sqrt{p_{j_{(n,m)}} \tau_p} \mathbf{W}_n \mathbf{h}_{j_{(n,m)}} + \mathbf{W}_n \bar{\mathbf{z}}_m \tag{13}$$



because  $\mathbf{W}_n \mathbf{h}_i = 0$  when users are in different subspaces. Let  $\mathbf{g}_{n,j(n,m)}$  be the projection of the true channel of the user  $j(n,m)$ , i.e.,

$$\mathbf{g}_{n,j(n,m)} \triangleq \mathbf{W}_n \mathbf{h}_{j(n,m)} \quad (14)$$

Then, the LS estimate of  $\mathbf{g}_{n,j(n,m)}$  can be obtained as

$$\begin{aligned} \hat{\mathbf{g}}_{n,j(n,m)} &= \frac{1}{\sqrt{p_{j(n,m)}} \tau_p} \bar{\mathbf{g}}_{n,j(n,m)} \\ &= \mathbf{g}_{n,j(n,m)} + \frac{1}{\sqrt{p_{j(n,m)}} \tau_p} \mathbf{W}_n \bar{\mathbf{z}}_m \end{aligned} \quad (15)$$

Based on the above algorithm, the CPU can find the users who are in the left null space of the accessed users in Stage 1. Allowing these users access into the system can greatly reduce inter-user interference. Here, we assume  $K - \tau_p$  users successfully accessed the system in the end. The proposed left-null-space-based access algorithm is summarized in Algorithm 1.

---

**Algorithm 1:** Left-Null-Space-Based Access

---

- 1 The remaining  $U - \tau_p$  users randomly transmit the pilot sequence to the APs.
  - 2 Given the received pilot matrix  $\mathbf{Y}_2^p$ , the CPU obtains  $\bar{\mathbf{h}}_m$  by using (10), where  $m = 1, \dots, \tau_p$ .
  - 3 Given  $\hat{\mathbf{h}}_k$ , compute  $\mathcal{N}(\hat{\mathbf{H}}_1^T)$  by, e.g., QR factorization. Obtain  $\mathbf{e}_n$  and then  $\mathbf{W}_n = \mathbf{e}_n \mathbf{e}_n^H$ , where  $n = 1, \dots, LN - \tau_p$ . Set  $\mathcal{T}_s = \emptyset$ .
  - 4 **for**  $1 \leq m \leq \tau_p$  **do**
  - 5 **for**  $1 \leq n \leq LN - \tau_p$  **do**
  - 6 Obtain  $\bar{\mathbf{g}}_{n,m}$  according to (11).
  - 7 Calculate the energy of  $\bar{\mathbf{g}}_{n,m}$ , i.e.,  $\mathbf{E}_{n,m} = \|\bar{\mathbf{g}}_{n,m}\|_2^2$ .
  - 8 **if**  $\mathbf{E}_{n,m} > \varepsilon$  **then**
  - 9 Put user  $j(n,m)$  into the list  $\mathcal{T}_s$ .
  - 10 **end**
  - 11 **end**
  - 12 **end**
  - 13 **final**
- 

In summary, the proposed massive access method has the following advantages. Firstly, the users accessed in Stage 1 use the orthogonal pilot sequences to access the system; hence, they have no interference and accurate channel estimation. Secondly, due to the orthogonality between  $\mathcal{N}(\hat{\mathbf{H}}_1^T)$  and  $\hat{\mathbf{H}}_1$ , users in Stage 2 will cause less interference to users in Stage 1. Finally, since the users accessed in Stage 2 are selected according to the bases and different bases are orthogonal to each other, the pilot contamination effect caused by pilot-sharing users will be significantly reduced in Stage 2. A similar method is to add a space-domain dimension, where users are able to share the same pilot with no interference. By doing so, the number of accessed users can be greatly increased with less interference.

#### 4.3. Stage 3: Detect the Data of All K-Accessed Users

According to (8) and (15), the estimation of the channel for all accessed users can be obtained. Finally, the CPU can detect the data of all users based on the estimated channel.

For users accessed in Stage 1, we use the SINR expression in (7) to calculate the achievable SE, where the MR combining with  $\mathbf{v}_k = \hat{\mathbf{h}}_k$  is used for data detection and  $\hat{\mathbf{h}}_k$  is given by (8).



To detect the data of the users accessed in Stage 2, the CPU needs to project the received signal  $\mathbf{y}$  into the subspace first. The received signal of the user  $j_{(n,m)}$  after projection can be obtained as

$$\begin{aligned} \mathbf{y}_{j_{(n,m)}} &= \mathbf{W}_n \sum_{i=1}^{K-\tau_p} \mathbf{h}_i s_i + \mathbf{W}_n \mathbf{z} \\ &= \mathbf{g}_{n,j_{(n,m)}} s_{j_{(n,m)}} + \sum_{i \in \mathcal{C}_n, i \neq j_{(n,m)}} \mathbf{g}_{n,i} s_i + \tilde{\mathbf{z}}_{j_{(n,m)}} \end{aligned} \quad (16)$$

where  $\mathcal{C}_n$  is the set of the users whose channels match the same base  $e_n$  and  $\tilde{\mathbf{z}}_{j_{(n,m)}} = \mathbf{W}_n \mathbf{z}$ . The second equation follows the fact that  $\mathbf{W}_n \mathbf{h}_i = 0$  when  $i \notin \mathcal{C}_n$ . This implies that users in the same group will not cause any interference, that interference only happens between inter-group users and that there are at most  $\tau_p$  elements in  $\mathcal{C}_n$ .

Since  $\mathbf{g}_{n,j_{(n,m)}}$  can be viewed as the equivalent channel of user  $j_{(n,m)}$ , we replace  $\mathbf{h}_{j_{(n,m)}}$  with  $\mathbf{g}_{n,j_{(n,m)}}$  and use MR combination i.e.,  $\mathbf{v}_{j_{(n,m)}} = \hat{\mathbf{g}}_{n,j_{(n,m)}}$  in (7) to obtain the effective SINR.

## 5. Numerical Results

In this section, we evaluate the uplink performance of our proposed massive access method. We consider a setup with  $U$  users who are independently and uniformly distributed at random within a square size of  $1 \text{ km} \times 1 \text{ km}$  and finally select  $K$  users from them to access the system.  $L = 100$  APs are deployed randomly within the square coverage area, and all APs are equipped with half-wavelength-spaced uniform linear arrays with  $N = 4$  antennas. We assume that all APs are deployed in urban environments with high user loads, roughly 10 m above the ground. The 3GPP Urban Microcell model in Tab. B.1.2.1-1 [30] with a 2 GHz carrier frequency is used to generate the large-scale propagation conditions, such as pathloss and shadow fading.

$$\beta_{kl} [\text{dB}] = -30.5 - 36.7 \log_{10} \left( \frac{d_{kl}}{1\text{m}} \right) + F_{kl} \quad (17)$$

where  $d_{kl}$  is the distance between the user  $k$  and AP  $l$  and  $F_{kl} \sim \mathcal{N}(0, 4^2)$  is the shadow fading. We present the key simulation parameters in Table 1.

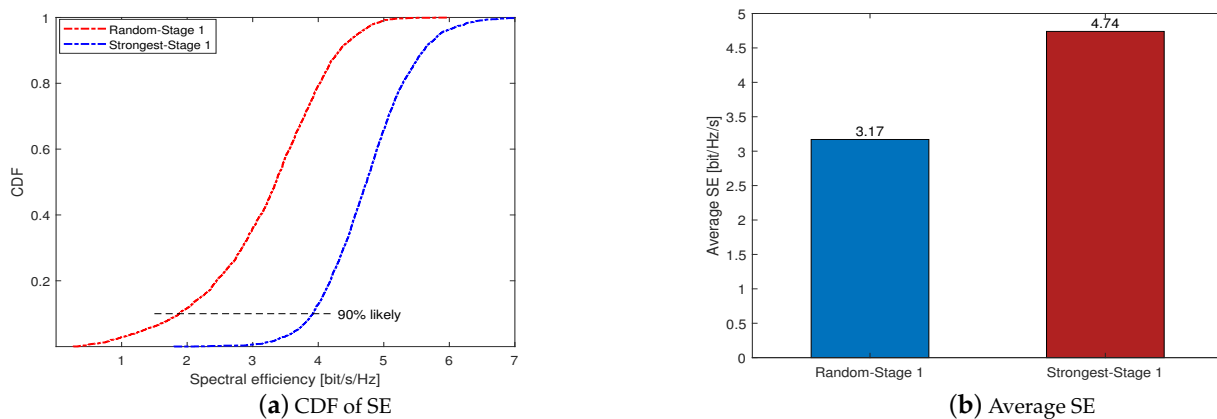
**Table 1.** Key simulation parameters.

Parameters	Value
Transmit powers	$p_k = 100 \text{ mW}$
System bandwidth	$B = 20 \text{ MHz}$
Noise power	$\sigma^2 = -96 \text{ dBm}$
Coherent bandwidth	100 KHz
Coherent time	20 ms
Total channel uses	$\tau_c = 200$
Pilot sequences	$\tau_p = 10$

In the simulations, we use “Strongest Users method” [18] and “K-Means method” [23] to denote the methods described in Section 3 and compare them with the proposed method.

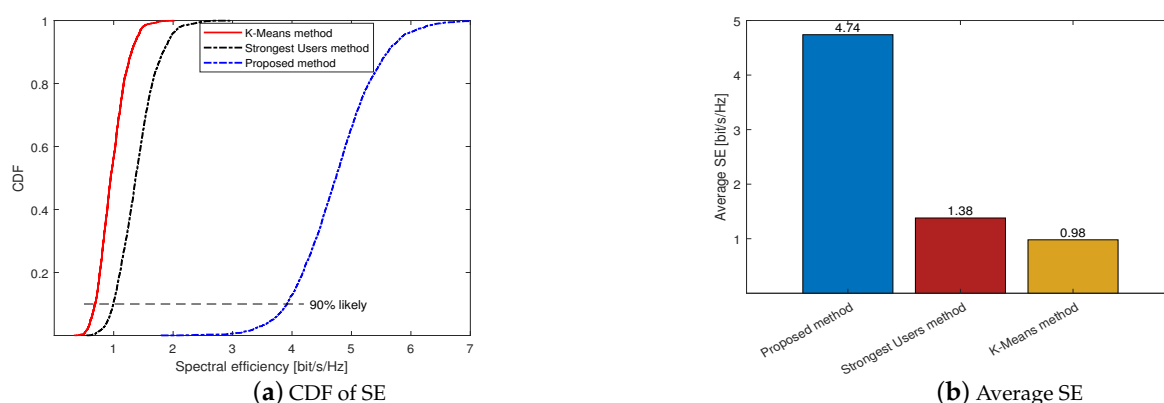
Figure 3 compares the two user selection schemes we proposed in Stage 1 under  $K = 40$  and  $U = 100$ . For the sake of description, we use “Random-Stage 1” and “Strongest-Stage 1” to denote the random selection scheme and the strongest channel large-scale fading coefficients user selection scheme, respectively. Figure 3a shows the cumulative distribution function (CDF) of the SE per user when using an LS estimator and MR combining. As expected, the Strongest-Stage 1 outperforms the Random-Stage 1, where the Strongest-Stage 1 curve is totally on the right of the Random-Stage 1 curve. At the 90% likely SE points,

the SE of the the Strongest-Stage 1 and the Random-Stage 1 are 3.91 and 1.87, respectively. Figure 3b shows that the average SE of the Strongest-Stage 1 and the Random-Stage 1 are 4.74 and 3.17, which is consistent with Figure 3a. Since the first scheme ensures the quality of the channel estimation, it can guarantee the accuracy of the obtained left null space. Hence, the strongest channel large-scale fading coefficients user selection scheme performs much better than the random selection scheme, and we use it as the default initial user access scheme in Stage 1 for the following simulation.



**Figure 3.** Comparison of different user selection schemes in Stage 1 under  $K = 40$  and  $U = 100$ .

Figure 4 shows the comparison of the three massive access methods, where we let  $K = 40$ ,  $U = 100$  and use an LS channel estimator and MR combining. It can be seen from Figure 4a that the proposed method outperforms the other two. For example, at the 90% likely SE points, the SEs of the proposed method, the Strongest Users method and the K-Means method are 3.91, 0.99 and 0.69, respectively. This means that the proposed method can achieve a higher SE than the other two methods for most users. Figure 4b shows the average SE of the proposed method, Strongest Users method and K-Means method, they are 4.74, 1.38 and 0.98, respectively.



**Figure 4.** Comparison of different massive access methods under  $K = 40$  and  $U = 100$ .

Figure 5 shows the average SE performance of the three methods under different  $K$ . We let the number of accessed users in Stage 2 be about one-third of the total users, i.e.,  $(U - \tau_p)/3$ . It can be seen from Figure 5 that the proposed method has a gentle descent and is always above the other two methods, which means that the performance of the proposed method outperforms the other two methods. When  $K = 50$ , the average SEs of the proposed method, the Strongest Users method and the K-Means method are 4.27, 1.08 and 0.77, respectively. Note that the average SE is a decreasing function regardless of

any massive access methods in the vast majority of intervals. The reason is that the larger number of users can cause stronger inter-user interference, which leads to a rapid reduction in the average SE.

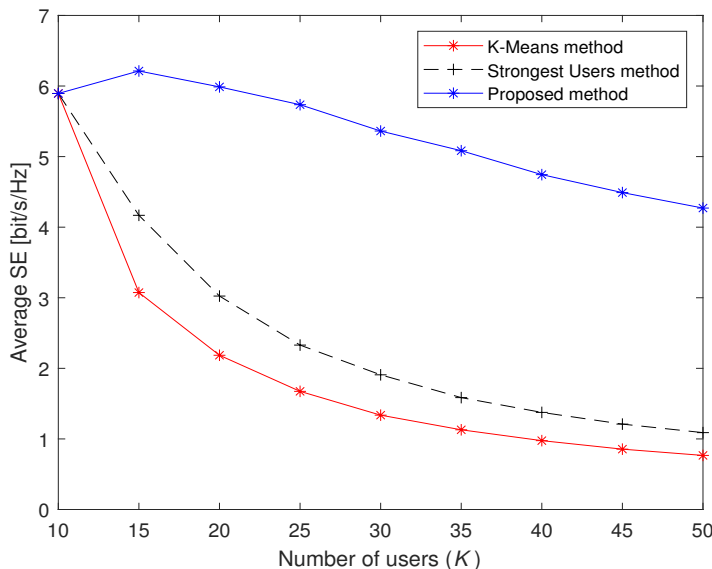


Figure 5. Average SE of three massive access methods under different the number of users K.

Figure 6 depicts the CDF of the sum SE over different massive access methods under  $K = 40$  and  $U = 100$ . The result shows that the proposed massive access method outperforms the other two methods, which is echoed by the results in Figure 4. The reason is that the users in our proposed massive access method are less influenced by interference, so it can improve the sum SE compared with the other two methods.

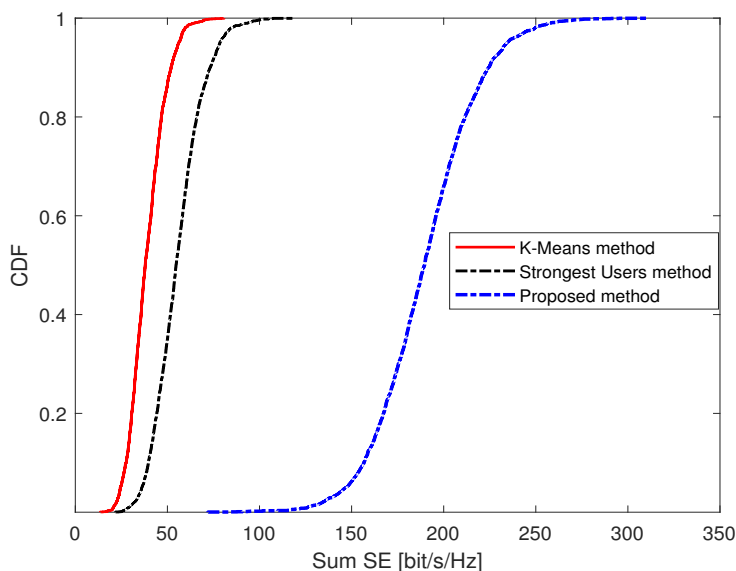
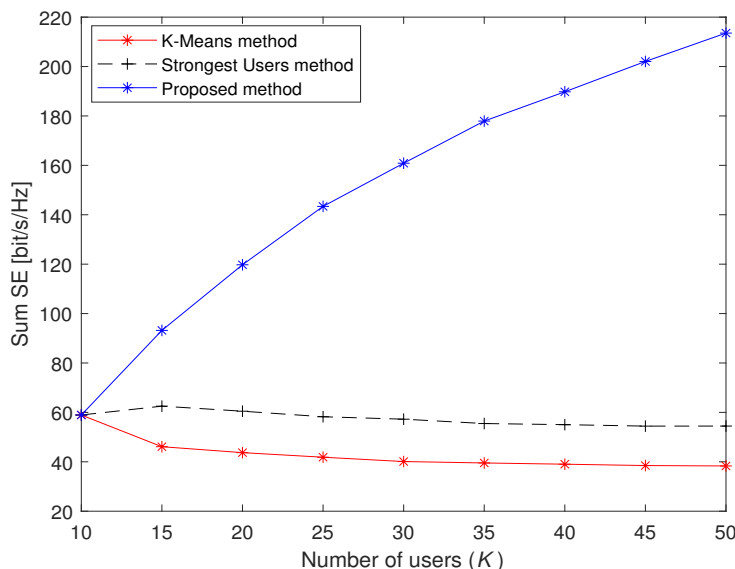


Figure 6. The CDF of the sum SE under different massive access methods under  $K = 40$  and  $U = 100$ .

Figure 7 shows the sum SE comparison of the three methods under different K. The setting of U is the same as that in the simulations in Figure 5. When  $K = 40$ , the sum SE of the proposed method, the Strongest Users method and the K-Means method are 189.77, 55.03 and 39.02, respectively. For the K-Means method, it is clear that the sum SE slowly decreases when the number of access users K increases. The Strongest Users method has the same distribution as the K-Means method, but it can achieve a higher sum SE. Compared with

the other two methods, the proposed method has much better performance since the curve of the proposed method is an increasing function of  $K$ . The reason is that the traditional massive access methods are greatly affected by interference; the more users connected to the system, the more obvious this phenomenon is. Since the proposed massive access method can alleviate the impact of multi-user interference through the left-null-space-based access algorithm, the sum SE can be increased in crowded scenarios.



**Figure 7.** Sum SE of three massive access methods under different the number of users  $K$ .

## 6. Conclusions

In this paper, we have proposed a left-null-space-based massive access method for cell-free massive MIMO systems. The key idea of the proposed method is to exploit the space-domain orthogonality. The proposed method first lets some of the users orthogonally access the system, and then the remaining users are given access to the system according to the left null space of the accessed users. The simulation results show that our proposed method performs much better than the existing methods, which can achieve higher SE by suppressing inter-user interference.

Under the current framework, all signal processing is performed on the CPU, which poses some challenges to the CPU's processing capacity. In addition, the massive data exchange between the AP and the CPU also places a great load on the backhaul network. Therefore, in future research, we will apply this method to a completely distributed small cell network, where all the channel estimations and data detection are performed locally at the APs. Since there are no experimental testbed cases used yet, we are looking forward to verifying these results in future studies.

**Author Contributions:** Conceptualization, D.H.; methodology, Y.W. and D.H.; software, Y.W.; validation, Y.W. and D.H.; formal analysis, Y.W.; investigation, Y.W.; resources, Y.W.; data curation, Y.W.; writing—original draft preparation, Y.W.; writing—review and editing, D.H.; visualization, Y.W.; supervision, D.H.; project administration, D.H.; funding acquisition, D.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China under Grant 62271155.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data that support the findings of this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Al-Fuqaha, A.; Guizani, M.; Mohammadi, M.; Aledhari, M.; Ayyash, M. Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications. *IEEE Commun. Surv. Tuts.* **2015**, *17*, 2347–2376. [[CrossRef](#)]
2. Misra, N. N.; Dixit, Y.; Al-Mallahi, A.; Bhullar, M. S.; Upadhyay, R.; Martynenko, A. IoT, Big Data, and Artificial Intelligence in Agriculture and Food Industry. *IEEE Internet Things J.* **2022**, *7*, 6305–6324. [[CrossRef](#)]
3. Yan, H.; Ashikhmin, A.; Yang, H. A Scalable and Energy-Efficient IoT System Supported by Cell-Free Massive MIMO. *IEEE Internet Things J.* **2021**, *8*, 14705–14718. [[CrossRef](#)]
4. Ngo, H.Q.; Ashikhmin, A.; Yang, H.; Larsson, E.G.; Marzetta, T.L. Cell-Free Massive MIMO Versus Small Cells. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 1834–1850. [[CrossRef](#)]
5. Zhang, J.; Chen, S.; Lin, Y.; Zheng, J.; Ai, B.; Hanzo, L. Cell-Free Massive MIMO: A New Next-Generation Paradigm. *IEEE Access* **2019**, *7*, 99878–99888. [[CrossRef](#)]
6. Ngo, H.Q.; Tran, L.N.; Duong, T.Q.; Matthaiou, M.; Larsson, E.G. On the Total Energy Efficiency of Cell-Free Massive MIMO. *IEEE Trans. Green Commun. Netw.* **2018**, *2*, 25–39. [[CrossRef](#)]
7. Nayebi, E.; Ashikhmin, A.; Marzetta, T.L.; Rao, B.D. Performance of cell-free massive MIMO systems with MMSE and LSFD receivers. In Proceedings of the 2016 50th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 6–9 November 2016; pp. 203–207. [[CrossRef](#)]
8. Matthaiou, M.; Yurduseven, O.; Ngo, H.Q.; Morales-Jimenez, D.; Cotton, S.L.; Fusco, V.F. The Road to 6G: Ten Physical Layer Challenges for Communications Engineers. *IEEE Commun. Mag.* **2021**, *59*, 64–69. [[CrossRef](#)]
9. Alsabah, M.; Naser, M.A.; Mahmmud, B.M.; Abdulhussain, S.H.; Eissa, M.R.; Al-Baidhani, A.; Noordin, N.K.; Sait, S.M.; Al-Utaibi, K.A.; Hashim, F. 6G Wireless Communications Networks: A Comprehensive Survey. *IEEE Access* **2021**, *9*, 148191–148243. [[CrossRef](#)]
10. Elijah, O.; Leow, C.Y.; Rahman, T.A.; Nunoo, S.; Iliya, S.Z. A Comprehensive Survey of Pilot Contamination in Massive MIMO—5G System. *IEEE Commun. Surv. Tutorials* **2016**, *18*, 905–923. [[CrossRef](#)]
11. Björnson, E.; Hoydis, J.; Sanguinetti, L. Massive MIMO networks: Spectral, energy, and hardware efficiency. *Found. Trends<sup>®</sup> Signal Process.* **2017**, *11*, 154–655. [[CrossRef](#)]
12. Jose, J.; Ashikhmin, A.; Marzetta, T.L.; Vishwanath, S. Pilot Contamination and Precoding in Multi-Cell TDD Systems. *IEEE Trans. Wirel. Commun.* **2011**, *10*, 2640–2651. [[CrossRef](#)]
13. Khansefid, A.; Minn, H. On Channel Estimation for Massive MIMO With Pilot Contamination. *IEEE Commun. Lett.* **2015**, *19*, 1660–1663. [[CrossRef](#)]
14. Ngo, H.Q.; Ashikhmin, A.; Yang, H.; Larsson, E.G.; Marzetta, T.L. Cell-Free Massive MIMO: Uniformly great service for everyone. In Proceedings of the 2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Stockholm, Sweden, 28 June–1 July 2015; pp. 201–205. [[CrossRef](#)]
15. Chen, S.; Zhang, J.; Björnson, E.; Zhang, J.; Ai, B. Structured Massive Access for Scalable Cell-Free Massive MIMO Systems. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 1086–1100. [[CrossRef](#)]
16. Hasan, M.; Hossain, E.; Niyato, D. Random access for machine-to-machine communication in LTE-advanced networks: Issues and approaches. *IEEE Commun. Mag.* **2013**, *51*, 86–93. [[CrossRef](#)]
17. Björnson, E.; de Carvalho, E.; Sørensen, J.H.; Larsson, E.G.; Popovski, P. A Random Access Protocol for Pilot Allocation in Crowded Massive MIMO Systems. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 2220–2234. [[CrossRef](#)]
18. Björnson, E.; de Carvalho, E.; Larsson, E.G.; Popovski, P. Random access protocol for massive MIMO: Strongest-user collision resolution (SUCR). In Proceedings of the 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur, Malaysia, 22–27 May 2016; pp. 1–6. [[CrossRef](#)]
19. Sørensen, J.H.; de Carvalho, E.; Stefanovic, Č.; Popovski, P. Coded Pilot Random Access for Massive MIMO Systems. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 8035–8046. [[CrossRef](#)]
20. de Carvalho, E.; Björnson, E.; Sørensen, J.H.; Larsson, E.G.; Popovski, P. Random Pilot and Data Access in Massive MIMO for Machine-Type Communications. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 7703–7717. [[CrossRef](#)]
21. De Carvalho, E.; Björnson, E.; Sørensen, J.H.; Popovski, P.; Larsson, E.G. Random Access Protocols for Massive MIMO. *IEEE Commun. Mag.* **2017**, *55*, 216–222. [[CrossRef](#)]
22. Wang, X.; Ashikhmin, A.; Wang, X. Wirelessly Powered Cell-Free IoT: Analysis and Optimization. *IEEE Internet Things J.* **2020**, *7*, 8384–8396. [[CrossRef](#)]
23. Attarifar, M.; Abbasfar, A.; Lozano, A. Random vs Structured Pilot Assignment in Cell-Free Massive MIMO Wireless Networks. In Proceedings of the 2018 IEEE International Conference on Communications Workshops (ICC Workshops), Kansas City, MO, USA, 20–24 May 2018; pp. 1–6. [[CrossRef](#)]
24. Zhu, X.; Dai, L.; Wang, Z.; Wang, X. Weighted-Graph-Coloring-Based Pilot Decontamination for Multicell Massive MIMO Systems. *IEEE Trans. Veh. Technol.* **2017**, *66*, 2829–2834. [[CrossRef](#)]
25. Atzeni, I.; Arnau, J.; Debbah, M. Fractional pilot reuse in massive MIMO systems. In Proceedings of the 2015 IEEE International Conference on Communication Workshop (ICCW), London, UK, 8–12 June 2015; pp. 1030–1035. [[CrossRef](#)]

26. Lee, W.S.; Ro, J.H.; You, Y.H.; Hwang, D.; Song, H.K. An Enhanced Precoder for Multi User Multiple-Input Multiple-Output Downlink Systems. *Appl. Sci.* **2020**, *10*, 4547. [[CrossRef](#)]
27. Iwakuni, T.; Maruta, K.; Ohta, A.; Shirato, Y.; Arai, T.; Iizuka, M. Inter-user interference suppression in time varying channel with null-space expansion for multiuser massive MIMO. In Proceedings of the 2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Hong Kong, China, 30 August–2 September 2015; pp. 558–562. [[CrossRef](#)]
28. Björnson, E.; Sanguinetti, L. Making Cell-Free Massive MIMO Competitive With MMSE Processing and Centralized Implementation. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 77–90. [[CrossRef](#)]
29. Fan, W.; Zhang, J.; Björnson, E.; Chen, S.; Zhong, Z. Performance Analysis of Cell-Free Massive MIMO Over Spatially Correlated Fading Channels. In Proceedings of the ICC 2019—2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019; pp. 1–6. [[CrossRef](#)]
30. 3GPP. Further advancements for E-UTRA physical layer aspects (release 9). In *3GPP, TS 36.814*; 2017. Available online: [https://www.3gpp.org/ftp/Specs/archive/36\\_series/36.814/](https://www.3gpp.org/ftp/Specs/archive/36_series/36.814/) (accessed on 30 May 2023)

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.