Channel Estimation using In-Band Pilots for Cell-Free Massive MIMO

Chan-Tong Lam, Ke Wang, and Benjamin K. Ng
Faculty of Applied Sciences, *Macao Polytechnic University*
Macao SAR, China
Email: {ctlam, ke.wang, bng}@mpu.edu.mo

***Abstract*—In this paper, a novel channel estimation technique using in-band pilots (IBPs) for uplink transmission is proposed for cell-free (CF) massive multiple-input multiple-output (MIMO) networks. In particular, compared to traditional superimposed pilot schemes in the time domain, the proposed IBP scheme is operated in the frequency domain, which is more effective and robust due to the removal of interference from the data frequencies at the pilot location. We first introduce a system model for the CF massive MIMO with the IBP design. Then, analytical expressions for the mean square error and the normalized mean square error (NMSE) are derived for performance evaluation. Analytical and simulated results reveal that the proposed IBP scheme for CF MIMO outperforms the conventional regular and superimposed pilots in terms of NMSE. Besides, other important insights include the IBP scheme works well in high transmit power regions, and this new channel estimation strategy is more suitable for the scenario with a large number of users since it is not that sensitive to the number of users.**

***Index Terms*—channel estimation, in-band pilots, superimposed pilots, cell-free massive MIMO, NMSE**

I. Introduction

Among candidate technologies for the next-generation communication system, cell-free (CF) massive multiple-input multiple-output (MIMO) has been envisioned as a promising solution to enhance wireless transmission efficiencies and provide a larger coverage [1]–[3]. User-centric CF networks exploit the benefits of the physical layer from cellular massive MIMO, ultra-dense network, and coordinated multi-point with joint transmissions, resulting in a network architecture that provides almost uniformly high data rates in a given geographical area [2]–[4]. In [5], the authors for the first time showed that the CF massive MIMO system can significantly outperform small-cell systems in terms of throughput, mainly because the CF system is much more reliable for shadow fading correlation. Besides, in [6], the authors illustrated that a centralized implementation with optimal minimum mean square error (MSE) processing not only maximizes the spectral efficiency (SE) but largely reduces fronthaul signaling compared to the standard distributed approach. Moreover, the authors in [7] revealed that the CF system and the legacy cellular system could coexist, which establishes that the gradual implementation of access points (APs) in conventional cellular systems can decrease the performance gap between users and improve the average SE of the system.

This work was funded by the research funding of the Macao Polytechnic University, Macao SAR, China (Project no. RP/ESCA-02/2020).

For the typical CF network, channel estimations can be done locally at the APs or be delegated to the central processing units (CPUs) [4]. Similar to the channel estimation for the conventional massive MIMO system, the estimation process for the CF communication system also suffers from pilot contaminations resulting from non-orthogonal pilots [1], [2], [4]. Some early CF massive MIMO studies [5], [8] assumed that all pilot signals are transmitted at full power during the training phase. This may, however, increase the pilot contamination especially when a user equipment (UE) has poor channel responses. Mai *et al*. [9] showed that in order to enhance the accuracy of the channel estimation and therefore the overall performance of CF massive MIMO, the most appropriate pilot power should be distributed to corresponding UEs during the training phase. Besides, opportunistic AP selections [10] can also bring a significant performance boost.

It should be emphasized that, for the CF network, removing the pilot contamination and improving the performance can not simply rely on power controls or AP selections [1]–[4]. Pilot designs of the CF massive MIMO system (e.g., enlarging the size of orthogonal pilots) are also of great significance to network operators. One way to increase the number of orthogonal pilots is to superimpose pilot signals on data signals, at the expense of data interference [1]–[3]. Upadhya *et al*. [11], [12] investigated the downlink throughput performance of a massive MIMO system that employs superimposed pilots (SPs) for channel estimation. The result shows that the SP strategy can not only enhance SE but also reduce the pilot contamination effect. Verenzuela *et al*. [13] derived a closed-form SE expression for an uplink massive MIMO system with SP-aided channel estimations, and proved that utilizing the SP can significantly reduce the pilot contamination. Zhang *et al*. [14] introduced the SP to the CF massive MIMO system and demonstrated that this pilot design method can also work well in increasing system performance and diminishing the pilot contamination.

Different from the previous works [11]–[14], in this paper, we bring a novel pilot design, named in-band pilots (IBPs) [15], to the CF network for further exploring the potential of the CF communication system. In particular, the pilot is imposed on the transmitted data signal in the frequency domain, rather than the traditional SP which is operated in the time domain [11]–[14]. We reveal that compared to the conventional SP strategy, the IBP method is more beneficial

to the CF network for enhancing SE and reducing pilot contamination. Our contributions are as follows:

- Firstly, we introduce a new channel model of the CF massive MIMO network, which includes the IBP design for releasing the capability of the CF system;
- Then, we derive analytical expressions for the MSE and the normalized mean square error (NMSE) of the channel estimation using the IBP in the CF network;
- Moreover, we obtain an analytical expression for the estimated data frequency for the UE;
- Finally, we numerically compare the NMSE performance of the regular pilot, the SP, and the IBP.

The rest of the paper is organized as follows. At first, Sec. II describes the system model. Then, in Sec. III, the least square (LS) channel estimation with IBP, along with the derivations of the MSE and the NMSE, are provided. Data detection with channel estimation using IBP is then described in Sec. IV. Finally, simulation results and conclusions are presented in Sec. V and Sec. VI, respectively.

II. System Model

We consider an uplink CF massive MIMO network with K single-antenna UEs and L APs, each equipped with N antennas. Both UEs and APs are arbitrarily distributed over the coverage area. We assume dynamic cooperation clustering [4] is used, i.e. each UE is served by a subset of the APs, depending on the UE's needs. Besides, orthogonal pilot signals are considered in this paper. During the uplink data (or pilot) transmission, all APs will receive a superposition of the signals sent from all UEs. Similarly, all UEs will receive signals from all other UEs. Assume C_u is the length of each data symbol (i.e., the block size), then the received signal at AP l, denoted as $\mathbf{Y}_l^{'} \in \mathbb{C}^{N\times C_u}$, can be written as,

$$\mathbf{Y}_l^{'} = \sum_{k=0}^{K-1} \sqrt{\mu_k}\mathbf{h}_{l,k}\mathbf{s}_{l,k}^{\mathsf{T}} + \mathbf{W}_l, \tag{1}$$

where $(\cdot)^{\mathsf{T}}$ is the transpose operation, μ_k denotes the transmit power for the kth user, $\mathbf{h}_{l,k} \in \mathbb{C}^{N\times 1}$ contains the channel state information (CSI) between the kth UE to the N antennas of the lth AP, $\mathbf{s}_{l,k} \in \mathbb{C}^{C_u\times 1}$ denotes the data signal for the kth user at the lth AP and $\mathbf{W}_l \in \mathbb{C}^{N\times C_u}$ is the additive white Gaussian noise (AWGN) at the lth AP with each column distributed as $\mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I_N})$ and mutually independent of the other columns. We assume $\mathbf{h}_{l,k}$ is distributed as [15]

$$\mathbf{h}_{l,k} \sim \mathcal{CN}(\mathbf{0}, \beta_{l,k}\mathbf{I}_N), \tag{2}$$

where $\beta_{l,k}$ denotes the large-scale path-loss coefficient which depends on the user location and is assumed to be known at the AP. We further assume that all elements in $\mathbf{h}_{l,k}$ are mutually independent of each other and that $\mathbf{h}_{l,k}$ is constant during the transmission of each symbol. For the received signal with time-domain superimposed pilots at AP l, denoted as $\mathbf{Y}_l \in \mathbb{C}^{N\times C_u}$, can be written as [14],

$$\mathbf{Y}_l = \sum_{k=0}^{K-1} \sqrt{\mu_k}\mathbf{h}_{l,k}\left(\rho_k\mathbf{x}_k + \lambda_k\mathbf{p}_k\right)^{\mathsf{T}} + \mathbf{W}_l, \tag{3}$$

where ρ and λ denote the scaling factor for the data and pilot signal, respectively, $\mathbf{x}_k \in \mathbb{C}^{C_u\times 1}$ and $\mathbf{p}_k \in \mathbb{C}^{C_u\times 1}$, are the data and pilot signal transmitted from the kth user.

To insert IBP pilots, we assume that the channel response is constant during the coherence interval and that a unique orthogonal pilot can be assigned for each of the K users. For the kth user, the pilots can be superimposed in the frequency domain onto its data frequencies as follows,

$$\mathbb{S}_k = \mathbf{Q}_k\mathbb{X}_k + \lambda_k\mathbb{P}_k, \tag{4}$$

where $\mathbb{X}_k$ and $\mathbb{P}_k$ represent the data and pilot frequencies, respectively, for the kth user in the lth cell, λ_k is the scaling factor for the power of the kth pilot frequency $\mathbb{P}$, and $\mathbf{Q}_k \in \mathbb{R}^{C_u\times C_u}$ is a location matrix for the allocation of pilots for the kth user[1]. $\mathbf{Q}_k$ can be defined as a square diagonal matrix with the following main diagonal elements,

$$\mathbf{q}_k = \begin{bmatrix} \gamma_k & \cdots & \gamma_k & \rho_k & \gamma_k & \cdots & \gamma_k \end{bmatrix}, \tag{5}$$

where ρ_k is the scaling factor for the power of the kth data frequency where the pilot frequency is superimposed and γ_k is the scaling factor for the remaining data frequencies. The pilot is inserted only at the k-th frequency because we assume flat fading during the coherence interval. ρ_k is used to control the power allocation at the k-th frequency and it is normally set to 0 to avoid interference from the data tone. Therefore, we assume that all the data frequencies at the non-pilot locations are scaled by the same scaling factor γ_k and that[2] [15]

$$\frac{C_u - 1}{C_u}\gamma_k^2 + \frac{1}{C_u}\rho_k^2 + \lambda_k^2 = 1. \tag{6}$$

By taking discrete Fourier transform of the received signal $\mathbf{Y}_l^{'}$ in (1), we have $\mathbf{Y}_l^{'}$ in the frequency domain $\mathbb{Y}_j \in \mathbb{C}^{N\times C_u}$, given as

$$\mathbb{Y}_l = \sum_{k=0}^{K-1} \sqrt{\mu_k}\mathbf{h}_{l,k}\mathbb{S}_k^{\mathsf{T}} + \mathbb{W}_l, \tag{7}$$

where $\mathbb{S}_{l,k}^{\mathsf{T}} \in \mathbb{C}^{1\times C_u}$ are the transmitted signal (data, pilots, or both) in the frequency domain and $\mathbb{W}_l \in \mathbb{C}^{N\times C_u}$ are the noise samples in the frequency domain.

By substituting (4) into (7), the frequency domain received signal with IBP at the lth AP can be written as,

$$\mathbb{Y}_l = \sum_{k=0}^{K-1} \sqrt{\mu_k}\mathbf{h}_{l,k}\left(\mathbf{Q}_k\mathbb{X}_k + \lambda_k\mathbb{P}_k\right)^{\mathsf{T}} + \mathbb{W}_l, \tag{8}$$

[1]Comparing with the time domain SP, the robustness of IBP results from the complete removal of interference from the data frequency at the k-th pilot frequency when ρ_k is set to 0.

[2]For example, when block size $C_u = 100$, the scaling factor for the remaining data frequencies $\gamma_k = \sqrt{1-\kappa}$ where $\kappa = 0.4$ is the ratio of each data symbol to the total transmit power [14]. Besides, $\gamma_k = \sqrt{((C_u - C_u\cdot\lambda_k^2 - \rho_k^2))/(C_u - 1)}$, and if $\rho_k = 0$, then we have $\gamma_k = 0.636$ and $\lambda_k = 0.775$.

where $\mathbb{X}_k \in \mathbb{C}^{C_u \times 1}$ represents the data frequencies and $\mathbb{P}_k \in \mathbb{C}^{C_u \times 1}$ denotes the pilot frequencies. We assume the pilots are orthogonal in the frequency domain, i.e.,

$$\mathbb{P}_k^{\mathsf{H}} \mathbb{P}_p = C_u \delta_{k,p}, \tag{9}$$

where $(\cdot)^{\mathsf{H}}$ is conjugate transpose operation, $\mathbb{P}_k^{\mathsf{T}} = [0\ 0 \cdots P_k \cdots 0]$, $P_k \in \mathbb{C}$ with a magnitude of $\sqrt{C_u}$ and $\delta_{k,p}$ is defined as

$$\delta_{k,p} = \begin{cases} 1, & \text{if } k = p \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

III. Channel Estimation using In-Band Pilots

In this section, we obtain the LS and Linear minimum mean square error (LMMSE) estimations of the channel. At first, the LS estimation of the channel for the kth user in the lth AP can be written as[3],

$$\begin{aligned} \hat{\mathbf{h}}_{l,k} &\triangleq \frac{1}{C_u \sqrt{\mu_k} \lambda_k} \mathbb{Y}_l \mathbb{P}_k^* \\ &= \frac{1}{C_u \sqrt{\mu_k} \lambda_k} \left(\sum_{m=0}^{K-1} \sqrt{\mu_m} \mathbf{h}_{l,m} \lambda_m \mathbb{P}_m^{\mathsf{T}} \mathbb{P}_k^* \right. \\ &\quad \left. + \sum_{m=0}^{K-1} \sqrt{\mu_m} \mathbf{h}_{l,m} (\mathbf{Q}_m \mathbb{X}_m)^{\mathsf{T}} \mathbb{P}_m^* + \mathbb{W}_l \mathbb{P}_m^* \right) \\ &= \mathbf{h}_{l,k} + \frac{\rho_k}{C_u \lambda_k} \chi_k^{(k)} P_k^{(k)*} \mathbf{h}_{l,k} \\ &\quad + \frac{\gamma_k P_k^{(k)*}}{C_u \lambda_k \sqrt{\mu_k}} \sum_{\substack{m=0 \\ m \neq k}}^{K-1} \sqrt{\mu_m} \chi_m^{(k)} \mathbf{h}_{l,m} + \frac{\mathbb{W}_l \mathbb{P}_k^*}{C_u \lambda_k \sqrt{\mu_k}}, \end{aligned} \tag{11}$$

where $\chi_m^{(k)}$ denotes the frequency response at the kth location of $\mathbb{X}_m \in \mathbb{C}^{C_u \times 1}$ for the mth user. $P_k^{(k)}$ denotes the pilot frequencies for the kth location of $\mathbb{P}_k$ for the kth user. From the 2nd and 3rd terms of (11), we can see that the channel is contaminated at the kth frequency from all UEs if $\rho_k \neq 0$.

Using (11), the LMMSE of $\mathbf{h}_{l,k}$ can be then found as

$$\begin{aligned} \tilde{\mathbf{h}}_{l,k} &\triangleq \mathbf{R}_{\mathbf{h}\hat{\mathbf{h}}} \boldsymbol{\Psi}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1} \hat{\mathbf{h}}_{l,k} \\ &= A_{l,k} \left(\mathbf{h}_{l,k} + \frac{\rho_k}{C_u \lambda_k} \chi_k^{(k)} P_k^{(k)*} \mathbf{h}_{l,k} + \right. \\ &\quad \left. \frac{\gamma_k P_k^{(k)*}}{C_u \lambda_k \sqrt{\mu_k}} \sum_{\substack{m=0 \\ m \neq k}}^{K-1} \sqrt{\mu_m} \chi_m^{(k)} \mathbf{h}_{l,m} + \frac{\mathbb{W}_l \mathbb{P}_k^*}{C_u \lambda_k \sqrt{\mu_k}} \right), \end{aligned} \tag{12}$$

where $\mathbf{R}_{\mathbf{h}\hat{\mathbf{h}}} = \mathbb{E}\{\hat{\mathbf{h}}_{l,k} \mathbf{h}_{l,k}^{\mathsf{H}}\}$, $\boldsymbol{\Psi}_{\hat{\mathbf{h}}\hat{\mathbf{h}}} = \mathbb{E}\{\hat{\mathbf{h}}_{l,k} \hat{\mathbf{h}}_{l,k}^{\mathsf{H}}\}$ and

$$\begin{aligned} A_{l,k} &\triangleq \mathbf{R}_{\mathbf{h}\hat{\mathbf{h}}} \boldsymbol{\Psi}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1} \\ &= \frac{\beta_{l,k}}{\beta_{l,k} + \frac{\rho_k^2}{C_u \lambda_k^2} \beta_{l,k} + \frac{\gamma_k^2}{C_u \lambda_k^2} \sum_{\substack{m=0 \\ m \neq k}}^{K-1} \mu_m \beta_{l,m} + \frac{\sigma^2}{C_u \lambda_k^2 \mu_k}}. \end{aligned} \tag{13}$$

[3]It should be emphasized that since $\rho_k = 0$, then the second term of the right hand of (11) vanishes. For SP, since the time domain signal is spread over the data signal over the entire time slot, it is impossible to completely remove the interference from the data signal itself. This is the main advantage of using IBP over SP.

Using (12), the MSE of the LMMSE can be calculated as

$$\begin{aligned} \mathsf{MSE}_{l,k} &\triangleq \frac{1}{N} \mathbb{E}\left\{ \|\Delta \mathbf{h}_{l,k}\|^2 \right\} \\ &= \frac{1}{N} \mathbb{E}\left\{ (\tilde{\mathbf{h}}_{l,k} - \mathbf{h}_{l,k})^{\mathsf{H}} (\tilde{\mathbf{h}}_{l,k} - \mathbf{h}_{l,k}) \right\} \\ &= \frac{1}{N} \mathbb{E}\left\{ \tilde{\mathbf{h}}_{l,k}^{\mathsf{H}} \tilde{\mathbf{h}}_{l,k} - \tilde{\mathbf{h}}_{l,k}^{\mathsf{H}} \mathbf{h}_{l,k} - \mathbf{h}_{l,k}^{\mathsf{H}} \tilde{\mathbf{h}}_{l,k} + \mathbf{h}_{l,k}^{\mathsf{H}} \mathbf{h}_{l,k} \right\} \\ &= A_{l,k}^2 \left(\beta_{l,k} + \frac{\rho_k^2}{C_u \lambda_k^2} \beta_{l,k} + \frac{\gamma_k^2}{C_u \lambda_k^2} \sum_{\substack{m=0 \\ m \neq k}}^{K-1} \mu_m \beta_{l,m} \right. \\ &\quad \left. + \frac{\sigma^2}{\lambda_{j,m}^2 C_u} \right) - 2 A_{l,k} \left(\beta_{l,k} + \frac{\rho_k}{\sqrt{C_u} \lambda_k} \beta_{l,k} \right) + \beta_{l,k} \\ &= A_{l,k} \beta_{l,k} - 2 A_{l,k} \beta_{l,k} - 2 A_{l,k} \frac{\rho_k}{\sqrt{C_u} \lambda_k} \beta_{l,k} + \beta_{l,k} \\ &= \left(1 - A_{l,k} - 2 A_{l,k} \frac{\rho_k}{\sqrt{C_u} \lambda_k} \right) \beta_{l,k} \\ &= \beta_{l,k} - A_{l,k} \left(1 + \frac{2\rho_k}{\sqrt{C_u} \lambda_k} \right) \beta_{l,k}, \end{aligned} \tag{14}$$

where

$$\Delta \mathbf{h}_{l,k} \triangleq \tilde{\mathbf{h}}_{l,k} - \mathbf{h}_{l,k}. \tag{15}$$

From [14], the normalized mean square error (NMSE) for IBP is given as

$$\begin{aligned} \mathsf{NMSE} &= \frac{\sum_{l,k} \mathbb{E}\left\{ \|\Delta \mathbf{h}_{l,k}\|^2 \right\}}{\sum_{l,k} \mathbb{E}\left\{ \|\mathbf{h}_{l,k}\|^2 \right\}} \\ &= \frac{\sum_{l,k} \beta_{l,k} - A_{l,k} \left(1 + \frac{2\rho_k}{\sqrt{C_u} \lambda_k} \right) \beta_{l,k}}{\sum_{l,k} \beta_{l,k}}. \end{aligned} \tag{16}$$

Note that when $\rho_k = 0$ and $\mu_m = 1, \forall m$, the NMSE for IBP, denoted as NMSE_0, is given as

$$\mathsf{NMSE}_0 = \frac{\sum_{l,k} \beta_{l,k} - A'_{l,k} \beta_{l,k}}{\sum_{l,k} \beta_{l,k}}, \tag{17}$$

where $A'_{l,k}$ is given as

$$A'_{l,k} = \frac{\beta_{l,k}}{\beta_{l,k} + \frac{\gamma_k^2}{C_u \lambda_k^2} \sum_{\substack{m=0 \\ m \neq k}}^{K-1} \beta_{l,m} + \frac{\sigma^2}{C_u \lambda_k^2}}. \tag{18}$$

From (17), we see that the larger the block size C_u and the larger of the scaling factor at the kth location λ_k is, the smaller the channel estimation error.

IV. Data Detection with In-band Pilots

We assume that maximum ratio combining is used for data detection, which is conducted at the central server after receiving the projection of the frequency domain received signal onto the estimated CSIs from all APs via the fronthaul

links. The estimated data frequency for the kth user, denoted as $\tilde{X}_k^T$, can be calculated as

$$\begin{aligned}\tilde{X}_k^{\mathsf{T}} &= \frac{1}{N}\sum_{l=0}^{L-1}\tilde{\mathbf{h}}_{l,k}^{\mathsf{H}}\mathbb{Y}_l \\ &= \frac{1}{N}\sum_{l=0}^{L-1}\tilde{\mathbf{h}}_{l,k}^{\mathsf{H}}\left(\sum_{m=0}^{K-1}\sqrt{\mu_m}\mathbf{h}_{l,m}\left(\mathbf{Q}_m\mathbb{X}_m+\lambda_m\mathbb{P}_m\right)^{\mathsf{T}}+\mathbb{W}_l\right) \\ &= \frac{1}{N}\sum_{l=0}^{L-1}\tilde{\mathbf{h}}_{l,k}^{\mathsf{H}}\sum_{m=0}^{K-1}\sqrt{\mu_m}\mathbf{h}_{l,m}(\mathbf{Q}_m\mathbb{X}_m)^{\mathsf{T}} \\ &+ \frac{1}{N}\sum_{l=0}^{L-1}\tilde{\mathbf{h}}_{l,k}^{\mathsf{H}}\sum_{m=0}^{K-1}\sqrt{\mu_m}\lambda_m\mathbf{h}_{l,m}\mathbb{P}_m^{\mathsf{T}}+\frac{1}{N}\sum_{l=0}^{L-1}\tilde{\mathbf{h}}_{l,k}^{\mathsf{H}}\mathbb{W}_l.\end{aligned} \tag{19}$$

Using (15), (19) can be written as

$$\begin{aligned}\tilde{X}_k^{\mathsf{T}} &= \frac{1}{N}\sum_{l=0}^{L-1}\sum_{m=0}^{K-1}\sqrt{\mu_m}\mathbf{h}_{l,k}^{\mathsf{H}}\mathbf{h}_{l,m}(\mathbf{Q}_m\mathbb{X}_m)^{\mathsf{T}} \\ &+ \frac{1}{N}\sum_{l=0}^{L-1}\sum_{m=0}^{K-1}\sqrt{\mu_m}\Delta\mathbf{h}_{l,k}^{\mathsf{H}}\mathbf{h}_{l,m}(\mathbf{Q}_m\mathbb{X}_m)^{\mathsf{T}} \\ &+ \frac{1}{N}\sum_{l=0}^{L-1}\sum_{m=0}^{K-1}\lambda_m\sqrt{\mu_m}\mathbf{h}_{l,k}^{\mathsf{H}}\mathbf{h}_{l,m}\mathbb{P}_m^{\mathsf{T}} \\ &+ \frac{1}{N}\sum_{l=0}^{L-1}\sum_{m=0}^{K-1}\lambda_m\sqrt{\mu_m}\Delta\mathbf{h}_{l,k}^{\mathsf{H}}\mathbf{h}_{l,m}\mathbb{P}_m^{\mathsf{T}} \\ &+ \frac{1}{N}\sum_{l=0}^{L-1}\mathbf{h}_{l,k}^{\mathsf{H}}\mathbb{W}_l+\frac{1}{N}\sum_{l=0}^{L-1}\Delta\mathbf{h}_{l,k}^{\mathsf{H}}\mathbb{W}_l.\end{aligned} \tag{20}$$

Assuming that the zero-mean and mutually independent channel vectors are asymptotically orthogonal when the number of antennas N grows without bound, we have [11],

$$\lim_{N\to\infty}\frac{\mathbf{h}_{l,k}^H\mathbf{h}_{q,r}}{N}=\beta_{l,k}\delta_{l,q}\delta_{k,r},\forall l,k,q,r, \tag{21}$$

where $\beta_{l,k}$ denotes the large-scale path-loss coefficient for the kth user to the l AP. Using (21), (20) can be rewritten as

$$\begin{aligned}\tilde{X}_k^{\mathsf{T}} &= \sqrt{\mu_k}(\mathbf{Q}_k\mathbb{X}_k)^{\mathsf{T}}\sum_{l=0}^{L-1}\beta_{l,k}+\lambda_k\sqrt{\mu_k}\mathbb{P}_k^{\mathsf{T}}\sum_{l=0}^{L-1}\beta_{l,k} \\ &+ \frac{1}{N}\sum_{l=0}^{L-1}\sum_{m=0}^{K-1}\sqrt{\mu_m}\Delta\mathbf{h}_{l,k}^{\mathsf{H}}\mathbf{h}_{l,m}(\mathbf{Q}_m\mathbb{X}_m)^{\mathsf{T}} \\ &+ \frac{1}{N}\sum_{l=0}^{L-1}\sum_{m=0}^{K-1}\lambda_m\sqrt{\mu_m}\Delta\mathbf{h}_{l,k}^{\mathsf{H}}\mathbf{h}_{l,m}\mathbb{P}_m^{\mathsf{T}} \\ &+ \frac{1}{N}\sum_{l=0}^{L-1}\mathbf{h}_{l,k}^{\mathsf{H}}\mathbb{W}_l+\frac{1}{N}\sum_{l=0}^{L-1}\Delta\mathbf{h}_{l,k}^{\mathsf{H}}\mathbb{W}_l.\end{aligned} \tag{22}$$

To remove the contribution of the pilots and the scaling factor to the estimated data frequencies, (22) can be further simplified as[4]

[4]Using (23), the SINR for SE calculation can be calculated by treating the first term as the data signal term and the remaining four terms as the interference. This will be investigated in future works.

$$\begin{aligned}\tilde{\mathbb{X}}_k^{\mathsf{T}} &= \frac{1}{B_k}\left(\tilde{X}_k^{\mathsf{T}}-\lambda_k B_k\mathbb{P}_k^{\mathsf{T}}\right)\mathbf{Q}_k^{-1} \\ &= \mathbb{X}_k^{\mathsf{T}}+\frac{1}{B_kN}\sum_{l=0}^{L-1}\sum_{m=0}^{K-1}\sqrt{\mu_m}\Delta\mathbf{h}_{l,k}^{\mathsf{H}}\mathbf{h}_{l,m}(\mathbf{Q}_m\mathbb{X}_m)^{\mathsf{T}}\mathbf{Q}_k^{-1} \\ &+ \frac{1}{NB_k}\sum_{l=0}^{L-1}\sum_{m=0}^{K-1}\lambda_m\sqrt{\mu_m}\Delta\mathbf{h}_{l,k}^{\mathsf{H}}\mathbf{h}_{l,m}\mathbb{P}_m^{\mathsf{T}}\mathbf{Q}_k^{-1} \\ &+ \frac{1}{NB_k}\sum_{l=0}^{L-1}\mathbf{h}_{l,k}^{\mathsf{H}}\mathbb{W}_l\mathbf{Q}_k^{-1}+\frac{1}{NB_k}\sum_{l=0}^{L-1}\Delta\mathbf{h}_{l,k}^{\mathsf{H}}\mathbb{W}_l\mathbf{Q}_k^{-1}\end{aligned} \tag{23}$$

where $B_k=\sqrt{\mu_k}\sum_{l=0}^{L-1}\beta_{l,k}$. By taking the IDFT of $\tilde{\mathbb{X}}_k^{\mathsf{T}}$, we can obtain the estimated data symbols.

V. Simulation Results

In this section, numerical simulations are offered to validate the analytical results in previous sections. A square area of D m $\times$ D m is considered as the main cell [14] and the wrap-around technique [4] is applied to mimic a large network deployment without edges. Unless otherwise stated, we assume that all APs are deployed uniformly at random in the coverage area [4]. The large-scale fading coefficient $\beta_{l,k}$ can be expressed as [5], [14]

$$\beta_{l,k}\triangleq\begin{cases}-L-35\log_{10}(d_{l,k})+\sigma_1, & \text{if } d_{l,k}>d_1\\ -L-15\log_{10}(d_1)-20\log_{10}(d_{l,k})+\sigma_2, & \\ \qquad\qquad\text{if } d_0<d_{l,k}\le d_1 & \\ -L-15\log_{10}(d_1)-20\log_{10}(d_0), & \text{if } d_{l,k}\le d_0,\end{cases}$$

where $d_{l,k}$ is the distance between the l-th AP and the k-th UE, σ_1 and σ_2 are shadow fading factors, d_0 and d_1 are reference distances, and L can be defined as [5], [14]

$$\begin{aligned}L\triangleq&46.3+33.9\log_{10}(f)-13.82\log_{10}(h_{\text{AP}})\\ &-(1.1\log_{10}(f)-0.7)h_{\text{UE}}+(1.56\log_{10}(f)-0.8),\end{aligned}$$

where f is the carrier frequency (in GHz), h_{AP} and h_{UE} are the antenna heights (in m) of the AP and the UE, respectively. We ignore other noise figures in this paper and have no power control scheme. Default setup parameters in this section can be found in Table I unless otherwise specified.

In order to show the effectiveness of the IBP scheme, we provide two baselines for comparison, i.e., the regular pilot (RP) strategy[5] [5] and the SP strategy [3], [11], [12]. Note that the same as [14], we assume the length of pilot training for the RP is 20% of the SP's, and the pilot symbol has 40% transmit power, i.e., $\kappa=0.4$. Besides, due to space limitations, numerical evaluations for data detection in Sec. IV are left open for future works.

TABLE I
SIMULATION PARAMETERS [4], [5], [14]

Parameters	Values
Carrier frequency (f)	1.9 GHz
Bandwidth	20 MHz
AP height (h_{AP})	15 m
UE height (h_{UE})	1.65 m
D, d_1, d_0	1000, 50, 10 m
AP number (L)	100
UE number (K)	240
Antenna number in each AP (N)	4
Block size (C_u)	100
Transmit power (μ_k)	0 dBm
AWGN noise power	-80 dBm
Shadow fading factors (σ_1 and σ_2)	3 dB, 8 dB
Data symbol power ratio (κ)	0.4
Realization number	1000

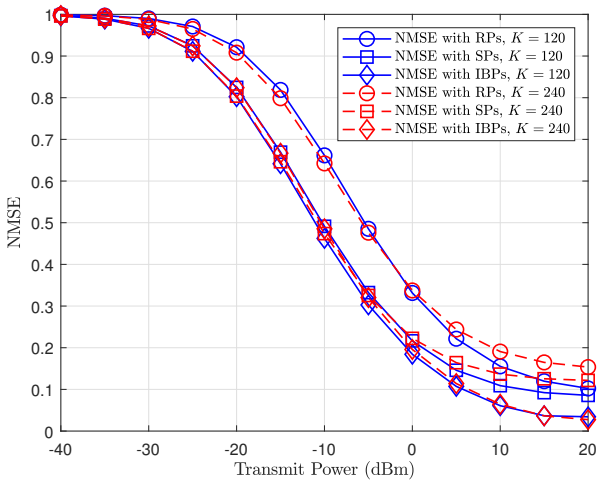

Fig. 1. NMSE vs. transmit power against different numbers of UEs K.

A. NMSE vs. Transmit Power

Fig. 1. illustrates the NMSEs by using the RP, the SP, and the IBP schemes, for the transmit power ranging from -40 dBm to 20 dBm against different numbers of UEs K. Note that the lines for the IBP scheme are generated according to Eq. (16). All NMSEs first decrease and then saturate to specific asymptotic bounds when the transmit power continues to increase, as we expected. In particular, when the transmit power is -40 dBm, all schemes lose their effectiveness, but as the power reaches -30 dBm, the SP, and the IBP are beneficial compared to the conventional RP scheme, owing to the fact that the pilot contamination is mitigated with the SP and the IBP schemes. When the transmit power approaches 0 dBm, the IBP scheme performs the best to the other two baselines since it operates the pilot in the frequency domains. This reveals that the IBP scheme works well in the high transmit power region. Besides, it is noteworthy that, interestingly, the IBP scheme performance is not sensitive to the number of UEs, K. However, when K = 120., the two baselines perform better than with K = 240. In summary, although the SP and the IBP schemes sacrifice part of the uplink transmit power, they perform better than the RP scheme. More importantly, because the superimposed process is operated in the frequency domain, the IBP scheme is more effective than the SP scheme, i.e., the IBP provides a lower NMSE.

[5]RP means pilots are sent separately from the data to avoid pilot contamination if orthogonal pilots are used. Compared with the SP, the RP has lower SE because it occupies time/frequency slots.

B. NMSE vs. Block Size

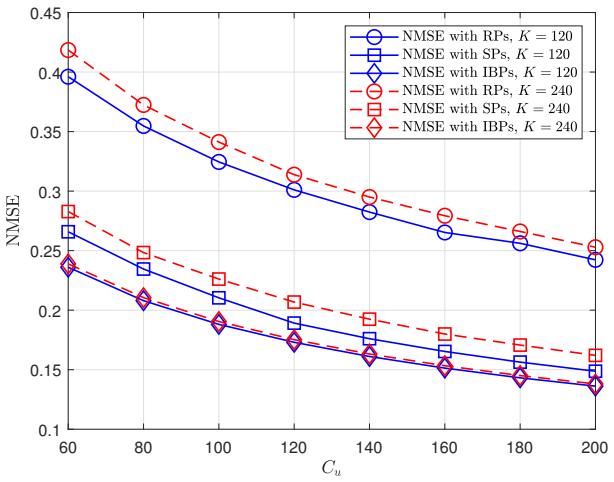

Fig. 2. NMSE vs. block sizes C_u against different numbers of UEs K.

Fig. 2. demonstrates the NMSEs by using the RP, the SP, and the IBP schemes, for the block size C_u ranging from 60 to 200 against different numbers of UEs K. Clearly, a larger C_u provides more accurate results for all three schemes. Thus, the NMSEs all decrease when the block size increases. Besides, as we expected, the IBP performs the best. This is because it manages the superimposed pilots more effectively. Once again, the IBP lines for K = 120 and K = 240 almost overlap, which implies that the IBP scheme is not sensitive to the number of UEs when we consider the uplink transmission.

C. NMSE vs. Pilot Scaling Factor

Fig. 3. depicts the NMSE vs. pilot scaling factor λ in Eq. (6) for different K. It can be easily observed that increasing pilot length is always favorable to the SP and the IBP schemes but has no impact on the RP scheme. Interestingly, when λ expands, the gap between K = 120 and K = 240 of the SP scheme becomes narrow. However, the IBP scheme keeps the most promising results both in K = 120 and K = 240. This result reveals that the SP scheme performance more relies on λ. Thus, when the scenario has more users and a small λ, the IBP scheme is more suitable.

Finally, in Fig. 4., suppose $\kappa \in [0.1, 0.9]$, we plot the three-dimensional illustration for NMSE vs. γ vs. λ to clarify the relationships among them. According to Eq. (6), if $\rho = 0$, then the scaling factors should be satisfied with $\frac{C_u-1}{C_u}\gamma^2 + \lambda^2 = 1$. In other words, if λ increases, then γ decreases and vice versa.

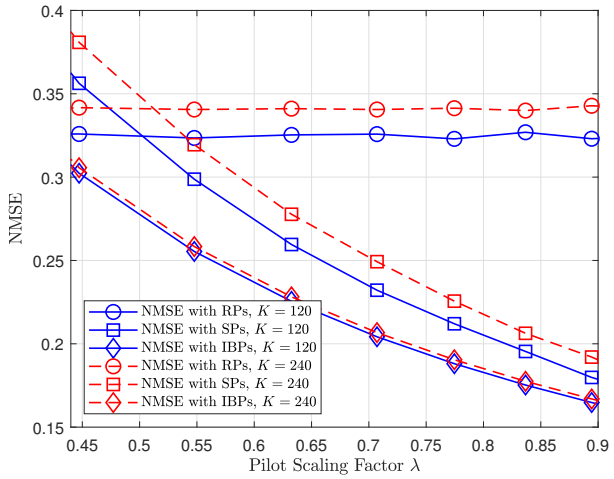

Fig. 3. NMSE vs. pilot scaling factor λ against different numbers of UEs K.

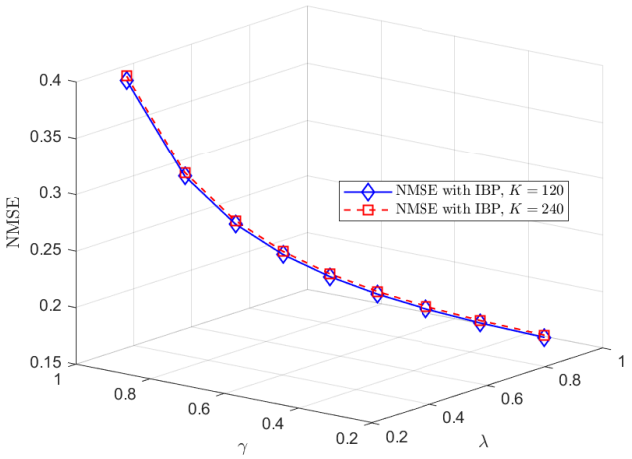

Fig. 4. NMSE vs. λ vs. γ against different numbers of UEs K. Note that in this case, $\kappa \in [0.1, 0.9]$.

As the figure shows, a lower γ and a higher λ performs better. However, the achievable rate may decrease if we enlarge the λ. Besides, the lines for $K = 120$ and $K = 240$ are almost the same, which indicates the robustness of the IBP scheme.

VI. Conclusion

A channel estimation approach using the IBP scheme has been proposed in this paper. The proposed strategy works in the frequency domain, as opposed to other conventional approaches that function in the time domain. Analytical and simulated results have confirmed that the IBP scheme is more effective and robust in enhancing performance and reducing pilot contamination, even when the number of UE increases. It is worth noting that in orthogonal frequency-division multiplexing (OFDM) systems [16], pilots are inserted in both the time and frequency domains. However, it occupies radio resources, which leads to a lower SE. Moreover, if the SP is to be used in the OFDM, one cannot completely remove the subcarrier for accommodating the pilot frequencies.

For future works, achievable SE analysis and power control designs for the IBP scheme are worth investigating. Besides, reconfigurable intelligent surfaces (RISs) - aided CF massive communication systems have attracted attention recently since the RIS can improve the system's coverage, increase the minimum rate, and boost the per-user throughput [17]. Consequently, the application of the IBP scheme in the RIS-assisted CF massive network is also very promising. Moreover, channel codes, such as low density parity check codes, can be incorporated into the proposed scheme to reveal the potential of practical implementations.

References

[1] H. He, X. Yu, J. Zhang, S. Song, and K. B. Letaief, "Cell-free massive mimo for 6g wireless communication networks," *J. Commun. Infor. Netw*, vol. 6, no. 4, pp. 321–335, 2021.

[2] S. Elhoushy, M. Ibrahim, and W. Hamouda, "Cell-free massive mimo: A survey," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 492–523, 2021.

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

[4] Ö. T. Demir, E. Björnson, L. Sanguinetti *et al.*, "Foundations of user-centric cell-free massive mimo," *Foundations and Trends® in Signal Processing*, vol. 14, no. 3-4, pp. 162–472, 2021.

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

[6] E. Björnson and L. Sanguinetti, "Making cell-free massive mimo competitive with mmse processing and centralized implementation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 77–90, 2019.

[7] T. Kim, H. Kim, S. Choi, and D. Hong, "How will cell-free systems be deployed?" *IEEE Commun. Mag.*, vol. 60, no. 4, pp. 46–51, 2022.

[8] T. M. Hoang, H. Q. Ngo, T. Q. Duong, H. D. Tuan, and A. Marshall, "Cell-free massive mimo networks: Optimal power control against active eavesdropping," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4724–4737, 2018.

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

[10] W. Jiang and H. D. Schotten, "Opportunistic ap selection in cell-free massive mimo-ofdm systems," in *2022 IEEE 95th Vehicular Technology Conference:(VTC2022-Spring)*. IEEE, 2022, pp. 1–5.

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

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

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

[14] Y. Zhang, X. Qiao, L. Yang, and H. Zhu, "Superimposed pilots are beneficial for mitigating pilot contamination in cell-free massive mimo," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 279–283, 2021.

[15] C.-T. Lam, B. Ng, K. Wang, and E. Law, "Channel estimation using in-band pilots for uplink massive mimo," in *2021 IEEE 21st International Conference on Communication Technology (ICCT)*. IEEE, 2021, pp. 268–273.

[16] X. Lin, J. Li, R. Baldemair, J.-F. T. Cheng, S. Parkvall, D. C. Larsson, H. Koorapaty, M. Frenne, S. Falahati, A. Grovlen, and K. Werner, "5g new radio: Unveiling the essentials of the next generation wireless access technology," *IEEE Commu. Stand. Mag.*, vol. 3, no. 3, pp. 30–37, 2019.

[17] X. Ma, D. Zhang, M. Xiao, C. Huang, and Z. Chen, "Cooperative beamforming for ris-aided cell-free massive mimo networks," *IEEE Trans. Wireless Commun.*, 2023.